

Randomized Trial of Intensive Early Intervention for Children With Pervasive Developmental Disorder

Tristram Smith

Washington State University

Annette D. Groen

Metropolitan State College of Denver

Jacqueline W. Wynn

University of California, Los Angeles

Young children with pervasive developmental disorder were randomly assigned to intensive treatment or parent training. The intensive treatment group (7 with autism, 8 with pervasive developmental disorder not otherwise specified—NOS) averaged 24.52 hours per week of individual treatment for one year, gradually reducing hours over the next 1 to 2 years. The parent training group (7 with autism, 6 with pervasive developmental disorder NOS) received 3 to 9 months of parent training. The groups appeared similar at intake on all measures; however, at follow-up the intensive treatment group outperformed the parent training group on measures of intelligence, visual-spatial skills, language, and academics, though not adaptive functioning or behavior problems. Children with pervasive developmental disorder NOS may have gained more than those with autism.

After years of debating whether or not early intervention helps children with developmental delays (Weinberg, 1989), researchers have largely come to agree on a middle ground: Early intervention is beneficial for many children, but gains tend to be limited (e.g., Scarr & Arnett, 1987). For example, many researchers in the area of early intervention for children with developmental disabilities have shown that such intervention prevents declines in intellectual development and may reduce family stress (e.g., Guralnick, 1998), though children continue to display substantial delays.

Nevertheless, there have been reports of larger improvements. Of particular interest, in peer-reviewed studies, seven independent groups of investigators have described dramatic gains with early intervention for children with autism (reviewed by Smith, 1999). In all studies, interventions were based on applied behavior analytic research and theory (Green, 1996) and were intensive (15 to 40 hours per week). Reported gains have included average increases of approximately 20 points in IQ (Harris, Handleman, Gordon, Kristoff, & Fuentes, 1991; Lovaas, 1987; Sheinkopf & Siegal, 1998) and other stan-

dardized test scores (Anderson, Avery, DiPietro, Edwards, & Christian, 1987; Birnbrauer & Leach, 1993; Hoyson, Jamison, & Strain, 1984; McEachin, Smith, & Lovaas, 1993), as well as less restrictive school placements (Fenske, Zalski, Krantz, & McClannahan, 1985; Lovaas, 1987).

Such results may not only enhance the outlook for children with autism but also raise optimism about the extent to which children with other developmental disorders may benefit from early intervention (Guralnick, 1998). However, the validity of the results has been a topic of intense debate. A study by Lovaas and colleagues (Lovaas, 1987; McEachin, Smith, & Lovaas, 1993) has garnered particular attention. Lovaas evaluated three groups of children with autism who were under 4 years old at intake. The experimental group ($n = 19$) received intensive treatment, which consisted of 40 hours per week of one-to-one, in-home, applied behavior analytic intervention for 2 or more years. One control group ($n = 19$) received minimal treatment (10 hours per week or less); a second control group ($n = 21$) was treated at other agencies and had no contact with Lovaas's clinic. Though the three groups did not appear to differ at intake, the intensively treated children substantially outperformed the children in control groups at age 7. Their mean IQ was 83 compared to 52 and 58, respectively. Also, 9 of 19 received passing grades without special assistance in classes for typically developing children compared to only 1 of 40 in the control groups. Moreover, at a follow-up conducted when the children averaged 12 years of age, the intensively treated children maintained their gains and also functioned more satisfactorily than did minimally treated children on measures of adaptive behavior and personality (McEachin et al., 1993).

McEachin et al. (1993) identified a number of strengths of this study, including (a) experimental and control groups that did not differ on 19 of 20 intake variables, (b) intake and follow-up evalu-

ations conducted by blind examiners independent of the study, (c) reliance on treatment approaches developed from extensive research on reducing maladaptive behaviors and enhancing skills in children with autism (cf. Newsom & Rincover, 1989), (d) use of a detailed treatment manual (Lovaas et al., 1981) and associated videotapes to standardize the interventions that children received, and (e) follow-ups to assess maintenance of treatment gains conducted many years after termination of treatment. However, others identified many possible flaws, notably the following (Gresham & MacMillan, 1997; Schopler, Short, & Mesibov, 1989): First, assignment to groups was based on whether or not therapists were available to provide intensive treatment rather than on a more arbitrary procedure, such as the use of a random numbers table. Thus, assignment could have been biased. Second, because children were referred to outside examiners, they received a variety of different tests rather than a uniform assessment protocol. Hence, assessment results may have been unreliable. Third, selection criteria such as IQ cut-offs may have been unduly restrictive, yielding a sample with an unusually favorable prognosis. Fourth, the large amount of treatment and the level of expertise required for proper implementation may have been too much for other professionals to duplicate, too stressful for most children and families to tolerate, and too costly for funding agencies to support. Lovaas and colleagues concurred with the first two of these criticisms, though doubting the importance of the second. They disputed the other criticisms but emphasized the need for replication to confirm the results (Lovaas, Smith, & McEachin, 1989; Smith & Lovaas, 1997; Smith, McEachin, & Lovaas, 1993).

Anderson et al. (1987), Birnbrauer and Leach (1993), and Sheinkopf and Siegal (1998) conducted partial replications of the study by Lovaas and colleagues. Children in these studies received fewer hours of treatment (18 to 25 hours

per week vs. 40 hours) from less experienced personnel than in the Lovaas study. All studies showed substantial average increases in nonverbal IQ (22 to 29 points), but gains in other areas were smaller than those reported by Lovaas (1987).

The present study was designed to extend this literature. Children received early intervention based on the same treatment manual used by Lovaas (1987), implemented by personnel who met the qualifications specified in that study and were independent of Lovaas (1987). However, because of concerns about cost of service delivery and stress on children and families, intervention was made less intensive than that in the Lovaas study, as described later in the *Treatment* section. To address criticisms of previous research and increase methodological rigor, we conducted a fully randomized clinical trial with uniform, comprehensive assessment protocols for all participants. To evaluate treatment efficacy for a wider range of children, we studied not only children with autism but also children with pervasive developmental disorder not otherwise specified (NOS). Because such children are often viewed as having "mild autism" (Towbin, 1997), they were hypothesized to be appropriate candidates for the intervention we provided.

Method

Participants

All referrals to the UCLA Young Autism Project between 1989 and 1992 who met the following criteria were enrolled in the study: (a) chronological age (CA) between 18 and 42 months at the time of referral, (b) residence within a one-hour drive of the research/treatment site (the UCLA Young Autism Project), (c) IQ ratio between 35 and 75, (d) diagnosis of autism or pervasive developmental disorder NOS, and (e) absence of major medical problems other than autism or mental retardation (e.g., cerebral

palsy, blindness or deafness, known genetic disorders such as Down syndrome, or neurological conditions such as uncontrolled seizure disorders). Twenty-eight children met these criteria and participated in the study, including 14 diagnosed with autism and 14 diagnosed with pervasive developmental disorder NOS. There were no dropouts among this group of children. However, one child's family declined participation at intake, and 8 other children were excluded (4 because they did not have a diagnosis of autism or pervasive developmental disorder NOS; 2 because they scored below the IQ cutoff; and 2 because they were in foster care, without a permanent residence in which to provide the home-based services offered to children in this study).

Diagnosis for all 28 participants was made independently of the study by licensed psychologists at the California State Regional Centers (a state agency that coordinates services for individuals with developmental disabilities). Nineteen participants had also received a second, independent diagnosis prior to entry into the study (8 from the UCLA Neuropsychiatric Institute; 3 from White Memorial Hospital; 3 from former clinic supervisors at the UCLA Young Autism Project who had become licensed, doctoral psychologists and were blind to the children's previous diagnostic history and independent of the study; 1 from the University of Southern California Medical Center; 1 from Children's Hospital; and 1 from Cedars-Sinai Hospital). The second diagnosis was identical to the Regional Center diagnosis for all participants except one, who was diagnosed with autism at the UCLA Neuropsychiatric Institute but pervasive developmental disorder NOS at the Regional Center (and, hence, was classified as having pervasive developmental disorder NOS for the purposes of this study).

Table 1 summarizes background information on all participants in the two groups (intensive treatment and parent training, described in *Treatment*), as re-

ported by each child's primary caregiver on the Family Background Questionnaire (Siegal & Elliott, 1988). The groups appeared similar on all variables. Participants had diverse ethnic and socioeconomic backgrounds, consistent with the general population in the Los Angeles area. In addition, they resembled other populations of children with pervasive developmental disorder in terms of sex ratio (Smith, 1997) and frequency of medical conditions (Rutter, Bailey, Bolton, & LeCouteur, 1994).

Design

Children were assigned to intensive treatment or parent training based on the following matched-pair, random assignment procedure: Once intake assessments had been completed on 4 to 8 children, they were divided into two cohorts, those with a diagnosis of autism and those with a diagnosis of pervasive developmental disorder. Their first names and Bayley IQs were then given to an independent statistician, who had no other information about the children. The statistician paired the children in each cohort based on IQ (the two highest forming one pair,

the next two forming another pair, etc.). Finally, using a random numbers table, he assigned one member of each pair to the intensive treatment group and the other to the parent training group. The unequal size of these groups resulted from a design change. We originally intended to study not only children with autism or pervasive developmental disorder NOS but also children who had mental retardation without pervasive developmental disorder. Hence, for subject assignment we divided the children into two cohorts: those with autism and those with other diagnoses. We then assigned them to groups as just described. Because our sample included only 4 children with mental retardation and no pervasive developmental disorder [1 assigned to intensive treatment and 3 to parent training], we could not draw reliable inferences about these children's responsivity to treatment and, hence, removed them from the study after they had completed their participation. Intake and follow-up data on the children were obtained and are available from the first author. The data do not alter the results of the significance testing presented later in the *Results* section.

Table 1
Background Information by Group

Characteristic	Intensive treatment (<i>n</i> = 15)	Parent training (<i>n</i> = 13)
Intake CA (in months) ^a	36.07 (6.00)	35.77 (5.37)
Follow-up CA (in months) ^a	94.07 (13.17)	92.23 (17.24)
Diagnosis ^b	7/8	7/6
Boy: girl ratio	12:3	11:2
Ethnicity		
White	7	7
Hispanic	4	2
Black	1	3
Asian	3	1
Single parent household (<i>n</i>)	4	4
Household income ^c	\$40-50,000 (<\$10,000 to \$75-100,000)	\$40-50,000 (<\$10,000 to \$75-100,000)
Years of schooling ^c		
Mother	12 (10-16+)	15 (12-16+)
Father	13-14 (<6-16+)	15 (12-16+)
Siblings ^c	1 (0-4)	1 (0-2)
Medical conditions ^d	1	1
Motor delays ^e	2	1

^aMean (*SD*). ^bAutism/pervasive developmental disorder NOS (not otherwise specified). ^cMedian (range). ^dDefined as any prenatal, perinatal, or neurological condition that resulted in medical treatment or any other medical condition that resulted in hospitalization. Medical conditions in the intensive group: one child had a skull fracture at 4 months of age; in the parent training group, one child had tubercular meningitis at 15 months of age. ^eDefined as sitting independently after 8 months of age or walking independently after 16 months of age.

Treatment

The authors, who had a combined total of 10 years experience at the UCLA Young Autism Project under Lovaas's supervision prior to the outset of the study, directed treatment. While the study was ongoing, we received no input from Lovaas except quarterly, one-hour consultations pertaining to the UCLA Young Autism Project administrative issues and, hence, we were essentially independent. In both intensive treatment and parent training, children received intervention based on Lovaas et al.'s (1981) manual. The treatment progressed gradually and systematically from relatively simple tasks, such as responding to basic requests made by an adult, to more complex skills, such as conversing and making friends with peers. The implementation of empirically supported teaching approaches developed by behavior analysts was emphasized (Newsom & Rincover, 1989).

Intensive treatment group. The goal of intensive treatment was to maximize children's intellectual, adaptive, and socioemotional functioning and, thereby, reduce their subsequent need for special education services. For our purposes in the present study, *intensive treatment* was defined as 30 hours per week of intervention for each child for 2 to 3 years. Children received treatment from teams of 4 to 6 student therapists working under our close supervision. Each child's primary caregiver was asked to conduct 5 hours per week of treatment, working alongside a student therapist, for the first 3 months of treatment.

At the outset of treatment, student therapists relied primarily on a one-to-one, discrete trial format, implemented in children's homes, so that they could provide highly individualized interventions that maximized children's successes. At this stage, the children were not yet attending school. Once they spoke in short phrases; cooperated with verbal requests from others; played appropriately with toys; and acquired self-care skills, such as dressing and toileting (ap-

proximately one year after treatment onset, with large variation across children), the focus shifted to more naturalistic instruction in group settings, such as classrooms. They were entered into classrooms in the public schools for typically developing children. Student therapists served as aides to assist the children in adjusting to classrooms, and hours in one-to-one treatment gradually decreased as children progressed at school. If children did not master the aforementioned skills after 18 months of treatment, they were enrolled in special education classrooms, based on the Individualized Education Plan developed by school personnel. If requested by parents or school personnel, information on children's present level of performance and recommended educational goals was provided by staff from the UCLA Young Autism Project, but staff did not participate in decisions about classroom placement (see McEachin et al., 1993, and Smith, Donahoe, and Davis, in press, for additional details on transitioning to school and terminating services from the UCLA Young Autism Project).

The treatment in the present study was less intensive than that provided by Lovaas (1987) in four respects: First, the aim was to provide 30 rather than 40 hours per week of intervention, as done by Lovaas (1987). Second, treatment was phased out after 18 months for children who were progressing slowly (as defined in the preceding paragraph), rather than continuing for up to 10 years, as Lovaas (1987) did. Third, fewer requirements were placed on parents, who, in the Lovaas (1987) study, were asked to take a year off from their jobs to participate in their children's treatment. Finally, though contingent aversives were employed briefly with the first 4 children, they were then stopped for all children, whereas Lovaas (1987) employed this intervention more extensively.

Parent training group. The goal of parent training was to teach parents to use treatment approaches described in the Lovaas et al. (1981) manual and assist

them in using these approaches to help their children acquire skills. The children's families received two sessions per week of parent training, totaling 5 hours per week, in their homes for 3 to 9 months. Parent trainers had a minimum of one year of experience at the UCLA Young Autism Project, including 6 or more months instructing novice therapists. They received one hour per week of individual supervision from the first author, with additional supervision available as needed. The first author also met with the parents at the first training session, every 3 months thereafter, and at the termination of parent training. An emphasis was placed on collaboration between parents and parent trainers, with parents selecting goals for their children and parent trainers demonstrating ways to work toward these goals. In the first session of parent training, parents were asked to identify three skills they wanted their children to achieve. The first author and parent trainer then outlined a plan for helping the child acquire these skills, using the treatment approaches described by Lovaas et al. (1981). For example, if a parent indicated that a goal was for the child to acquire expressive language, the therapist described the sequence of steps presented by Lovaas et al. (1981) for achieving this goal (first obtaining cooperation with simple requests, followed by teaching imitation of nonverbal actions, increasing the child's rate of vocalizations, teaching imitation of sounds, chaining sounds together to form words, associating words with objects and events, and generalizing this vocabulary to everyday settings). In the next session, the parent trainer outlined the basic principles of discrimination learning, discrete trial formats, and functional analyses of maladaptive behaviors and assigned the chapters on these principles in Lovaas et al.'s (1981) manual. In subsequent sessions, the parent trainer worked for 2 to 3 minutes with the child on a particular treatment program (e.g., teaching imitation of nonverbal actions), requested that the parent work on the same program

with therapist present, then demonstrated another program, and so on. At the end of each 2 to 3 minutes of teaching, the parent trainer and parent gave each other feedback on their work with the child. Between sessions, parents were asked to work an additional 5 hours per week implementing the teaching programs set up with the parent trainer. Throughout parent training, children in this group were enrolled in special education classes in the public schools for 10 to 15 hours per week, with no direct involvement from the UCLA Young Autism Project.

Treatment fidelity. To check how much intervention children in the intensive treatment group actually received, student therapists recorded start and stop times for each treatment session. Their lead therapist summed these hours and reported them in monthly summaries of children's progress. One child was kept at low levels of service (10 to 15 hours per week) throughout treatment because therapy procedures appeared ineffective for him (as shown by lack of progress on the Early Learning Measure, described in *Measures*). Monthly summaries indicated that the other 14 children received a mean of 24.52 hours per week ($SD = 3.69$, range = 18.40 to 30.79) during the first year, with gradual reductions thereafter. Across all children in this group, the mean number of total treatment hours was 2,137.88 ($SD = 1,304.76$, range = 1,141.5 to 5,451.75), and the mean length of treatment was 33.44 months ($SD = 11.00$, range = 18 to 63). These results indicated that children received fewer hours of treatment per week, and the number of months of treatment was more variable than had been planned.

To ensure treatment quality in both intensive treatment and parent training, we implemented the same procedures used by Lovaas (1987; McEachin et al., 1993) for selecting, training, and supervising student therapists. In addition, to be allowed to continue at the UCLA Young Autism Project beyond their first 3 months, student therapists needed to obtain a passing grade on a test of knowl-

edge about the UCLA Young Autism Project treatment, achieve a satisfactory score in a standard behavior observation of their proficiency at discrete trials, and receive favorable ratings from their supervisors. To become supervisors (developing treatment plans for children and training novice student therapists under the direction of the authors), student therapists needed to accumulate a minimum of 1,500 hours of one-to-one treatment experience at the UCLA Young Autism Project, demonstrate mastery of research pertaining to applied behavior analytic treatment for children with pervasive developmental disorder, perform satisfactorily during a behavior observation of their skill at designing and implementing treatment plans, and obtain satisfactory ratings from other supervisors and from other children's parents (see Smith et al., in press, for more detail).

Measures

Most measures in the present study were standardized instruments designed to assess a broad range of skills for both typically and atypically developing children. Smith et al. (1993) recommended the use of such measures because they have more extensive psychometric data, cover more areas of functioning, and may have more clinical relevance than do measures specifically developed to assess functioning in children with pervasive developmental disorder (but see Mundy, 1993, for a different perspective). In addition to standardized instruments, however, the assessment also included one measure used in previous studies, namely, the type of classroom in which children were placed (Lovaas, 1987) and two measures developed specifically for evaluating treatment provided by the UCLA Young Autism Project (one assessing skill acquisition early in treatment and one assessing parental satisfaction with treatment).

Intellectual functioning. At intake and follow-up, the Stanford-Binet Intelli-

gence Scale, which was designed for individuals with a mental age (MA) of 2 years and older (Thorndike, Hagen, & Sattler, 1986) was administered to participants. If they did not achieve a basal (as defined in the manual), examiners administered the Bayley Scales of Infant Development-Mental Development Index (Bayley, 1969), which is a test given to children with developmental levels of 0 to 30 months. Both the Stanford-Binet and Bayley have been extensively used and validated for children with pervasive developmental disorder (Newsom & Hovanitz, 1997). In addition, children completed the Merrill-Palmer Scale of Mental Tests (Stutsman, 1948), a commonly used instrument that primarily assesses visual-spatial skills in children with developmental levels of 1.5 to 6.5 years and that has been shown to predict the later functioning of children with pervasive developmental disorder (e.g., Lord & Schopler, 1989). Children who did not answer any items correctly on this test were assigned a score of 18 months (the lowest MA covered by the test). Children who did not reach a ceiling were assigned a score of 78 months (the highest MA covered by the test).

Language functioning. At intake and follow-up, children completed the Reynell Developmental Language Scales (Reynell, 1990), which is used to assess comprehension (receptive language) and expressive language in children ages 1 to 6 years. This instrument is commonly administered to assess children with pervasive developmental disorder (Sparrow et al., 1997), though its psychometric properties have not been studied with this population. Children who did not answer any items correctly were assigned a score of 12 months in the Comprehension domain and 15 months in the Expressive Language domain (the lowest developmental ages covered by these scales). Children who did not reach a ceiling were assigned a score of 72 months (the highest developmental age covered by these scales).

Adaptive functioning. At intake and

follow-up, children's parents completed the survey form of the Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchetti, 1984), which is an interview given to caregivers regarding the everyday functioning of children ages 0 to 18 years in the areas of communication, daily living skills, and socialization. The Vineland is widely regarded as the instrument of choice for assessing adaptive functioning in children with pervasive developmental disorder (Newsom & Horowitz, 1997).

Socioemotional functioning. At follow-up, children's primary caregivers completed the Achenbach Child Behavior Checklist and their teachers completed a parallel form of this checklist, the Teacher Report Form (Achenbach, 1991). The Child Behavior Checklist and Teacher Report Form both measure behavior problems in the areas of social withdrawal, somatization, anxiety/depression, social problems, thought problems, attention problems, delinquency, and aggression. These two instruments are the most extensively studied measures of socioemotional functioning in children, but their utility for assessing children with pervasive developmental disorder has not been investigated.

Academic achievement. Children completed the Wechsler Individualized Achievement Test (Wechsler, 1992). This instrument is a well-normed measure of academic skills from preschool to adulthood.

Class placement. At follow-up, children's report cards and/or individualized education plans (IEPs) were obtained in order to determine their class placement: *regular classroom* (defined as placement in a classroom for typically developing children without an IEP or special services such as assistance from an aide), *regular classroom with support* (defined as placement in a regular classroom with an IEP or special services), or *self-contained classroom*. Class placement may be influenced by factors other than the child's functioning (e.g., parent advocacy or school district policy)

but provides a measure of "real world outcome" (Kazdin & Weisz, 1998, p. 29).

Progress in treatment. Children in the intensive treatment group completed the Early Learning Measure (available from the first author), which is a behavior observation instrument developed by Smith, Eikeseth, Buch, and Lovaas (1995) for assessing children's progress during the first 4 months of treatment. At intake, the examiner identified 40 instructions to which the child did not give the correct response, including 10 in each of the following four areas: receptive language (e.g., request to clap or wave), nonverbal imitation (e.g., imitating arms held up over the head), verbal imitation (e.g., imitating sounds of letters such as "ah" or "em"), and expressive language (e.g., stating "ball" when shown a ball). The instructions were re-administered at 1-month intervals during the first 4 months of treatment. Children's rate of acquisition of correct responses was used to predict outcome at follow-up. Leaf (1982) found that mastery of verbal imitation items in the first 3 months of treatment was strongly associated with outcome at age 7 years for 16 of the 19 intensively treated children studied by Lovaas (1987).

Parent evaluation. The Family Satisfaction Questionnaire (Smith, 1990), is a 20-item, written survey with questions on primary caregivers' appraisal of children's progress with treatment, quality of treatment, impact of treatment on the family, and relationship between the family and treatment staff. Respondents rated each item on a 7-point Likert scale. The psychometric properties of the Family Satisfaction Questionnaire have not yet been tested, apart from its correlations with other measures used in the present study (summarized in *Parent Ratings*).

Data Collection

Pretreatment evaluations occurred in the 3 months prior to treatment onset.

Follow-up evaluations occurred at a CA of 7 to 8 years. Parents completed surveys without a staff member present and were assured that their responses would not be disclosed to treatment personnel. Standardized tests were administered by doctoral students in clinical psychology at UCLA. These examiners had completed approximately 20 hours of training from the first or third author on the particular tests used in the study (as described by Smith et al., in press) and were blind to children's group assignment and treatment history. At pretreatment, 10 children had completed the Bayley with an examiner at an outside agency in the preceding 3 months. Therefore, to check reliability, we compared Bayley scores obtained by the UCLA Young Autism Project and those from the outside agency. Four of the 10 children were deemed untestable by the outside examiner but testable by the study examiner, who reported IQs ranging from 37 to 46. The 6 remaining children obtained an average IQ of 60 from the outside examiner and 57 from the study examiner, with a significant Pearson correlation between examiners, $r(6) = .90, p < .05$. The difference in IQ between examiners was 4 to 9 points for individual children. Thus, the UCLA examiner often obtained an IQ when the outside examiner did not; when both examiners obtained scores, these scores were similar, providing some evidence of reliability.

Senior staff members at the UCLA Young Autism Project (postgraduate or graduate students with a minimum of 2 years treatment experience) administered the Early Learning Measure. They were not blind to the child's treatment history, having interacted infrequently with the child (once a week or less). However, all administrations were videotaped and scored by an independent rater who was blind to the child's treatment history, and interrater reliability was evaluated by having a second independent, blind rater score 50% of the videotapes (Cohen's $\kappa = .86$).

Results

Comparison Between Groups

To test for differences between the intensive treatment and parent training groups at intake, we conducted two-tailed pooled variance t tests on IQ, Merrill-Palmer score, total Reynell score (Comprehension plus Expressive Language), and Vineland composite score. Another set of t tests was performed for individual scale scores on the Reynell and Vineland. Each set of t tests was Dunn-Bonferroni corrected for a family-wise error of .10. This procedure was also applied to follow-up data for these measures, with the modification that tests were one-tailed.

Table 2 presents intake and follow-up data. At intake, the groups appear to have been closely matched on all variables, with no statistically significant between-group differences. As expected, children displayed major developmental delays. For example, no child achieved a basal on the Stanford-Binet. Moreover, 23 of 28 children (82%) appeared to be nonverbal; they demonstrated no receptive or expressive vocabulary and uttered no words on the Reynell (13 children assigned to intensive treatment and 10 assigned to parent training). Also, 14 (50%) obtained a raw score of 0 on the Merrill-Palmer (6 intensive and 7 parent training children). Vineland scores also tended to be far below the national average.

At follow-up, as shown in Table 2, the intensive treatment group had a statistically significant advantage over the parent training group in IQ, visual-spatial skills (as measured by the Merrill-Palmer), and language development (as assessed by the total score on the Reynell, though not the scores on either of the two subscales), but not adaptive behavior in everyday settings (as indexed by the Vineland). The between-group differences on the Merrill-Palmer and Reynell may have been underestimated because of ceiling effects. Eight intensively treated children achieved the highest possible

Table 2
Means and SDs of Children's Standardized Test Scores by Group and Time

Measure	Intensive treatment						Parent training					
	Autism		PDD NOS ^a		Total		Autism		PDD NOS		Total	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
IQ ^a												
Intake	51.00	13.94	50.13	9.11	50.53	11.18	50.71	14.24	50.67	14.79	50.69	13.88
Follow-up	55.29	24.09	76.25	20.69	66.49	24.08	52.57	22.83	47.13	17.89	49.67*	19.74
Merrill-Palmer ^b												
Intake	21.29	4.65	21.88	4.64	21.60	4.49	23.57	6.11	20.00	4.43	21.92	5.50
Follow-up	54.43	23.60	68.63	13.41	64.33	18.74	45.67	23.64	52.67	20.60	49.17*	21.43
Reynell ^b												
Comprehension												
Intake	15.14	4.91	12.00	0.00	13.47	3.60	14.86	4.88	12.33	0.82	13.69	3.73
Follow-up	39.00	26.00	46.25	19.98	42.87	22.29	32.14	19.04	34.20	15.66	33.00	16.86
Expressive												
Intake	15.29	0.76	15.00	0.00	15.13	0.52	17.00	3.46	15.50	1.2	16.31	2.69
Follow-up	40.57	25.13	48.00	23.05	44.53	23.48	32.14	20.93	41.00	22.3	36.23	21.19
Total (Rec+Exp)												
Intake	30.43	5.59	27.00	0.00	28.60	4.07	31.86	8.32	27.83	2.04	30.00	6.34
Follow-up	79.57	41.99	94.25	41.99	87.40	46.21	64.29	39.25	57.20	21.48	61.33*	31.88
Vineland ^c												
Communication												
Intake	61.00	5.69	55.75	4.40	58.20	5.56	62.29	5.74	61.67	7.06	62.00	6.11
Follow-up	66.14	33.97	69.38	28.55	67.87	30.08	64.43	15.72	56.50	19.44	60.77	17.26
DLS ^d												
Intake	75.57	8.62	65.00 [†]	4.11	69.93	8.37	68.43	8.66	73.17	14.59	70.62	11.50
Follow-up	65.57	30.85	63.88	22.49	62.33	25.76	64.71	15.89	61.00	19.47	63.00	16.97
Socialization												
Intake	66.00	8.17	59.25	6.40	62.40	7.82	69.14	3.76	69.17	12.98	69.15	8.75
Follow-up	66.57	28.50	66.13	23.06	66.33	24.78	67.00	15.33	71.12	19.88	68.92	16.94
Composite												
Intake	68.86	11.16	59.2	4.99	63.44	9.35	65.00	9.54	65.40	10.41	65.17	9.44
Follow-up	62.14	32.84	60.44	25.11	61.19	29.72	61.86	16.83	53.80	16.84	58.50	16.58

^aBayley scored as ratio IQ. Stanford-Binet as deviation IQ. ^bDevelopmental age (months). ^cStandard score (population $M[SD] = 100[15]$).

^dDevelopmental Language Scales. ^ePervasive development delay not otherwise specified.

[†] $p < .05$, Intensive Treatment > Parent Training.

[‡] $p < .05$, PDD NOS > Autism.

Merrill-Palmer score, compared to 1 child in parent training. Also, 4 intensively treated children obtained the maximum scores in both Comprehension and Expressive Language on the Reynell; no child in parent training did so. (One child in parent training did reach the maximum in Expressive Language, but not Comprehension). However, floor effects appeared to apply about equally to the two groups. Four intensively treated and 5 parent training children did not achieve a basal on the Stanford-Binet; 1 intensively treated and 2 parent training children tested as nonverbal on the Reynell; and 1 intensively treated and 2 parent training children had substantially delayed visual-spatial skills, as evidenced by raw scores of 0 on the Merrill-Palmer. In addition, as indicated in Table 2, there was little difference between groups in mean Vineland scores.

The intensively treated group had significantly less restrictive school placements than did the parent training group (4 regular ed and 24 regular ed with support vs. no regular ed and 3 regular ed with support, respectively). The remaining children were placed in self-contained classes (9 from the intensive treatment group and 11 from the parent training group). The intensively treated group also had higher WIAT scores than did the parent training group. The mean standard score was 75.71 ($SD = 21.31$) for the intensively treated group and 58.44 ($SD = 18.43$) for the parent training group. Table 3 presents results for the only other test given at follow-up, the Child Behavior Checklist. It can be seen from the table that there was little difference between groups in behavior problems, as reported by parents and teachers.

Table 3
Results of Follow-Up Assessment by Group on the Child Behavior Checklist

Measure	Intensive treatment				Parent training			
	Parent		Teacher		Parent		Teacher	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Child Behavior Checklist <i>T</i> score								
Withdrawal	59.33	10.26	61.89	7.04	60.17	7.81	55.00	4.40
Somatization	56.11	8.16	52.33	4.95	56.11	8.16	54.86	8.47
Anxiety/Depression	52.22	5.24	54.22	5.26	59.67	11.59	54.57	4.08
Social Problems	60.11	13.46	59.78	9.59	64.33	11.34	57.43	8.02
Thought Problems	67.11	10.82	64.67	13.62	64.47	12.74	62.57	7.55
Attention Problems	64.78	10.32	64.89	12.80	67.50	4.18	61.57	9.29
Delinquency	54.67	9.24	53.44	6.39	59.00	6.42	54.00	5.13
Aggression	56.11	9.10	60.00	10.81	59.67	10.41	55.71	5.53

Comparison Between Children With Autism and Pervasive Developmental Disorder NOS

Within each group, children with autism were compared to those with pervasive developmental disorder NOS. The statistical procedures were the same as those used to compare intensive treatment to parent training, with the exception that all tests were two-tailed. Table 2 summarizes the data for children with autism and those with pervasive developmental disorder NOS in each of the two treatment groups. As shown, intake test scores of children with autism were similar to those of children with pervasive developmental disorder NOS in the intensive treatment group and in the parent training group. Differences between diagnostic categories in follow-up test scores also did not reach statistical significance. However, visual inspection of the results indicates that within the intensive training group, children with pervasive developmental disorder NOS tended to obtain higher scores than did children with autism. For example, children with pervasive developmental disorder NOS averaged 17 IQ points above those with autism. Given this finding and the low statistical power to detect differences between diagnostic categories (due to small cell sizes and large within-group variability), it is possible that important

differences existed in the follow-up scores achieved by children with pervasive developmental disorder NOS and those achieved by children with autism.

Prediction of Outcome in the Intensive Treatment Group

Two of the 15 intensively treated children (1 with autism and 1 with pervasive developmental disorder NOS) met the criteria used by McEachin et al. (1993) for classifying children as "best outcome" (placement in regular classes without special services and IQ > 85). An additional 2 children in this group (1 with autism and 1 with pervasive developmental disorder NOS) met the placement criterion while scoring just below the IQ cutoff. These 4 children performed in the average range on all other tests, with the exception that clinically significant behavior problems were reported for 1 child on the Child Behavior Checklist by the parent (though not the teacher). By contrast, only 1 child in the parent training group scored in the average range on any test at follow-up (a child with pervasive developmental disorder NOS who scored 87 on the Stanford-Binet and 93 on the WIAT, but below 75 on all other tests, and who had a full-time, individual aide in a regular class).

Though Lovaas and Smith (1988) reported that best-outcome children

tended to have higher intake IQs than did non-best-outcome children, this finding was not replicated in the present study. Of the 4 intensively treated children described in the preceding paragraph, 1 scored above the mean IQ at intake, 1 scored at the mean, and 2 scored below. More generally, the correlation between intake and follow-up IQ in the intensive treatment group was small and not statistically significant, Pearson $r(15) = .08$. Further, unprotected Pearson correlations indicated that intake IQ did not significantly correlate with any other outcome variable. In addition, unprotected Pearson correlations for each intake measure (including both demographic variables and standardized test scores) with each follow-up measure revealed only three statistically significant associations: intake Merrill-Palmer with follow-up Merrill-Palmer (.43), intake Reynell with follow-up Reynell (.36), and intake Reynell with follow-up Vineland (.48). Given the large number of correlations performed, even these statistically significant findings may have been spurious. Overall, then, intake data were poor predictors of follow-up scores.

Lovaas and Smith (1988) suggested that the Early Learning Measure would predict follow-up tests more strongly than would intake standardized tests. Specifically, they proposed that mastery of verbal imitation on the Early Learning Measure 3 months after treatment onset would predict who achieved average functioning at follow-up. Eight of the 15 intensively treated children (including all 4 children who scored in the average range on a majority of tests at follow-up) met this mastery criterion. Thus, the criterion identified the 4 children who achieved average functioning on most follow-up tests but also yielded false positives. Three of these 4 children also demonstrated mastery of expressive labels at 3 months, whereas no other child did so. Hence, mastery of expressive labels may also have been associated with outcome. However, the other two scales (Receptive Actions and Nonverbal Imitation)

were not associated with outcome because almost all children (13 of 15) showed mastery 1 month after treatment onset.

Parent Ratings

Table 4 summarizes parent ratings for each of the 20 items on the Family Satisfaction Questionnaire. As shown in the table, parents in both groups tended to report that children improved. On average, they described moderate gains. All parents expressed ongoing concern about their children's communication skills, but some reported that other behaviors were no longer a problem: social skills (1 in intensive treatment), play and leisure skills (3 in intensive treatment, 1 in parent training), tantrums and aggression (4 in intensive treatment, 2 in parent training), self-stimulatory behaviors (4 in intensive treatment, 2 in parent training), self-help (3 in intensive treatment, 1 in parent training). No ratings indicated that children's behavior worsened, apart from one report of a slight worsening of aggression for a child in intensive treatment. Spearman correlations between parent ratings and follow-up IQ, Merrill-Palmer, Reynell, and Vineland were moderately high, with r s ranging from .24 to .67. However, correlations between parent ratings of a particular behavior and standardized tests designed to measure the same behavior tended to be no higher than correlations across behaviors. For example, the correlation coefficient for parent ratings of communication and post-treatment Reynell score was .47, but the coefficients between these ratings and IQ, Merrill-Palmer, and Vineland were .50, .67, and .62, respectively. Hence, parent ratings and follow-up test scores may have had convergent but not discriminant validity. This finding indicates that parent ratings were generally consistent with objective measures of children's progress but reflected children's overall improvement rather than improvement in a particular behavior (e.g., communication, social skills).

Table 4
Parent Ratings of Treatment by Group

Rating	Intensive treatment		Parent training	
	Mean	SD	Mean	SD
Child's progress ^a				
Language/communication	2.41	.67	2.50	0.84
Social skills	2.58	.67	2.50	.55
Play skills	2.42	1.00	2.33	.82
Aggression/tantrums	2.25	1.22	2.17	1.17
Self-stimulatory/ritualistic behaviors	2.25	1.42	2.00	1.27
Self-help	1.92	.67	2.00	1.27
Workload ^b				
For child	4.08	.67	4.33	1.03
For parent	3.91	.29	5.00	1.10
Stress during treatment ^c				
For child	4.33	.66	3.83	1.84
For parent	3.08	1.68	4.60	1.34
Quality of treatment ^d				
Treatment methods	1.27	.67	2.00	1.27
Therapist motivation and concern	1.75	.87	1.20	.45
Comparison with services from other agencies	1.75	1.06	2.80	1.10
Recommendation of treatment to others	1.58	1.73	1.20	.45
Worth time/effort	1.25	.62	1.20	.45
Would enroll again	1.17	.39	1.00	.00
Overall opinion	1.17	.39	1.00	.00

Note. There was no statistically significant between-group differences on any question.

^a1 = no longer a problem, 4 = no change, 7 = much worse. ^b1 = much too little, 4 = just right, 7 = too much. ^c1 = greatly reduced, 4 = no effect, 7 = greatly increased. ^d1 = very positive, 4 = neutral, 7 = very negative.

As also shown in Table 4, parents in both groups tended to give very favorable ratings for quality of treatment, impact of treatment on the family, and relationship between the family and treatment staff. One parent in the parent training group reported that the treatment was stressful for her. However, there were no other negative ratings.

Discussion

Consistent with previous studies based on the treatment manual we used (McEachin et al., 1993), intensively treated children outperformed children in a parent training group at follow-up on measures of intelligence, visual-spatial ability, language, and academic achievement. Also, as a group, they had less restrictive school placements. Extending previous reports, results of the present study reveal that parents in both groups held highly positive views about the services their children received. Also, children with pervasive developmental disorder NOS benefited at least as much from intensive treatment as did children

with autism. Despite these favorable outcomes, between-group differences in follow-up IQ were roughly half that reported by McEachin et al. (16 vs. 31 points) as was the proportion of children placed in regular classes without special services (27% vs. 47%). Further, in the present study, intensively treated children did not differ from children in the parent training group on standardized tests of behavior problems and adaptive functioning in everyday settings at follow-up, whereas McEachin et al. (1993) found substantial advantages for intensively treated children on these variables.

The present study retained features that were strengths of previous research (manualized treatments based on extensive research and supervised by experienced personnel, blind examiners, and long-term follow-up). We also incorporated improvements, particularly the use of a true experimental design and a comprehensive assessment battery that was uniform across children. Still, the study has limitations. The small sample size and heavy tailed, skewed distributions of scores precluded conducting some statis-

tical procedures, such as factorial analyses of variance to examine diagnosis-by-treatment interactions, that might have helped in interpretation of the results, and it yielded low power to detect predictors of treatment response. The assessment instruments included only one measure of social skills (the Vineland Socialization Domain), a parent satisfaction questionnaire with untested psychometric properties, no measure of children's or parent's quality of life, and no measure of parent's participation in treatment. Also, the assessment measures used in the study consisted mostly of instruments developed for both typically and atypically developing children rather than ones specifically developed for children with developmental disabilities (e.g., the Achenbach Child Behavior Checklist instead of the Aberrant Behavior Checklist—Aman & Singh, 1986). The addition of measures designed for children with developmental disabilities might have been useful. Finally, the study lacked a standardized diagnostic instrument (and had no follow-up diagnostic assessment at all). A standardized instrument now exists for identifying young children with autism (Lord, 1995), though not for differentiating between autism and pervasive developmental disorder NOS (Myhr, 1998).

We are currently participating in multisite research projects aimed at overcoming these the limitations (Smith et al., in press). Nevertheless, we believe that the present study supports several conclusions. First, results confirm that some children with pervasive developmental disorder may make large gains with early intervention. Second, in view of the differences observed between groups at follow-up, intensive treatment may be more effective than is parent training, even when parent training incorporates many features recommended by professionals (e.g., individualized, collaborative services in the family home). Third, given that children with pervasive developmental disorder NOS obtained outcomes at least as positive as those

obtained by children with autism, intensive early intervention may be especially beneficial for children with pervasive developmental disorder NOS. Fourth, because parents who varied greatly in socioeconomic status and other demographic variables consistently evaluated treatment favorably, a wide range of parents may experience the treatment as helpful rather than stressful, despite its high intensity. Fifth, in view of the positive ratings given by parents in both intensive and parent training groups, the differences between groups are likely to have arisen from specific aspects of the intervention rather than placebo factors (e.g., the credibility of intensive treatment or the warm relationships it fostered between families and staff). Finally, because the Early Learning Measure appeared to be more strongly associated with outcome than were any of the intake standardized tests, measures of children's skill acquisition early in treatment warrant further scrutiny as predictors of outcome.

Although some conclusions are supported by results of the study, new questions are also raised. First, considering that we used the same treatment manual as did Lovaas and colleagues (Lovaas, 1987; McEachin et al., 1993), why were our results less favorable? One possibility, of course, is that the treatment is really not as effective as Lovaas (1987) reported. Nevertheless, another possibility is that methodological differences between studies accounted for the differences. For example, despite the use of assessment procedures that yielded scores for children who might otherwise have been classified as untestable (see *Initial Analyses*), intensively treated children had a lower intake IQ and level of language than those in Lovaas's investigation and most other published outcome studies (cf. Smith, 1999). As an illustration, the mean intake IQ of intensively treated children was 50 in the present study compared to 63 in Lovaas (1987). Neither IQ nor language strongly predicted follow-up scores in the present

study, but they have been predictive in other studies (Smith, 1999) and, hence, may have been a factor in the relatively modest results reported here. Studies with large samples and without IQ cutoffs are needed to determine whether these intake variables are reliably associated with outcome in intensive early intervention.

Second, changes in the treatment may have undermined its effectiveness. For example, we directed treatment in the present study, whereas treatment in the Lovaas (1987) study was overseen by its originators (Lovaas et al., 1981). Because we had many years of training prior to the study and employed rigorous training and evaluation procedures to ensure the quality of therapy (Smith et al., in press), we expected that the change in treatment directors would not affect the results, but our results suggest the need for further research on this issue. Another change in the present study was that we reduced the intensity of intervention, as described in *Treatment*. This change may have lowered treatment efficacy. For example, the intervention we provided was focused on skill-building in a one-to-one setting before generalizing to school and other everyday settings. Therefore, phasing out treatment for children who were progressing slowly may have prevented them from improving their performance on measures of everyday functioning, such as the Vineland and Child Behavior Checklist. Reducing parental involvement for intensively treated children may also have contributed to the lack of improvement on these measures. The relatively low number of treatment hours, relative to Lovaas (1987), may have limited gains on all follow-up measures. Direct investigations of treatment intensity are needed to evaluate these possibilities.

Children received an average of 25 hours of treatment per week rather than the 30 hours that we intended to provide. Our impression is that this shortfall resulted from staff shortages, scheduling conflicts, and illnesses. Thus, the logistics of arranging for intensive treatment

turned out to be more formidable than we had anticipated and may pose more of an obstacle to replication than we have previously acknowledged (e.g., Smith & Lovaas, 1997). Employing paid staff instead of students may ensure a steadier supply of therapists, and scheduling extra hours may compensate for missed sessions. Both would increase expenses, but the intervention may still be cost effective (Jacobson, Mulick, & Green, 1998).

A final, critical question is whether it is appropriate to provide intensive treatment to achieve the outcomes we have reported. We think the answer is plainly *yes* for the children who achieved average functioning on most follow-up measures but less clear for the remaining children. The latter children increased their rate of skill acquisition when they entered treatment yet did not make gains in standardized test scores at follow-up. One possible reason for this finding is that the children may have made advances while in treatment but regressed afterward. Another is that the children may have reached a plateau in treatment that they would have eventually reached without treatment. Alternatively, they may have continued to acquire skills more rapidly than they would have without treatment, but their skill acquisition did not raise standardized test scores. Unfortunately, our data do not provide a basis for deciding among these possibilities (and there may be others). Each possibility, however, suggests that the treatment is potentially appropriate for children who do not achieve average functioning but needs substantial additions or modifications (e.g., identifying continuation services that enable children to maintain gains, determining when children reach a plateau and making a transition to another program at that time, and setting less ambitious goals for treatment).

Although the results were more mixed than in some previous studies, the present study substantiates the view that intensive early intervention can be a powerful intervention. Of particular impor-

tance may be the finding that children with pervasive developmental disorder NOS gained as much as or more than did children with autism. This finding suggests that intensive early intervention may be effective not only for autism but also for other pervasive developmental disorder.

References

- Achenbach, T. M. (1991). *Integrative guide for the 1991 Child Behavior Checklist/4-18, YSR, and Teacher Report Form profiles*. Burlington: University of Vermont, Department of Psychiatry.
- Aman, M. G., & Singh, N. N. (1986). *Aberrant Behavior Checklist: Manual*. E. Aurora, NY: Slosson.
- Anderson, S. R., Avery, D. L., DiPietro, E. K., Edwards, G. L., & Christian, W. P. (1987). Intensive home-based intervention with autistic children. *Education and Treatment of Children, 10*, 352-366.
- Bayley, N. (1969). *Bayley Scales of Infant Development*. New York: Psychological Corp.
- Birnbrauer, J. S., & Leach, D. J. (1993). The Murdoch Early Intervention Program after two years. *Behaviour Change, 10*, 63-74.
- Fenske, E. C., Zalski, 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.
- Green, G. (1996). Early behavioral intervention for autism: What does research tell us? In C. Maurice (Ed.), *Behavioral intervention for young children with autism* (pp. 29-44). Austin, TX: Pro-Ed.
- Gresham, F. M., & MacMillan, D. L. (1997). Autistic recovery? An analysis and critique of the empirical evidence on the Early Intervention Project. *Behavioral Disorders, 22*, 185-201.
- Guralnick, M. J. (1998). Effectiveness of early intervention for vulnerable children: A developmental perspective. *American Journal on Mental Retardation, 102*, 319-345.
- Harris, S., Handleman, J., Gordon, R., Kristoff, B., & Fuentes, F. (1991). Changes in cognitive and language functioning of preschool children with autism. *Journal of Autism and Developmental Disabilities, 21*, 281-290.
- Hoyson, M., Jamieson, B., & Strain, P. S. (1984). Individualized group instruction of normally developing and autistic-like children: A description and evaluation of the LEAP curriculum model. *Journal of the Division of Early Childhood, 8*, 157-181.
- Jacobson, J. W., Mulick, J. A., & Green, G. (1998). Cost-benefit estimates for early intensive behavioral intervention for young children with autism: General models and single state case. *Behavioral Interventions, 13*, 201-226.
- Leaf, R. B. (1982). Predicting outcome in the UCLA Autism Project. In O. I. Lovaas (Chair), *The UCLA Autism Project*. Symposium conducted at the annual meeting of the American Psychological Association, Washington, DC.
- Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry, 36*, 1365-1382.
- 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., Ackerman, A. B., Alexander, D., Firestone, P., Perkins, J., & Young, D. (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.
- Lovaas, O. I., Smith, T., & McEachin, J. J. (1989). Clarifying comments on the young autism study: Reply to Schopler, Short, and Mesibov. *Journal of Consulting and Clinical Psychology, 57*, 165-167.
- 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.
- Myhr, G. (1998). Autism and other pervasive developmental disorders: Exploring the dimensional view. *Canadian Journal of Psychiatry, 43*, 589-595.
- Newsom, C., & Hovanitz, C. A. (1997). Autistic disorder. In E. J. Mash & L. G. Terdal (Eds.), *Assessment of childhood disorders* (3rd ed., pp. 408-452). New York: Guilford.
- Newsom, C., & Rincover, A. (1989). Autism. In E. J. Mash & R. A. Barkley (Eds.), *Treatment of childhood disorders* (pp. 286-346). New York: Guilford.

- Reynell, J. K. (1990). *Reynell Developmental Language Scales*. Los Angeles: Western Psychological Association.
- Rutter, M., Bailey, A., Bolton, P., & LeCouteur, A. (1994). Autism and known medical conditions: Myth and substance. *Journal of Child Psychology and Psychiatry*, 35, 311-322.
- Scarr, S., & Arnett, J. (1987). Malleability: Lessons from intervention and family studies. In J. J. Gallagher & C. T. Ramey (Eds.), *The malleability of children* (pp. 71-84). Baltimore: Brookes.
- Schopler, E., Short, A., & Mesibov, G. B. (1989). Relation of behavioral treatment to "normal functioning": Comment on Lovaas. *Journal of Consulting and Clinical Psychology*, 57, 162-164.
- Sheinkopf, S., & Siegal, B. (1998). Home-based behavioral treatment for young autistic children. *Journal of Autism and Developmental Disorders*, 28, 15-23.
- Siegal, B., & Elliott, G. R. (1988). *Stanford Developmental History Questionnaire*. Stanford, CA: Author.
- Smith T. (1990). *Family Satisfaction Questionnaire*. Unpublished instrument. (Available from Dr. T. Smith, Department of Psychology, Washington State University, PO Box 644820, Pullman, WA 99164-4820)
- Smith, T. (1997, June 23). Girls with pervasive developmental disorders. *Medscape Mental Health*, 2, 1-9.
- Smith, T. (1999). Outcome of early intervention for children with autism. *Clinical Psychology: Science and Practice*, 6, 33-49.
- Smith, T., Donahoe, P. A., & Davis, B. J. (in press). The UCLA treatment model. In J. S. Handleman & S. L. Harris (Eds.), *Preschool education programs for children with autism* (2nd ed.). Austin, TX: Pro-Ed.
- Smith, T., Eikeseth, S., Buch, G., & Lovaas, I. (1995). *Early learning measure*. Unpublished test. (Available from Dr. T. Smith, Department of Psychology, Washington State University, PO Box 644820, Pullman, WA 99164-4820)
- Smith, T., & Lovaas, O. I. (1997). The UCLA Young Autism Project: A reply to Gresham and MacMillan. *Behavioral Disorders*, 22, 202-218.
- Smith, T., McEachin, J. J., & Lovaas, O. I. (1993). Comments on replication and evaluation of outcome. *American Journal on Mental Retardation*, 97, 385-391.
- Sparrow, S. S., Bella, D. A., & Cicchetti, D. V. (1984). *The Vineland Adaptive Behavior Scales*. Circle River, MN: American Guidance Service.
- Sparrow, S., Marans, W., Klin, A., Carter, A., Volkmar, F. R., & Cohen, D. J. (1997). Developmentally based assessments. In D. J. Cohen & F. R. Volkmar (Eds.), *Handbook of autism and pervasive developmental disorders* (2nd ed., pp. 411-447). New York: Wiley.
- Stutsman, R. (1948). *Guide for administering the Merrill-Palmer Scale of Mental Tests*. New York: Harcourt, Brace & World.
- Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale* (4th ed.). Chicago: Riverside.
- Volkmar, F. R., Klin, A., Siegal, B., Szatmari, P., Lord, C., Campbell, M., Freeman, B. J., Cicchetti, D. V., Rutter, M., Kline, W., Buitelaar, J., Hattab, Y., Fombonne, E., Fuentes, J., Werry, J., Stone, W., Kerbeshian, J., Hoshino, Y., Bregman, J., Loveland, K., Szymanski, L., & Towbin, K. (1994). Field trial for autistic disorder in *DSM-IV*. *American Journal of Psychiatry*, 151, 1361-1367.
- Wechsler, D. (1992). *Manual for the Wechsler Individual Achievement Test*. San Antonio, TX: Psychological Corp.
- Weinberg, R. A. (1989). Intelligence and IQ: Landmark issues and great debates. *American Psychologist*, 44, 99-104.
- White, K. R. (1985-1986). Efficacy of early intervention. *Journal of Special Education*, 41, 401-416.

Received 10/19/99, accepted 2/8/00.

This research was supported in part by Department of Education Grant No. H133G80103 (Intensive Early Intervention for Children With Mild to Moderate Mental Retardation) and by the UCLA Regents Account No. 4-444040-LS-60090. The authors thank Scott Wright and Kristen Villone for serving as case supervisors as well as Eric Holman for performing the random assignment. Requests for reprints should be sent to Tristram Smith, Department of Psychology, Washington State University, PO Box 644820, Pullman, WA 99164-4820 (E-mail: tristram@mail.wsu.edu).