

Treatment

Early Intensive Behavioral Treatment: Replication of the UCLA Model in a Community Setting

HOWARD COHEN, Ph.D.

Valley Mountain Regional Center, Stockton, CA

MILA AMERINE-DICKENS, M.S.

Central Valley Autism Project, Modesto, CA

TRISTRAM SMITH, Ph.D.

Department of Pediatrics, University of Rochester Medical Center, Rochester, NY

ABSTRACT. Although previous studies have shown favorable results with early intensive behavioral treatment (EIBT) for children with autism, it remains important to replicate these findings, particularly in community settings. The authors conducted a 3-year prospective outcome study that compared 2 groups: (1) 21 children who received 35 to 40 hours per week of EIBT from a community agency that replicated Lovaas' model of EIBT and (2) 21 age- and IQ-matched children in special education classes at local public schools. A quasi-experimental design was used, with assignment to groups based on parental preference. Assessments were conducted by independent examiners for IQ (Bayley Scales of Infant Development or Wechsler Preschool and Primary Scales of Intelligence), language (Reynell Developmental Language Scales), nonverbal skill (Merrill-Palmer Scale of Mental Tests), and adaptive behavior (Vineland Adaptive Behavior Scales). Analyses of covariance, with baseline scores as covariates and Year 1-3 assessments as repeated measures, revealed that, with treatment, the EIBT group obtained significantly higher IQ ($F = 5.21$, $p = .03$) and adaptive behavior scores ($F = 7.84$, $p = .01$) than did the comparison group. No difference between groups was found in either language comprehension ($F = 3.82$, $p = .06$) or nonverbal skill. Six of the 21 EIBT children were fully included into regular education without assistance at Year 3, and 11 others were included with support; in contrast, only 1 comparison child was placed primarily in regular education. Although the study was limited by the nonrandom assignment to groups, it does provide evidence that EIBT can be successfully implemented in a community setting. *J Dev Behav Pediatr* 27:145-155, 2006. Index terms: *autism, early intervention, applied behavior analysis, behavioral treatment.*

The design and implementation of methodologically rigorous treatment studies are daunting tasks and, in the area of treatment for autism spectrum disorders, often emotionally charged and publicly vetted as well. Matching groups on a variety of important measures, including severity of disability, individual characteristics of the child, multiple important socio-familial and environmental factors, as well as controlling multiple treatment issues such as fidelity, intensity and length of treatment and pre-determining appropriate outcome measures are all challenging (and expensive). Moving treatment studies from the laboratory setting into the community presents additional hurdles, yet this is ultimately the setting in which the efficacy of treatment models needs to be evaluated. Cohen and colleagues

are to be commended for implementing a community-based treatment study with matched samples, documentation of treatment fidelity, and comprehensive 3-year follow-up. However, the setting was based in a community program that is mandated to provide treatment to families of children with autism spectrum disorders who are then free to accept a plan or not, which prohibited random assignment to treatment. This introduced potential bias in their groups, with more educated and dual parent families in the EIBT group. There are strengths as well as limitations in this study. Although it does not resolve the controversies that continue regarding the "best" treatments for young children with ASD, we include it because of the critical need for evaluation of treatment approaches. The reviewers pointed out the limitations in this community approach as well as its strengths. The reader is encouraged to look at both in reviewing this article. We hope that it will inspire others to do these vitally needed treatment effectiveness studies. —Editor

Received September 2005; accepted February 2006.
Address for reprints: Mila Amerine-Dickens, M.S., 1317 Oakdale Rd., Suite 800, Modesto, CA 95355; e-mail: mamerine-dickens@cvap.org.

In an era when Autistic Spectrum Disorder (ASD) was viewed as largely untreatable,¹ Ivar Lovaas' 1987 outcome study² became a pivotal event that provided optimism about behavioral interventions for ASD. Almost half (9 of 19) of the children with autism who began intensive behavioral treatment prior to the age of 4 years from the UCLA/Lovaas clinic (40 hours per week for 2 or more years) were fully included into regular education and showed significant gains in intellectual achievement. A follow-up study of the same children showed sustained gains.³ This finding, coupled with a general trend toward earlier diagnosis of ASD (under 3 years of age)⁴ and the recent exponential increase in documented cases of ASD,⁵ made Lovaas' results even more influential and replication of his research more compelling.

Replication of the UCLA/Lovaas Model involves the following key elements⁶: (1) clinical internship and training on the UCLA/Lovaas Model of intervention under the direction of qualified supervisors; (2) implementation of the model for 35 to 40 hours per week throughout the year, including one-to-one instruction, peer play training sessions, inclusion into regular education classrooms, and generalization activities; (3) parent training to foster the child's acquisition and generalization of skills; and (4) annual outcome measures.

Several studies have partially replicated the UCLA/Lovaas Model. In the only randomized clinical trial, 28 children with ASD received either intensive behavioral treatment or parent training.⁷ The intensive treatment group averaged 25 hours per week in the first year which faded over the next 1 to 2 years. The comparison group participated in 10 to 15 hours per week of special education classes and received 5 hours per week of parent training for 3 to 9 months. The intensive children outperformed the comparison children on intellectual, visual-spatial, and academic measures. However, gains were substantially smaller than in Lovaas' original study. For example, the between-group IQ difference at follow-up was 16 points compared to the 31 reported by Lovaas. In other partial replications of the UCLA model, children with ASD obtained 15 to 35 hours per week of treatment and obtained results similar to those reported in the randomized clinical trials^{8,9}; similar results also have been reported for other early intensive behavioral treatment (EIBT) models with about 25 hours per week of treatment.^{10,11}

Concerns have been expressed about the difficulty of offering treatment at this level of intensity to community samples,¹² and mixed results of EIBT in community settings have been reported. One investigation indicated a lack of significant improvements in a sample of 66 children with ASD.¹³ A multiple baseline study of 6 children found clear short-term gains but equivocal long-term effects.¹⁴ However, a third study reported that an EIBT group ($n = 29$) in a community agency made statistically significant gains in all areas of development except motor skills, relative to 2 comparison groups.¹⁵ Moreover, 13 of the 29 EIBT children (45%) achieved IQs in the average to above average range. In the first replication of the UCLA Model that included all of the elements identified by Lovaas, 11 of 23 children with ASD (48%) achieved full inclusion into regular education and

IQ scores greater than 85.¹⁶ However, the study did not have a comparison group.

Although these studies generally confirm that EIBT is effective, differing results across studies and methodological limitations such as the absence of comparison groups in many reports weaken the ability to truly validate the optimism generated by the initial Lovaas study. Accordingly, the present study was an attempt to fully replicate that study in a community setting. Research questions included the following: (1) Can the Lovaas/UCLA model be replicated in a community setting? (2) What outcomes do children with ASD achieve with this intervention?

METHODS

Participants

Participants were 42 children in 2 groups: The early intensive behavioral treatment (EIBT) group ($n = 21$) received 35 to 40 hours of behavioral intervention, 47 weeks per year, for 3 or more years. The comparison group ($n = 21$) received services from local public schools. In accord with the UCLA Young Autism Project multisite research replication protocol, participation criteria for both groups included (1) primary diagnosis of autistic disorder or pervasive developmental disorder not otherwise specified based on an evaluation by an independent licensed psychologist and confirmed by the Autism Diagnostic Interview-Revised,¹⁷ (2) pretreatment IQ above 35 on the Bayley Scales of Infant Development-Revised (BSID-R),¹⁸ (3) chronological age between 18 and 42 months at diagnosis and under 48 months at treatment onset, (4) no severe medical limitation or illness including motor or sensory deficits that would preclude a child from participating in 30 hours per week of treatment, (5) residence within 60 km of the treatment agency, (6) no more than 400 hours of behavioral intervention prior to intake, and (7) parent's agreement to participate actively in parent training and generalization and to have an adult present during home intervention hours.

In addition to the 21 participants in each group, there were 5 dropouts who were excluded from the data analyses (3 in the EIBT group and 2 in the comparison group). One EIBT participant moved out of the area at 17 months into treatment and was unavailable for follow-up; 2 withdrew their participation, 1 at 3 months and the other at 18 months. Dropouts were similar to completers with regard to age of diagnosis (24, 36, and 22 months), baseline IQ (42, 44, and 44), and 1-year IQ (58 and 61; score unavailable for participant who dropped out after 3 months). Two comparison children were dropped because parents either declined annual testing of their child or could not be contacted. All other eligible referrals enrolled in the study, completed yearly follow-up assessments, and were included in the data analyses.

All treatment in both groups was provided at no cost to families. Funding was split between 2 public agencies: (1) the Valley Mountain Regional Center (VMRC; Stockton, CA) and (2) the child's Special Education Local Planning Area (SELPA) of residence. VMRC is contracted by the California Department of Developmental Services to

identify and coordinate services for individuals with developmental disabilities; its catchment area includes San Joaquin, Stanislaus, Calaveras, Amador, and Tuolumne Counties. SELPAs are contracted by the California Department of Education to provide special education instruction.

Design

Inasmuch as VMRC and SELPA had a mandate to provide free and appropriate services, legal and ethical considerations precluded random assignment of children to groups. Therefore, a quasi-experimental design was used. A comparison group was formed by identifying children who met participation criteria for EIBT and whose parents chose other services. Specifically, for each EIBT participant, a file review was initiated at VMRC to identify a matching child who was not receiving EIBT; the first identified child was then added to the comparison group. Comparison children were followed prospectively and received the same annual assessments as EIBT children.

To ensure that choices were available to families and that families were aware of these choices, VMRC and SELPA 6, along with nonpublic educational agencies and parents, developed an ongoing collaborative program (Autism Connection).¹⁹ The Early Autism Diagnostic Clinic (EADC) was created by the Autism Connection (1) to provide expert evaluations for autism and related disorders (or referrals to other experts in the area) and (2) to bring together local clinicians, VMRC, parents, school district representatives, and advocates to communicate directly with each other, at the EADC, rather than requiring the parents to endure separate meetings. At the time of diagnosis, an educational consultant from the EADC and a representative from the school district of residence presented the family, orally and in writing, a Matrix of Educational Options developed by the Autism Connection. This matrix delineates the service agencies in the child's area of residence and their eligibility criteria, along with the roles and responsibilities of parents, service providers, and funding agencies in implementing interventions.

Options included special education settings, Autistic Spectrum Disorder (ASD) classes, speech and language services, occupational therapy, genetic counseling, behavior intervention services, grief counseling, Early Start programs for children under 3 years old, and EIBT Programs, including the agency in this study (Central Valley Autism Project; CVAP) and other EIBT providers. During the enrollment period (1995–2000), the number of other EIBT providers ranged from 1 to 3. At times when CVAP did not have openings, the education consultant and school representative removed CVAP from the Matrix. EADC educational consultant and school representatives were otherwise independent of the study.

Treatment Procedures: EIBT Group

EIBT consisted of 35 to 40 hours per week of intervention based on Lovaas' UCLA treatment model.^{2,6,20} Seventeen of the 21 participants remained in EIBT for 3 years. Four others ended EIBT prior to 3 years but

completed follow-up assessments and are included in the statistical analyses; 1 completed the intervention protocol and was fully included in regular education at Year 2, whereas 3 others were transferred to other services (2 after 6 months and 1 at Year 2) because their progress did not meet specific, predetermined developmental markers for continuing intervention. Markers at 6, 12, 24, and 36 months were identified collaboratively by Autism Connection.²¹ For example, at 24 months, the IEP team considered whether the child showed one or more signs of progress such as the following: (1) the child's standardized cognitive testing indicated steady growth or near-average functioning; (2) objective data collected on EIBT instruction demonstrated that the child was mastering new skills; (3) objective data revealed an increase in the child's frequency of initiating language or peer interaction; or (4) the child was included in a general education placement with similar-aged peers for systematically increasing increments of time and was acquiring age-appropriate pre-academic skills.

The EIBT agency, CVAP, met all criteria for replication of Lovaas' UCLA treatment model and participated in a multicenter study supported by the National Institute of Mental Health. The UCLA model relies exclusively on behavioral techniques such as unambiguous instruction, shaping through positive reinforcement of successive approximations, systematic prompting and fading procedures, discrimination learning, and careful task analysis. Positive reinforcers such as edibles, sensory and perceptual objects are used initially but soon replaced by social reinforcers such as praise, tickles, hugs, and kisses. Ongoing data collection is performed to monitor skill acquisition, generalization, and frequency of problem behaviors. The intervention protocol consists of 3 primary components: (1) In-home 1:1 instruction, (2) peer play training, and (3) regular education classroom inclusion. No aversive interventions were used throughout the study.

Initially, the In-Home 1:1 Intervention Component is implemented 35 to 40 hours per week for children older than 3 years, and 20 to 30 hours per week for children younger than 3 years. The focus is on establishing foundational and spontaneous communication. The main teaching format is discrete trials,²² but generalization activities and community outings are also part of the 35 to 40 hours per week of instruction. In discrete trials, the tutor works individually with a child in a distraction-free setting and administers 3 to 8 trials in a sitting, with 1- to 2-minute breaks between sittings, for approximately 50 minutes each hour. The remaining 10 minutes of each hour are devoted to generalization activities. These activities include structured play, in which the child has opportunities to apply skills initially mastered in the 1:1 setting (e.g., labeling toys or taking turns with the tutor during a game), and incidental teaching, in which situations were arranged to encourage initiation of language (e.g., placing preferred objects in sight but out of reach). Skill mastery in discrete trials was defined as 90% accuracy across 2 days of intervention, across 2 or more tutors. Concept mastery was defined as 90% accuracy of 5 to 10 novel items probed and mastered within a concept. After mastery, skills and concepts were

systematically generalized to other more naturalistic settings and maintained by available contingencies in the natural environment. To facilitate generalization, community outings occurred 3 to 5 times per week. The UCLA curriculum was used for teaching the initial foundation skills including compliance, imitation, early receptive and expressive language, visual spatial skills, and self-help.^{6,20}

At approximately 1 year into the behavioral intervention, the distribution of the 35 to 40 hours per week is typically as follows: 26 to 31 of home instruction, 3 to 5 hours of peer play, and 6 to 9 hours at preschool. Thereafter, the home component gradually decreases, whereas other components gradually increase based upon the child's inclusion in the classroom.

As part of the generalization of skills and behaviors to the natural environment, the peer play component is initiated 3 to 5 sessions per week with a typically developing peer for 15 to 60 minutes per session when the child has mastered prerequisite skills: verbal response to questions, on topic statements, simple play skills, and turn taking.^{2,6,20} Skills mastered in the 1:1 setting are systematically generalized to a social/play setting with a peer of similar age. A trained tutor facilitates mastered activities for the child and peer (e.g., conversation, pretend play with toys, or turn-taking games) and prompts the peer to engage the child with subtle cues such as whispers in the peer's ear, visual signals, or indirect questions. When the child is 90% accurate initiating with peers across 3 or more peers for 18 to 24 months, additional children are presented at one time to form a group play setting.

At about the time that peer play training is initiated, the child enters a teacher-directed structured regular education preschool setting.² Initially, trained tutors accompany the child to school to assist the teaching staff with gaining instructional control, generalizing mastered skills to the school setting, and learning classroom skills. The tutor functions as a classroom aide and not as a 1:1 aide for the child. Initial goals for inclusion center on generalizing skills to a novel, yet structured environment. As the child achieves independent responding during specific activities (e.g. circle time, center time, and so forth), as determined by data, the shadow tutor is faded. Activities requiring social skills and behaviors are always the last to fade in the process.

When children have achieved typical levels of academic functioning in the classroom and participate without the assistance of a shadow tutor during teacher-directed activities, they still may require the assistance of the shadow tutor during social opportunities throughout the school day for an additional 2 to 3 years. Thus, an intervention with reduced hours both at home and in school may extend into the early primary grades. School hours focus on generalization of social skills and friendship development. As the child's rate of independent social interaction increases, the intervention hours are successively reduced to 0. Subsequently, consultation to the family and the school setting continue 1 to 2 hours per month for up to 1 to 2 years. Home hours focus on play sessions with peers and gradually transition to typical play dates with peers without the presence of a tutor. Periodic

standardized assessments continue until the child is 18 years old.

During the course of the study, there was a growing recognition that many children who made significant gains in the first 2 years of treatment required training beyond the UCLA curriculum to develop mutually satisfying social relationships, enhance their understanding of social meanings, understand and interpret other's perspectives/knowledge/cognition/beliefs, and ultimately respond appropriately to social behaviors of peers and others. To address this need, overt social behaviors were operationally defined, both verbal (e.g., conversational skills, such as responding to statements or questions asked by others, reciprocal statements, initiating conversation, inquiring about others, remaining on topic, and sustaining conversation) and non-verbal (e.g., interpreting and responding to other's facial expressions, emotional states, voice tone, or body language), and initially taught in a discrete trial format, using the same behavioral principles and methodology described above, with an emphasis on a quick transition to generalized teaching to a social context, using incidental teaching and video modeling as tools for generalization.

Staff and Parent Training. To ensure proficiency in implementing the UCLA model, 5 CVAP staff members each completed 3- to 4-month internships at UCLA, and consultants from UCLA made on-site visits 2 to 4 times per year for the first 3 years of the study period, with frequent telephone contacts between visits (typically once per week). During this period, a random sample of 12 CVAP tutors were videotaped and scored by blind raters for adherence to UCLA procedures. The level of adherence by CVAP tutors was found to be nonsignificantly higher than adherence by tutors employed at UCLA.²³

One UCLA-trained individual served as CVAP site director, responsible for oversight of each child's intervention; she holds a master's degree in clinical psychology/applied behavior analysis and is a Board Certified Behavior Analyst. Clinic supervisors trained and provided ongoing performance feedback to tutors. Supervisors were graduate students in behavior analysis or master's level clinicians with 2 or more years of experience in providing EIBT. Tutors were recruited from the community and were the main providers of direct services. Supervisors and tutors were assigned to each EIBT participant based on openings in their schedule and geographic location.

To become a supervisor, individuals had to meet prespecified, objective criteria, including high ratings based on direct observation of their implementation of EIBT interventions, favorable evaluations from families and staff members, satisfactory performance on a test of skill at curriculum development, and oral and written demonstration of their knowledge of applied behavior analysis and ASD.²⁴ Tutors had to pass a rigorous behavior observation assessment of their accuracy in conducting discrete trial training (DTT) and oral tests of their knowledge of the UCLA treatment manual.

Parents were encouraged to be involved in all levels of intervention. At the beginning of treatment, all parents attended a 12- to 18-hour training workshop across 2 to 3 days on behavioral principles and intervention methods. Thereafter, they participated in weekly training sessions to

generalize their child's newly established skills to the natural environment. Parents provided ongoing information regarding their child's current level of functioning both in and out of intervention sessions, and they were asked to be active participants in their child's intervention, although there was no requirement for parents to provide any direct intervention hours.

Treatment Procedures: Comparison Group

Participants in the Comparison Group received community services that their families selected from the Matrix of Educational Options. At intake, 1 comparison child, under 3 years old, received an Early Start Autism Intervention Program, which emphasized learning readiness skills with both the parent and child. This child received less than 9 hours per week of a discrete trial program in his or her home, until the age of 3. Two comparison children received a home-based developmental intervention that ranged from 1 to 4 hours a week. At age 3, these 3 children were enrolled in a public school Special Day Class (SDC). Seventeen children who were 3 and above at intake were enrolled in SDC in the public schools. No records were available for 1 child. The instructional methodology in the SDC placements was eclectic, the child/teacher ratios varied from 1:1 to 3:1, and the classes operated for 3 to 5 days per week, for up to 5 hours per day. Related services such as speech, occupational, and behavioral therapy to these children varied from approximately 0 to 5 hours per week. Three of the children spent brief sessions (up to 45 minutes per day) mainstreamed in regular education. Due to the diverse interventions provided to the comparison group, it was not possible to monitor treatment fidelity for this group.

Assessment

At pretreatment, a licensed psychologist at EADC who was independent of the study administered a standardized behavior observation,²⁵ parent interview, and developmental tests, including the BSID-R, Merrill-Palmer Scale of Mental Tests,²⁶ Reynell Developmental Language Scales,²⁷ and Vineland Adaptive Behavior Scales.²⁸ The BSID-R extrapolated table was used to generate a standard score for children who obtained an IQ below 50.²⁹ Administration of the BSID-R began at the starting point for the child's chronological age (or at the highest starting point for the test if the child was older than 42 months). The examiner administered each successive item after the starting point to establish a basal and ceiling; if the child did not obtain a basal on these items, the examiner administered each preceding item in succession until a basal was achieved and then followed rules in the test manual for establishing the ceiling.

From the evaluation, the psychologist made a DSM-IV diagnosis of autism or Pervasive Disorder, Disorder Not Otherwise Specified (PDDNOS).³⁰ Subsequently, the diagnosis was confirmed by the Autism Diagnostic Interview-Revised (ADI-R),¹⁷ administered by a certified examiner employed by CVAP. The developmental tests (but not the ADI-R) were repeated in annual follow-up evaluations. If

a participant performed at the ceiling of the BSID-R, this test was replaced with the Wechsler Preschool and Primary Scales of Intelligence.³¹ Follow-up evaluations were conducted by an independent, self-employed, highly-skilled, licensed, child evaluator. VMRC made the referral and funded the evaluations. The referral to the evaluator consisted only of the name of the child, birth date, parent's names, and telephone number.

Data Analysis

IQ was the main measure of treatment response in previous EIBT studies⁶⁻¹⁶ and was designated as the primary outcome measure in the present study. Secondary outcome measures were the Merrill-Palmer Scale of Mental Tests, Reynell Language Comprehension, Reynell Expressive Language, Vineland Adaptive Behavior Scales, and classroom placement.

To test our main hypothesis that the EIBT group would differ from the comparison group on outcome measures, we performed a repeated-measures analysis of covariance (ANCOVA) for each measure, with pretreatment score as the covariate and Year 1, Year 2, and Year 3 scores as the repeated dependent measures. Consistent with standard assumptions for an ANCOVA,³² analyses of skew and kurtosis, as well as visual inspection, were consistent with a normal distribution in our data. Hyunh-Feldt epsilon tests confirmed that the data showed compound symmetry ($\epsilon > .90$), unless otherwise noted in Results.

As is usual in outcome studies with repeated measures, a few participants had missing data at one or more time points. For each outcome measure, we employed the standard procedure of removing participants with missing data from the analysis.³² This procedure is appropriate when missing data are random or unbiased. We used visual inspection to confirm that the missing data were unbiased (e.g., the data were not primarily from participants who had unfavorable outcomes or who did not complete the full 3 years of intervention), and *Results* show the number of participants retained for each analysis.

In as much as the EIBT and comparison groups differed on several demographic variables (mother education, father education, and diagnosis), we explored whether adding these variables as covariates in the ANCOVA model would change the interpretation of the results. These analyses need to be interpreted with caution because they involve a larger number of variables than is usually considered appropriate for the relatively small sample size in the present study. However, they provided some information on whether or not the groups differed when we statistically controlled for demographic variables.

When an ANCOVA revealed a between-group difference on an outcome measure, we hypothesized that the EIBT group would show an increase in scores from Year 1 to Year 2 to Year 3, whereas scores in the comparison group would remain stable. To test this hypothesis, we examined whether the ANCOVA yielded a statistically significant Group \times Time interaction; if so, we performed planned comparisons to test for an increase from Year 1 to Year 3 in the EIBT group.

Table 1. Background Information for the EIBT Group (n = 21) and Comparison Group (n = 21)

	EIBT	Comparison
Demographics		
Male/Female	18:3	17:4
Diagnosis (Autism/PDDNOS)*	20:1	15:6
Age at diagnosis [(M(SD))]	30.2 (5.8)	33.2 (3.7)
Mother education, yr [(M(SD))]*	15.3 (2.9)	13.1 (1.6)
Father education, yr [(M(SD))]*	15.8 (2.9)	11.8 (2.3)
Two-parent household (yes/no)*	21:0	14:7
Pretreatment Test Scores [(M(SD))]		
IQ	61.6 (16.4)	59.4 (14.7)
Merrill-Palmer	82.4 (17.3)	73.4 (11.9)
Reynell		
Language Comprehension	51.7 (15.2)	52.7 (15.1)
Expressive Language	52.9 (14.5)	52.8 (14.4)
VABS		
Composite	69.8 (8.1)	70.6 (9.6)
Communication	69.4 (11.8)	65.0 (6.8)
Daily Living	73.2 (9.2)	72.7 (12.5)
Socialization	70.3 (10.9)	75.1 (13.0)

EIBT indicates early intensive behavioral treatment; Reynell, Reynell Developmental Language Scales; VABS, Vineland Adaptive Behavior Scales; PDDNOS, Pervasive Disorder, Disorder Not Otherwise Specified.

*Significant difference between EIBT and comparison group ($p < .05$).

To examine the clinical significance of the results, we ascertained the number of participants in each group who achieved scores in the average range at follow-up on each measure. We also sought to identify pretreatment measures that were associated with later scores in the average range. Therefore, for the EIBT group, we conducted *t*-tests to compare pretreatment scores of participants who scored in the average range across all measures to pretreatment scores of the remaining participants.

RESULTS

Pretreatment

Table 1 summarizes the demographics and pretreatment scores of the early intensive behavioral treatment

(EIBT) and comparison groups. The gender make-up mirrors the 4:1 male to female ratio in Autistic Spectrum Disorder (ASD).³¹ Twenty of 21 EIBT children (95%) and 15 of 21 comparison children (71%) were diagnosed with Autistic Disorder. This difference was statistically significant, $t(40) = 2.13$, $p < .05$. The remaining children were classified with Pervasive Disorder, Disorder Not Otherwise Specified (PDDNOS). Age of diagnosis was 20 to 41 months, with the EIBT group averaging 3 months younger than the comparison group (a difference that was not statistically significant). Also, as shown in Table 1, although not a requirement for participation in the EIBT program, parents had significantly more education and were significantly more likely to be married than comparison parents. IQ, Merrill-Palmer, Reynell, and Vineland scores did not differ significantly between groups; scores in both groups indicated developmental delays comparable to other samples of children with ASD.³⁰

Outcome

Table 2 presents the results of the analysis of covariance (ANCOVA) tests for each outcome measure, whereas Figure 1 presents the means and 95% confidence intervals for each group at intake, Year 1, Year 2, and Year 3. As shown in Table 2, there was a significant difference between groups on the primary outcome measure, IQ. Figure 1 reveals that the mean IQ in the EIBT group increased 25 points, from 62 at pretreatment to 87 at Year 3. Interestingly, the mean IQ in the comparison group also increased, from 59 at pretreatment to 73 at Year 3.

The EIBT and comparison groups did not differ significantly on the Merrill-Palmer. Both groups displayed a mean increase of 13 points from intake to Year 3 on this measure. Figure 1 suggests that the groups may not have been matched at pretreatment, as the mean for the EIBT was 82 compared to 73 in the comparison group. A post hoc analysis indicated that this difference approached statistical significance, $t(35) = 1.87$, $p = .07$. Also, the assumption of compound symmetry was questionable for this variable, with Huynh-Feldt $\epsilon = .85$; because the

Table 2. Analyses of Covariance Testing for Differences Between the EIBT and Comparison Groups on Outcome Measures

Measure	N		Sums of Squares (Between Subjects)			MSE	F
	E	C	Group	Covariate	Error		
IQ	21	19	4,229.91	12,046.14	30,042.41	811.96	5.21*
Merrill-Palmer	21	16	246.27	15,613.74	20,657.91	626.00	ns
Reynell							
Language Comprehension	21	19	3,750.25	17,523.60	36,312.08	981.41	3.82**
Expressive Language	20	19	3,413.57	13,590.90	52,495.66	1,458.21	ns
VABS							
Composite	20	20	3,897.52	1,589.31	18,385.69	496.91	7.84***
Communication	20	20	3,937.71	2,937.53	25,994.10	722.06	5.45*
Daily Living	20	20	2,527.14	2,229.25	14,207.49	394.65	6.40*
Socialization	20	20	1,857.84	21.66	16,130.41	460.87	4.03**

N indicates number of participants included in the analysis; E, EIBT group; C, comparison group; ns, not statistically significant; MSE, mean square of errors (between subjects); Reynell, Reynell Developmental Language Scales; VABS, Vineland Adaptive Behavior Scales.

* $p < .05$; ** $p < .10$; *** $p < .01$.

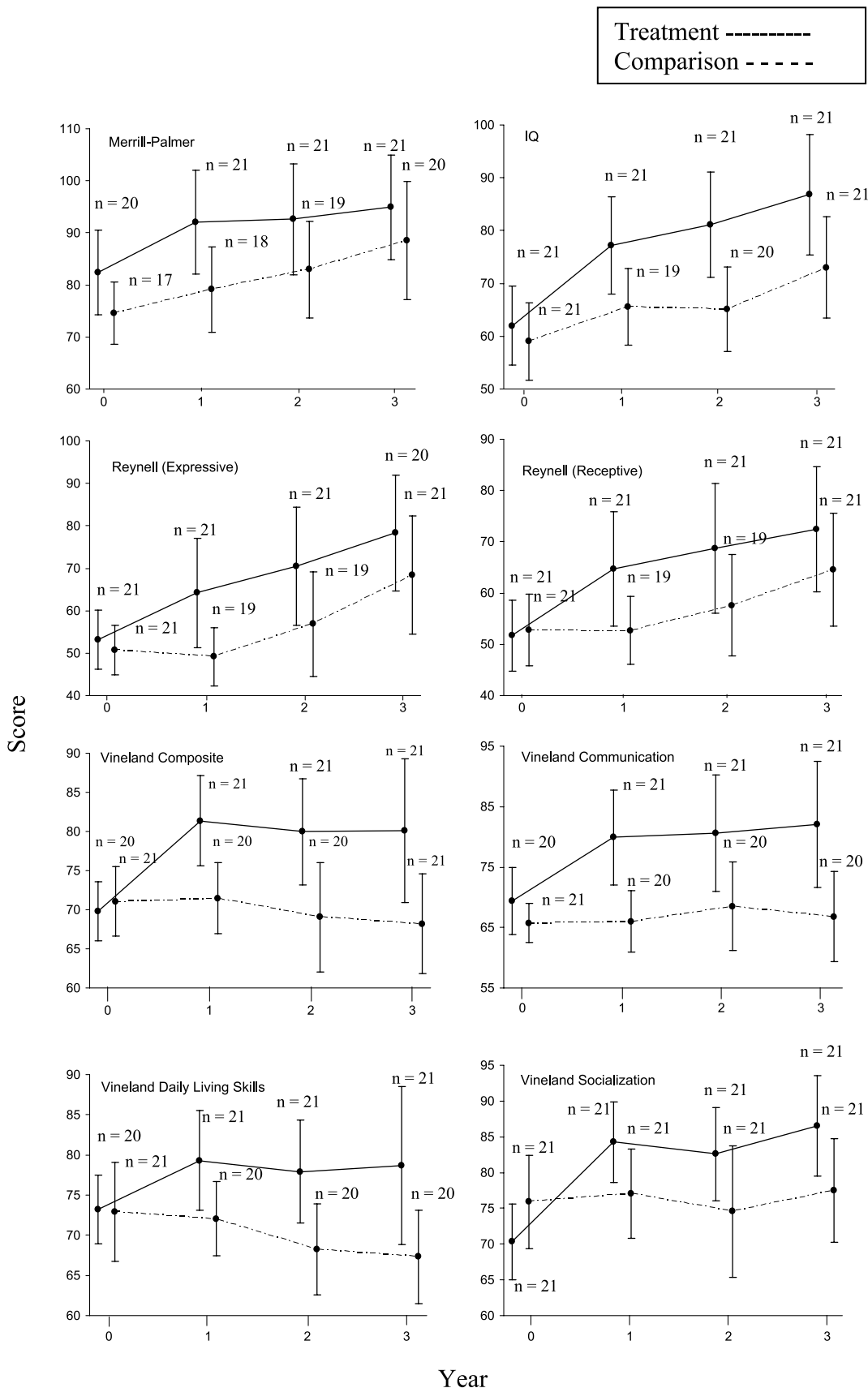


FIGURE 1. Mean and 95% confidence interval for pretreatment (Year 0) and follow-up (Years 1–3).

Table 3. Number of Children in the Average Range on each Outcome Measure for the EIBT Group (n = 21) and Comparison Group (n = 21)

Measure	EIBT	Comparison	<i>p</i>
IQ	12	7	ns
Language Comprehension ^a	8	4	ns
Expressive Language ^a	9	6	ns
VABS Composite ^b	8	3	.10
School Placement	6	0	.001

^aReynell Developmental Language Scales.

^bVineland Adaptive Behavior Scales.

ANCOVA did not approach statistical significance, alternate analyses were not attempted.

There was a trend toward a significant difference in Reynell Language Comprehension ($p = .06$). The mean score in the EIBT group increased 20 points, from 52 at pretreatment to 72 at Year 3; the mean score in the comparison group increased 9 points, from 53 at pretreatment to 62 at Year 3. The EIBT group also had a larger increase from pretreatment to Year 3 in Reynell Expressive Language (53–78, compared to 51–66), but this difference was not statistically significant ($p = .13$). The failure to find a significant difference may indicate that EIBT did not have a meaningful effect on expressive language, or it may simply reflect low statistical power to detect an effect.

The EIBT and comparison groups differed significantly in the Vineland Adaptive Behavior Scales Composite. Consistent with this finding, the EIBT group demonstrated a mean increase of 9 points compared to a 4-point decline in the comparison group, as shown in Figure 1. Inasmuch as a difference was observed in the Composite, individual scales were also analyzed. Significant differences between groups were found in Communication and Daily Living Skills, and a trend was found for Socialization ($p = .05$). Figure 1 indicates that the changes in scores from pretreatment to Year 3 for each scale were similar to the change in Composite scores. These findings support the inference that the EIBT group had more advanced adaptive behavior skills than the comparison group at the time of the outcome assessments.

An analysis of classroom placement at year 3, between the 2 groups, revealed that 17 of the 21 EIBT children and 1 of the 21 comparison children were included into regular education classroom settings. Of the 17 EIBT children, 6 were fully included without assistance, 4 were fading the shadow tutor, and 7 required full shadows.

When mother's education, father's education, or diagnosis was added as a covariate to the ANCOVA model, ANCOVA was unaltered, except in one instance: With the father's education as a covariate, the difference between groups in IQ was not statistically significant ($p = .11$). It is unclear whether this finding indicates that father's education was a confound or reflects the limited statistical power for the analysis. When mother's education, father's education, and diagnosis were all added as covariates to the ANCOVA model, IQ, Reynell Language Comprehension,

and Vineland Composite continued to show a trend toward significance ($p = .09$ for all 3 outcome measures). In sum, the possibility that father's education was a confound in the analysis of IQ cannot be ruled out, but the remaining analyses indicated that reliable differences in outcome between groups remained after statistically controlling for inequalities at pretreatment.

None of the analyses for group \times time interactions were statistically significant. Thus, we did not confirm our hypothesis that the EIBT group would have increasing scores from Year 1 to Year 2 to Year 3, whereas scores in the comparison group would be stable. On the contrary, Figures 1 and 2 illustrates that although the EIBT group appeared to make larger increases than the comparison group from pretreatment to Year 1, both groups exhibited stable scores from Year 1 to Year 3 in IQ, Merrill-Palmer, and Vineland. Both groups may have exhibited similar increases in scores in Reynell Language Comprehension and Expressive Language from Year 1 to Year 3.

As shown in Table 3, more EIBT participants than comparison participants achieved follow-up scores in the average range for each measure, although this difference was significant only for school placement and showed a trend toward significance for the Vineland. Ten EIBT participants scored in the average range on all measures (6 of these 10 also were included in regular education without assistance, whereas the remaining 4 continued to receive shadowing in the regular education classroom). *t*-tests did not reveal any significant differences in pretreatment test scores for these 10 participants compared to the remaining 11 participants. For example, these 10 children had a mean pretreatment IQ of 66.6 (SD = 12.4) compared to 57.7 (SD = 19.0) for the remaining 11 children, $t(19) = 1.28$, ns. However, pretreatment Reynell Language Comprehension scores showed a trend toward a difference, with a pretreatment mean of 58.1 for the participants with the most favorable outcome compared to 45.9 for the other participants, $t(19) = 1.98$, $p = .06$.

DISCUSSION

The present study suggests that the UCLA/Lovaas Model of early intensive behavioral treatment (EIBT) can be implemented in a nonuniversity community-based setting. On the primary outcome measure of IQ, the EIBT group showed a gain of 25 points, which was statistically significant compared to the gain of 14 points in the comparison group. Similar effects were found on measures of adaptive behavior. Although language comprehension showed a trend towards significance, expressive language and nonverbal cognitive skill revealed no difference between groups. The increases in test scores are similar to those reported in Lovaas' original EIBT study^{2,3} and in some recent investigations.^{15,16} However, the difference between the EIBT group and the comparison group on outcome measures was smaller than that in other studies, as the comparison group also made gains.

An important limitation of the study is that, because treatment was funded by public agencies that were required to offer free and appropriate services, groups could not be randomly assigned, and a quasi-experimental design was used, with parents choosing which group their child entered. Although pretreatment test scores did not differ significantly between groups, other pretreatment variables did differ. The EIBT group had more children with autism and fewer with Pervasive Disorder, Disorder Not Otherwise Specified (PDDNOS) than did the comparison group. To the extent that PDDNOS is a milder diagnosis that may have a more favorable prognosis than autism,⁷ this difference may have favored the comparison group. However, the EIBT group also may have had an advantage in that it had more 2-parent families and better educated families than did the comparison group. These family variables have not been associated with outcome in previous studies,^{2,7} but they might have encouraged families to select EIBT over other interventions in the present study, even though all interventions were provided at no cost to families. In addition, these variables might have given the EIBT group an advantage by making it easier for families to participate in treatment sessions and facilitate generalization of skills outside of treatment. After statistically controlling for family variables, outcome analyses continued to show improved outcomes in the EIBT group relative to the comparison group. Nevertheless, statistical controls are not a satisfactory solution for preexisting group differences, especially given the relatively small sample size in the present study. A design with random assignment would have strengthened the study and allowed for more clearcut conclusions about whether EIBT is effective or not.

Further limitations pertain to the assessment protocol in the study. As previously noted, the comparison group received such diverse interventions that a measure of treatment fidelity could not be applied. Also, outside evaluators were employed by Valley Mountain Regional Center (VMRC) for pretreatment and follow-up assessments of participants. The referrals to the evaluators did not include information on group assignment or treatment history. However, to ensure that evaluators remained unaware of this information and to allow for checks on the reliability of test administration and scoring, evaluators who were employed by the study and conducted assessments at a research site (rather than in their clinical offices) might have been preferable. Another limitation is that the assessment protocol tested developmental level more rigorously than did the features of Autistic Spectrum Disorder (ASD). The inclusion of the Autism Diagnostic Observation Schedule (ADOS),³³ in addition to the Autism Diagnostic Interview–Revised (ADI-R) and clinical diagnosis, would have increased confidence in the initial diagnosis. Including a measure such as the ADOS in follow-up assessments would have indicated whether or not children continued to display behaviors indicative of ASD. Additional measures such as the Theory of Mind Test³⁴ also would help address this issue; Central Valley Autism Project (CVAP) is currently involved in a study to translate this test into English and standardize it in the

United States. Without such measures, the present study cannot address one of the most controversial issues raised by previous EIBT research—whether some children become indistinguishable from typically developing peers⁶ or whether they continue to display characteristics of ASD. An additional follow-up evaluation of study participants with the ADOS and Theory of Mind (TOM) Test is planned to fill in some of these gaps.

In this study, advanced behaviors associated with friendship initiation and maintenance, social skills, understanding of social meaning, and response to social behaviors were identified and treated, using the same discrete trial methodology as other behaviors, which consequently increased the duration of treatment beyond 3 years for many participants (usually for 2 additional years). Although this expansion of the treatment protocol reflects the contemporary view that the defining feature of ASD is an impairment in social reciprocity, it raises the question of whether the present study truly was a replication of the UCLA model. The treatment site met all of Lovaas' criteria for replication, and the first 2 years of intervention followed the model as it has been previously described.² The third year also followed the model, with the addition of the training in advanced social skills. Thus, results from Years 1 and 2 are directly comparable to those of previous studies, and results from Year 3 also reflect mostly the same interventions. Research on the specific effects of the additional social skills training is warranted, as it is acknowledged that such training was not included in previous studies. Also, although discrete trial training is a common approach to teaching social skills and has some empirical support,^{35,36} teaching methodologies other than discrete trials (e.g. video modeling, incidental teaching) also have empirical support and may have advantages such as generalizing more quickly to settings outside of treatment;²² thus, the question of how best to teach such skills may be another area for research.

Interestingly, although the EIBT protocol lasted for 3 years and, in some cases, was continued beyond that time, the nonsignificant group \times time interactions in the statistical analyses indicates that the EIBT group did not show reliable IQ increases relative to the comparison group after Year 1. A possible explanation is that most gains occurred in the first year of intervention. Alternatively, however, it is also possible that gains took place later in treatment but that the study measures were not sensitive to them.

Potential evidence for the latter view comes from the findings on classroom placement. A striking result was that, despite IQ gains in the comparison group, all participants but 1 remained primarily in a special education classroom setting, whereas most EIBT participants were included in regular education at least part of the day. Classroom placement is a controversial outcome measure because of concerns that it may reflect factors such as parent advocacy and school policy rather than the child's functioning.¹² However, the measure also may be an index of real-world academic and social competence.³⁷ If so, the differences between groups on this measure may be

attributable at least in part to the social skills training that EIBT participants received. In addition, it may suggest a need for a high number of treatment hours. Dismantling studies might help address these possibilities.

The initial collaborative funding efforts by VMRC and Special Education Local Planning Areas (SELPA) resulted in a sustainable treatment environment. Stable funding, effective guidelines and policies, and positive communication and working relationships were primary contributory variables to the feasibility of this study. Thus, this collaboration may be a useful model for other regions to employ. Other clinical strengths of this study included rigorous treatment quality control criteria, stringent staff training and evaluation standards, multiple internships at UCLA by supervising clinicians, precise programming for each individual child, advanced completion programming and skilled generalization training, yearly follow-ups by an independent evaluator using multiple outcome measures,

and a centralized process and standardized protocol for diagnosing children and informing families of EIBT and other intervention options available to them. Without such standards, outcomes may differ. Nevertheless, given the methodological limitations of the present research, there is a continued need for rigorous outcome studies comparing EIBT to control conditions or other interventions.

Acknowledgments. Preparation of this manuscript was partially supported by NIMH grants R01 MH 48663 (Multi-site Young Autism Project) and U54 MH066397 (Genotype and Phenotype of Autism). The authors thank Kym Cassaretto, M.S., CVAP Clinical Director, and CVAP staff for their clinical contributions and committed efforts on behalf of children and families. The authors also thank the following individuals for their assistance with the research: Mieke San Julian, Chanti Fritzsching, M.S., BCBA, and Angela Castro from CVAP; Schelley McDonald and Marie Overmyer from VMRC; and Suzannah Ferraioli from URCM. Preliminary reports were presented at the California Association for Behavior Analysis Conference in San Francisco, CA, February 20, 2004, and at the International Meeting For Autism Research, Boston, MA, May 6, 2005.

REFERENCES

- DeMyer MK, Hingtgen JN, Jackson RK. Infantile autism: a decade of research. *Schizophr Bull.* 1981;7:388–451.
- Lovaas OI. Behavioral treatment and normal educational and intellectual functioning in young autistic children. *J Consult Clin Psychol.* 1987;55:3–9.
- McEachin JJ, Smith T, Lovaas OI. Long-term outcome for children with autism who received early intensive behavioral treatment. *Am J Ment Retard.* 1993;97:359–372.
- Zwaigenbaum L, Bryson S, Rogers T, et al. Behavioral manifestations of autism in the first year of life. *Int J Dev Neurosci.* 2005; 23:143–152.
- California Department of Developmental Services. Department of Developmental Services Fact Book, 7th ed. California Department of Developmental Disabilities website. December, 2004. Available at www.dds.ca.gov/factsstats/factbook.cfm#pdf. Accessed August 16, 2005.
- Lovaas OI. *Teaching Individuals with Developmental Delays: Basic Intervention Techniques.* Austin, TX: PRO-ED; 2003.
- Smith T, Groen AD, Wynn JW. Randomized trial of intensive early intervention for children with pervasive developmental disorder. *Am J Ment Retard.* 2000;105:269–285.
- Anderson SR, Avery DL, DiPietro EK, et al. Intensive home-based early intervention with autistic children. *Educ Treat Child.* 1987; 10:352–366.
- Birnbrauer JS, Leach DJ. The Murdoch Early Intervention Program after two years. *Behav Change.* 1993;10:63–74.
- Harris SL, Handleman JS. Age and IQ at intake as predictors of placement for young children with autism: a four- to six-year follow-up study. *J Autism Dev Disord.* 2000;30:137–142.
- Weiss M. Differential rates of skill acquisition and outcomes of early intensive behavioral intervention for autism. *Behav Interv.* 1999;14:3–22.
- Schopler E, Short A, Mesibov G. Relation of behavioral treatment to “normal functioning”: comment on Lovaas. *J Consult Clin Psychol.* 1989;57:162–164.
- Bibby P, Eikeseth S, Martin NT, et al. Progress and outcomes for children with autism receiving parent-managed intensive interventions. *Res Dev Disabil.* 2002;23:81–104.
- Smith T, Buch GA, Gamby TE. Parent-directed, intensive early intervention for children with pervasive developmental disorder. *Res Dev Disabil.* 2000;21:297–309.
- Howard JS, Sparkman CR, Cohen HG, Green G, Sanislaw HA. Comparison of intensive behavior analytic and eclectic treatments for young children with autism. *Res Dev Disabil.* 2005;26:359–383.
- Sallows GO, Graupner TD. Intensive behavioral treatment for children with autism: four year outcome and predictors. *Am J Ment Retard.* 2005;110:417–438.
- Lord C. Follow-up of two-year-olds referred for possible autism. *J Child Psychol Psychiatry.* 1995;36:1365–1382.
- Bayley N. *Bayley Scales of Infant Development*, 2nd ed. San Antonio, TX: The Psychological Corporation; 1993.
- Cohen HG. Pyramid building: Partnership as an alternative to litigation. In: Lovaas OI ed. *Teaching Individuals with Developmental Delays: Basic Intervention Techniques.* Austin, TX: PRO-ED, 2003:375–386.
- Lovaas OI. *Teaching Developmentally Disabled Children: The ME Book.* Austin, TX: PRO-ED; 1981.
- Region 6 Autism Connection. *Early Intensive Behavioral Treatment 4-way Agreement.* Stockton, CA: Region 6 Autism Connection; 2004.
- Smith T. Discrete trial training in the treatment of autism. *Focus Autism Relat Disord.* 2000;16:86–92.
- Mortenson S, Smith T. Quality Control in the Multisite Young Autism Project. Paper presented at: Annual Meeting of the Association for Behavior Analysis; May 1996; San Francisco, CA.
- Davis BJ, Smith T, Donahoe P. Evaluating supervisors in the UCLA treatment model for children with autism: validation of an assessment procedure. *Behav Ther.* 2002;31:601–614.
- California Department of Developmental Services. *Best Practice Guidelines for Screening, Diagnosis, and Assessment. Ethological Observation Schedule (ETHOS).* Sacramento, CA: California Department of Developmental Services; 2002.
- Stutsman R. *Guide for Administering the Merrill-Palmer Scale of Mental Tests.* New York: Harcourt, Brace & World; 1948.
- Reynell JK. *Reynell Developmental Language Scales.* Windsor, England: Nfer-Nelson; 1990.

28. Sparrow SS, Balla DA, Cicchetti DV. *Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Service; 1984.
29. Robinson BF, Mervis CB. Extrapolated raw scores for the second edition of the Bayley Scales of Infant Development. *Am J Ment Retard*. 1996;100:666–671.
30. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed, Text Revision. Washington, DC: American Psychiatric Association; 2000.
31. Wechsler D. *Manual for the Wechsler Intelligence Scale for Children*, 3rd ed. San Antonio, TX: Psychological Corporation; 1991.
32. Nich C, Carroll K. Now you see it, now you don't: a comparison of traditional versus random-effects regression models in the analysis of longitudinal follow-up data from a clinical trial. *J Consult Clin Psychol*. 1997;65:252–261.
33. Lord C, Rutter M, DiLavore PC, et al. *Autism Diagnostic Interview Schedule*. Los Angeles: Western Psychological Services; 2001.
34. Steerneman P, Meesters C, Muris P. *TOM-Test*. Antwerpen-Appeldoorn: Garant; 2003.
35. Taylor BA, Jasper S. Teaching programs to increase peer interaction. In: Maurice C, Green G, Foxx M eds. *Making a Difference: Behavioral Intervention for Autism*. Austin, TX: Pro-Ed; 2001: 97–162.
36. Weiss MJ, Harris SL. Reaching out, joining. In: *Teaching Social Skills to Young Children with Autism*. Bethesda, MD: Woodbine House; 2001.
37. Kazdin A. Replication and extension of behavioral treatment of autistic disorder. *Am J Ment Retard*. 1993;97:382–383.