School Mental Health (2013) 5:221–232
DOI 10.1007/s12310-013-9103-7

ORIGINAL PAPER

A School-Based Preschool Program for Children with ASD: A Quasi-Experimental Assessment of Child Change in Project DATA

Ilene Schwartz · Carnot J. Thomas · Bonnie McBride · Susan Sandall

Published online: 16 March 2013

Abstract Project DATA (Developmentally Appropriate Treatment for Autism) is a school-based comprehensive treatment model for preschool children with autism spectrum disorder (ASD). This study explored 5 years of results from children participating in the model. Data from up to 69 preschool children independently diagnosed with ASD were utilized in determining pre-/post-program performance on a variety of outcome measures representative of functioning across developmental domains, receptive language skills, social skills, problem behaviors, and autism characteristics. Descriptive statistics are utilized to describe the participants at program entry and exit. Results of the paired-samples *t* tests indicate statistically significant positive program change on 13 of the 16 outcomes measures, especially those that measure core deficits common to ASD. Overall results suggest that Project DATA is a promising practice for providing school-based services to young children with ASD.

Keywords Autism · Comprehensive treatment model · Early intervention

I. Schwartz (✉) · S. Sandall
University of Washington, Seattle, WA, USA
e-mail: ilene@uw.edu

C. J. Thomas
Vashon Island School District, Vashon Island, WA, USA

B. McBride
University of Oklahoma Health Science Center, Oklahoma City, OK, USA

Introduction

Diagnoses of autism spectrum disorders (ASD) are on the rise in the USA, with prevalence rates estimated at approximately 1 in every 88 children in the USA (Autism and Developmental Disabilities Monitoring Network, 2012). ASD are more common in boys, with a prevalence rate of 1 in 54 compared to that in girls of 1 in 252. Although we do not know for sure why there are so many children with ASD being diagnosed, the functional outcome for early intervention providers and public schools is the same. There are greater numbers of children with diagnoses of ASD requiring services from our publicly funded systems, and these children and their families are beginning their services at younger ages. Since 2000, the number of students aged 3–21 years served under the diagnosis of ASD within the public early intervention and education system has increased from 93,650 students to 297,739 students, more than 300 % increase in less than 8 years (OSEP, 1995–2007).

While the incidence of people served under the ASD eligibility category affects all aspects of publicly funded programs, education, specifically public schools, may be absorbing the bulk of this increase. A free, appropriate public education, provided in the least restrictive environment possible, is an entitlement for every student with a disability in the USA. While the incidence of students served under the autism eligibility has skyrocketed across all ages served, the 3- to 5-year-old or preschool population has displayed the most rapid rate of growth (OSEP, 1995–2007). With this increase in students diagnosed with ASD enrolled in public services, the demand for high-quality educational programming for students with ASD has become a national concern. This need for high-quality, early childhood educational opportunities is highlighted by

research findings indicating that young children with ASD respond extremely well to high-quality, early intensive behavioral intervention (e.g., Dawson & Rogers, 2010; Levy, Kim, & Olive, 2006; Lovaas, 1987; Reichow & Wolery, 2009; Rickards, Walstab, Wright-Rossi, Simpson, & Reddihough, 2007; Rogers & Vismara, 2008; Strain & Bovey, 2011).

Although there has been an increase in the amount of published research investigating early intervention approaches for children with ASD, to date there has only been one published report of a randomized trial investigating the efficacy of a behavioral, school-based, inclusive model of intervention for children with ASD. Strain and Bovey (2011) evaluated the use of the LEAP model of early intervention in 28 preschool classrooms. The LEAP model is conducted in an inclusive preschool setting. In LEAP classrooms, evidence-based instructional strategies are embedded into ongoing classroom activities and routines. Through a documented system of peer-mediated instruction, the typically developing peers in the classroom play a major role in the intervention process. Intensity of instruction in the LEAP model is not measured by hours per week of intervention, but rather by the number of meaningful opportunities to respond in a context of supportive of natural social interaction. The results of the LEAP trial were extremely positive. Children participating in a high-fidelity application of the LEAP model made significant progress in all areas compared to the students in the business as usual classrooms. This study is important because it is the first evaluation of an early intervention approach implemented within the context of public school preschool classrooms. It is also the first study demonstrating the effectiveness of an intervention program that educates children with ASD alongside their typically developing peers, providing them the opportunity to learn from children with typical social and communicative behaviors and develop relationships with these children.

The published reports of effective programs for young children with ASD have almost exclusively consisted of reports from experimental programs providing one-on-one instruction at home or in clinics (e.g., Dawson et al., 2010; Lovaas, 1987; Sallows & Graupner, 2005). Although these reports have been important to inform and shape the state of the science, they do not necessarily inform or translate easily into the state of the practice. Implementing interventions in public school classrooms requires that the strategies and supports for interventionists be well developed and relatively easy to implement and are responsive to the social and environmental context of schools. This means that interventions must be able to be implemented across a school day and across the different school environments and activities and must not require the child with ASD to be working one-on-one with an adult at all times. The LEAP RCT provides an excellent model of an intervention that is scientifically rigorous and socially valid and demonstrated to result in statistically significant changes in valued behaviors for its participants.

There do exist a number of manualized approaches to working with young children with ASD. Many of these either build upon or were used to shape the recommendations of the National Research Council's Committee on Educational Interventions for Children with Autism (Lord et al., 2001). Often referred to as comprehensive treatment models (CTMs), these approaches attempt to increase students' skills in the deficit areas core to ASD through a host of inter-related educational practices. Odom et al. (2010) operationally defined CTMs by the following six criteria:

1. The model and its components must be described in a refereed publication
2. The model must have a procedural manual
3. The model must have a clear conceptual framework
4. The model must address multiple developmental or behavioral domains representative of the core characteristics of ASD
5. The model must be intensive (e.g., 25 h per week, 9–10 months of the year, with a planned set of activities)
6. The model must have been implemented in at least one site in the USA

Based on these six criteria, the authors identified 30 CTMs currently in use in the USA, and while the majority of these approaches were well operationalized, most had limited fidelity of implementation support and weak evidence of program efficacy. Only four of these CTMs were categorized as using applied behavior analysis and being implemented in inclusive settings. Project DATA (Developmentally Appropriate Treatment for Autism) is one of these (Schwartz, Sandall, McBride, & Boulware, 2004).

Project DATA was developed with assistance from the Department of Education, Office of Special Education (OSEP) Model Demonstration (Schwartz et al., 2004). The goal of the Project DATA CTM is to blend the strengths of different disciplines such as early childhood education, early childhood special education, and applied behavior analysis to develop, implement, evaluate, and replicate a school-based/center-based program for toddlers and preschoolers with ASD that is effective, sustainable, and responsive to the needs of consumers (i.e., families and school district personnel). Currently, the program is in its 17th year and is funded by the local school district. Since 1997, Project DATA has served over 100 preschoolers. Over 50 % of the graduates have made the transition into inclusive elementary settings.

At the core of the Project DATA model is a high-quality, inclusive, early childhood special education program that

includes opportunities and support every day and from the first day of the program to interact successfully with typical peers; functional assessment and evaluation strategies; appropriate curriculum to meet individual student needs; and effective, evidence-based instructional strategies. Project DATA utilizes a data-driven instructional model that involves the use of repeated standardized and curriculum-based measures (CBM) to create the information necessary to monitor progress toward individualized goals and facilitate data-based educational decisions. Additional components of the Project DATA model include additional time at school and explicit instruction to facilitate experiences in the classroom; technical and social support for families; collaboration with other family-negotiated services (i.e., outside service providers such as speech or behavioral therapists); and a quality-of-life-influenced curriculum.

To date, data were collected on individual children who have participated in Project DATA, but there has not been a major evaluation of the project efficacy. Although we do not have access to any comparison group data, the purpose of this study is to evaluate the growth over time on specific outcome measures (both norm-referenced and CBMs) of Project DATA graduates. Although Project DATA met the requirements set by Odom et al. (2010) to be considered a CTM and was in fact reviewed positively by the authors, limited efficacy data hindered the overall rating of the program. Therefore, the purpose of the current research is to provide data from 69 young children who participated in Project DATA to illustrate the effectiveness of the program. Pre- and post-data from measures will be examined to address the effectiveness of this approach to intervention for young children with ASD across valued developmental domains. Although we realize that this does not meet the standard of a randomized clinical trial, these descriptive pre-/post-data will be the first step toward demonstrating the effectiveness of this model.

Method

Participants and Setting

The current study included up to 69 preschool students with independent medical diagnoses of ASD who participated in Project DATA over the previous 5 years. Unlike many comprehensive treatment models (Odom et al., 2010), Project DATA is operated in collaboration with a local public school district. As part of the memorandum of understanding between the University of Washington and the school district, admissions to the program are controlled by the district. This means that any preschool child with a diagnosis of ASD who lived within the school district was eligible to attend Project DATA. No other entry criteria in addition to a diagnosis of ASD and referral by the school district were required for participation. If the school district assigned the student to Project DATA and parents consented, they participated in the project. This "open-door" admission policy was put in place intentionally to reflect the range of students with ASD that public schools are required to serve and, unlike many other programs for young children with ASD, did not require children to be able to reach a target score on standardized assessments at pre-testing. The age of entry for students who participated in Project DATA ranged from 38 to 67 months of age with an average age of 4 years and 5 months. The gender breakdown was 21 % female students and 79 % male students, which accurately reflects the male-to-female identification ratio of 4:1 in diagnoses of ASD in general population (Ehlers & Gillberg, 1993; Kogan et al., 2009). The racial demographics indicated that 80.9 % of the students served were Caucasian, 1.5 % were African-American, 14.7 % were Asian, and 2.9 % were considered "some other race" which would include those of Hispanic origin.

Children enrolled in Project DATA receive approximately 22 h a week of instruction at school. Ten hours a week the children participated in an inclusive preschool classroom with 16 children, 8 of whom were typically developing and 8 of whom had an IEP (no more than 4 of the children in any classroom had a diagnosis of ASD). For 12 h a week, the children participated in more intensive instructional programming. All instructional programming took place at the children's preschool. In addition to the 22 h a week of instruction, families received a monthly home visit. The key components of Project DATA are illustrated in Fig. 1.

There are five components to Project DATA. The core component is participation in, from the first day of intervention, an inclusive preschool program where children

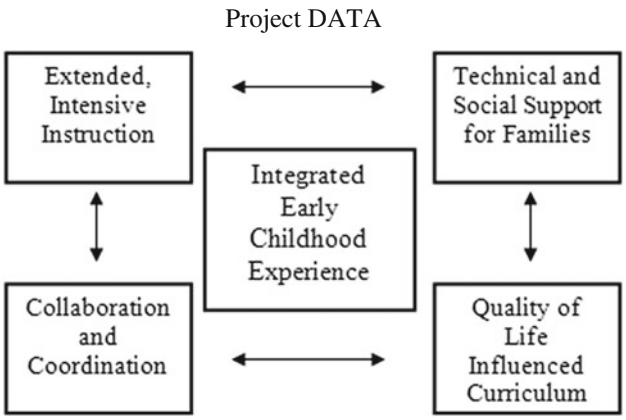

Fig. 1 Project DATA

have support from staff to interact successfully with typically developing peers. During this time, evidence-based behavioral strategies are used to embed specially designed instruction during ongoing classroom routines and activities. The second component of Project DATA is the extended, explicit instruction component, often referred to as an extended day program for young students with ASD. During these instructional hours, students received services and supports specific to their individualized needs and the core deficit areas of ASD with low student-to-teacher ratios. Focused individual and small-group interventions and activities are implemented during this time. Evidence-based instructional practices used in the extended day program include discrete trial training, incidental teaching, and social skills groups. During the extended day session, daily data were collected on skill acquisition. These data were graphed and used to make instructional decisions.

Technical support for families is the third component. This includes monthly home visits and a parent education program. The purpose of the monthly home visits was to assist families with issues that they were facing in home or in the community and to facilitate consistency in interventions in the home and educational environment. Addressing parent concerns was the first priority of the home visits. Services were provided by a trans-disciplinary educational team including masters-level special education teachers, speech language pathologists, occupational therapists, physical therapists, and behavior analysts. Parents also participated in a school-based parent education program where they were introduced to the basics of applied behavior analysis and explicit instruction.

Collaboration and coordination across providers is the fourth component. This involves frequent interaction between the student's classroom teacher and Project DATA staff. Every child participating in Project DATA had an IEP. Individualized intervention goals were set for every child based on commercially available curriculum-based assessments (Assessment, Evaluation, and Programming System—second edition [AEPS] Bricker et al., 2002); a skills checklist developed by the project to augment the AEPS (this skills checklist is available from the first author); and parent priorities. Staff members from the classroom and the extended day program meet weekly to ensure that interventions were being implemented consistently across settings.

The fifth component of the model is a quality-of-life-influenced curriculum. Project DATA is based on the philosophy that children with ASD are children first and that the goal of this intervention is to increase the quality of life of the children and their families. Therefore, when writing IEPs and developing instructional plans, we choose to teach skills that help children be more independent, be more successful in school and community settings, and interact more successfully with their families and friends.

Data Collection

In addition to the daily data that were used to assess the progress of individual children, data from four assessments were used to evaluate program effectiveness. The Childhood Autism Rating Scale (CARS) (Schopler, Reichler, & Renner, 1998) was used at program entry and program exit to evaluate the severity of autism symptoms of participants. The Peabody Picture Vocabulary Test—third edition (PPVT-III) (Dunn & Dunn, 1997), Assessment, Evaluation, and Programming System—second edition (AEPS) (Bricker et al., 2002), and Preschool and Kindergarten Behavior Scales—second edition (PKBS-2) (Merrell, 2003) were administered at the beginning and end of every school year to all the children in Project DATA. All data were collected by trained graduate students using direct administration, classroom observation, or teacher interview to complete the various measures. All assessors were trained according to the criteria on the administration, observation, and interview procedures, by senior grant staff.

CARS

The Childhood Autism Rating Scale (CARS; Schopler et al., 1998) is a measure of the degree of autistic characteristics that has been used diagnostically, which rates children's behavior on a scale from one to four, ranging from normal to severe, and yields a composite score ranging from non-autistic to mildly autistic, moderately autistic, or severely autistic. The scale is used to observe and subjectively rate 15 items: relationship to people; imitation; emotional response; body use; object use; adaptation to change; visual response; listening response; taste–smell-touch response and use; fear and nervousness; verbal communication; non-verbal communication; activity level; level and consistency of intellectual response; and general impressions. This scale can be completed by a clinician or teacher or parent, based on subjective observations of the child's behavior. Each of these 15 criteria is rated with a score of 1 = normal for child's age; 2 = mildly abnormal; 3 = moderately abnormal; and 4 = severely abnormal. Midpoint scores of 1.5, 2.5, and 3.5 are also used to create an 8-point evaluation scale. Total CARS scores range from 15 to 60; a minimum score of 30 is required to be considered an individual with ASD on this measure.

The CARS was used twice on all participants at program entry and program exit. The chosen CARS metric for analyses will be total raw scores. The CARS was completed by project staff based on classroom observation and teacher interview.

AEPS-2

The Assessment, Evaluation, and Programming System (AEPS)—second edition (Bricker et al., 2002) is a criterion-referenced, curriculum-based measure embedded within an overarching early childhood curriculum created for the use of children from birth through 6 years of age. The AEPS was created as an alternative to the use of standardized and normed measures that tend to yield scores or outcomes that provide little information for the development of educational goals or intervention strategies. In contrast to these standardized or normed measures, the AEPS was created as a comprehensive system that embeds a developmentally appropriate assessment measure that yields functional data within a larger developmentally based early childhood curriculum. Information derived from the embedded assessment is stated in functional behavioral terms that indicate the presence or non-presence of specified developmental criteria or targets that are described in functional behavioral descriptions. These data lend themselves to direct development of educationally relevant goals and interventions based on the skills that the given individual possesses and needs to acquire. Educators or interventionists can then access the overarching AEPS curriculum to create instructional programs and opportunities to teach and practice developing skills with either individual students or groups of students.

The AEPS was collected twice during each respective academic year in October and May. The chosen AEPS metric for analyses will be raw scores. The domain subscales of the AEPS have the following raw score maximums: Fine Motor = 30, Gross Motor = 34, Adaptive = 70, Cognitive = 108, Social Communication = 98, and Social = 94. On the AEPS, higher scores in a particular domain indicate greater levels of skill mastery within that domain. The AEPS was completed by the classroom teacher.

PPVT-III

The Peabody Picture Vocabulary Test—third edition (PPVT-III) (Dunn & Dunn, 1997) is an individually administered, standardized, norm-referenced assessment of receptive vocabulary for Standard English and a screening test of verbal ability for persons from 2 years and 6 months of age to 99 years of age. The PPVT-III consists of two parallel forms which each consist of 204 different stimulus words. The PPVT-III can be completed in less than 25 min and yielded scores ranging from raw score to nationally normed scale and standard scores.

The PPVT-III was conducted twice annually every year for the students who were in Project DATA. PPVT-III form A was utilized in November and form B was administered in May of respective academic years. Raw and standard scores will be utilized for analyses. The PPVT-III was administered by project staff.

PKBS-2

The Preschool and Kindergarten Behavior Scale—second edition (PKBS-2) (Merrell, 2003) is an individually administered, norm-referenced behavior rating system for children aged 3–6 years. The PKBS-2 consists of 76 questions divided among two scales indicative of social skills and problem behaviors. The social skills scale consists of 34 items and is broken down into three subscales: social cooperation, social interaction, and social independence. The problem behaviors scale includes 42 questions divided into two subscales: externalizing problems and internalizing problems. Separate scoring options are included for home or school raters. The PKBS-2 can be completed in less than 15 min as a questionnaire or in approximately 20 min of a structured interview. Scoring options range from raw score to normed scaled and standard scores.

The PKBS-2 was evaluated twice annually in November and May of every school year. The information was collected in structured interviews led by a trained graduate student with the individual student's lead classroom teacher. Results were scored using the school-based rating norms. The chosen PKBS-2 metric for analyses will be standard scores. The PKBS-2 was completed by the classroom teacher.

Results

To conduct a preliminary evaluation of the Project DATA model, assessments conducted at program entry and program exit were compared. Although every attempt was made to collect complete pre-/post-evaluation data on every participant, we do have some incomplete data sets. Pre-/post-scores for all outcome measures, the number of participants for whom we have data at each measurement point, and the results of paired-samples *t* tests on the average pre-/post-score for each measure are presented in Table 1.

CARS

On the CARS, a measure of ASD characteristics, students' ratings at program entry ranged from 17.5 (not considered on the autism spectrum) to 53.5 raw score points (in the severely autistic range), with a mean raw score of 36.76 (SD = 9.55) that would fall toward the high end of the mildly/moderately autistic. The mean at program exit decreased to 30.19 (SD = 8.11; range 16–51), which falls on the cut-score dividing ratings considered non-autistic from mildly/moderately autistic.

Table 1 Sample size (*N*), range, mean, and standard deviation of pre- and post-program outcome measures

Measure	Pre-program				Post-program			
	N	*M*	Range	*SD*	*N*	*M*	Range	*SD*
CARS*	65	36.76	14.5–53.5	9.55	64	30.19	16–51	8.11
PPVT-III raw score	69	33.19	0–108	30.30	63	46.41	0–107	31.03
PPVT-III standard score*	69	70.86	40–126	27.29	63	77.54	40–122	26.29
AEPS								
Fine Motor*	59	13.32	2–30	8.80	61	19.30	2–30	9.18
Gross Motor*	57	20.65	10–34	5.69	62	22.97	6–34	6.65
Adaptive*	62	34.63	8–66	13.60	65	40.46	4–65	14.35
Cognitive*	63	39.27	0–97	29.44	66	54.36	0–101	29.94
Social Communication*	60	39.22	0–98	29.54	59	52.08	0–98	30.11
Social*	62	41.82	12–80	18.76	66	50.36	7–89	20.60
PKBS—Social Skills								
Cooperation*	66	75.98	36–112	20.12	64	85.47	31–120	20.83
Interaction*	66	61.29	41–98	15.78	64	69.80	43–115	18.82
Independence*	66	61.91	24–103	21.01	64	74.81	36–115	21.34
Composite*	66	62.21	42–104	18.79	64	73.59	42–116	20.95
PKBS—Problem Behaviors								
Externalizing	66	98.77	81–122	10.36	64	97.81	79–126	11.45
Internalizing	66	110.29	82–135	11.05	64	107.33	82–140	10.60
Composite	66	104.36	60–125	10.92	64	102.83	82–127	9.37

CARS Childhood Autism Rating Scale; *PPVT-III* Peabody Picture Vocabulary Test—third edition; *AEPS* Assessment, Evaluation, and Programming System; *PKBS* Preschool and Kindergarten Behavior Scale

* Significant at <.01 level on paired-samples *t* tests based on pre-/post-testing

The results of the CARS pre-/post-test analysis indicated that there was a significant reduction in ASD characteristics over the course of the intervention. Paired-samples *t* tests indicated the difference from pre-/post-test rating on the CARS was statistically significant, $t(49) = 6.88$, $p < .01$. Not only did the students display a significant reduction in ASD characteristics, the exit score average was within one raw score point of the minimum required to be considered autistic by the CARS (a raw score on the CARS of 30 or more is required to be considered "mildly moderately autistic"; students scoring 29.5 or below are not considered autistic).

PPVT-III

On the PPVT-III, students' performance ranged from a standard score of 40 to 108. Twenty-two students at program entry could not perform on the PPVT-III and were assigned a raw score of 0 and the minimum possible standard score of 40. With these 22 students included in the sample, the mean standard score was 70.86 (SD = 27.29), which would indicate receptive language skills which are almost 2 standard deviations below the mean. At program exit, mean standard score was 77.54 (SD = 26.29; range 40–122), which would indicate receptive language skills which are almost 1.5 standard deviations below the mean. At program exit, there were only 6 students who could not perform on the measure, down from the 22 students who could not complete the task at the beginning of intervention.

While analyzing the pre-/post-program performance on the PPVT-III, it should be noted that there were 69 students who had a pre-program score on this measure, yet only 63 students had post-program results for this measure. Of the 6 students who did not have post-program PPVT-III scores, all exited the program without a final PPVT-III being collected. A comparison of the average pre-program rating on the PPVT-III raw scores for these 6 students without a post-program score to the 63 students with post-program results indicated that there was a slight difference in means on the pre-program PPVT-III raw score performance between the two groups [without PPVT-III post-program raw score = 34.50 (17.15) and with PPVT-III post-program raw score = 29.81(30.81)] but that the difference was not statistically significant [$t(67) = .38$, $p > .53$]. A comparison of the average pre-program rating on the PPVT-III standard scores for these 6 students without a post-program score to the 63 students with post-program results indicated that there was a slight difference in means on the pre-program PPVT-III standard score performance between the two groups [without PPVT-III post-program standard

score = 77.83 (SD = 18.10) and with PPVT-III post-program standard score = 67.52 (SD = 27.76)] but that the difference was not statistically significant [$t(67) = .89$, $p > .38$].

Finally, 22 of the study participants were assigned the lowest possible standard score of 40 at program entry as a result of their inability to perform the requisite practice items on the PPVT-III due to cognitive limitations or interfering behaviors. By the end of the program, only 6 students were unable to perform on the measure, though 14 of the students obtained the lowest possible standard score (40) due to limited receptive language capabilities when compared to same-aged peers. These data were included in the final analyses as they accurately reflect the true range of abilities of students served by Project DATA. When these 22 students' data were removed to more accurately reflect the true performance of students who were able to complete the administration of the PPVT-III, the mean pre-program PPVT-III raw score rose to 43.44 (SD = 26.45) and the mean pre-program PPVT-III standard score rose to 80.85 (SD = 23.35).

With these caveats acknowledged, the results of the PPVT-III pre-/post-test analysis of raw scores indicated that there was a significant gain in receptive language skills over the course of the intervention. Students averaged a raw score of 31.37 at program and exited the program with a mean raw score of 46.41. Paired-samples t tests indicated the difference from pre-/post-test performance on the PPVT-III raw score was statistically significant, $t(62) = 4.64$, $p < .01$.

Further analysis of the PPVT-III indicated that there was also significant growth in receptive language skills as measured by the standard score obtained. Standard scores are not typically utilized as measures of growth. Part of the calculation of standard scores depends on the chronological age of the student, and the scores are intended to remain consistent as a student ages and displays greater levels of task competence. Stated more simply, a student must display greater levels of task competence (or increased raw scores) over time to maintain a similar standard score. The DATA students had a mean standard score of 68.13 at program entry and averaged a standard score of 77.54 by the end of the intervention. Paired-samples t tests indicated the difference from pre-/post-test performance on the PPVT-III standard score was statistically significant, $t(62) = 2.94$, $p < .01$.

AEPS

Fine Motor

On the 3- to 6-year-old version of the AEPS, the DATA students' scores at program entry ranged from a raw score of 2 to 30 (which is the maximum point value for the domain) on the Fine Motor domain with an average raw score of 13.32 (SD = 8.80), 44 % domain mastery. At program exit, it ranged from a raw score of 2 to 30 on the Fine Motor domain with an average raw score of 19.30 (SD = 8.18), 64 % domain mastery. Analysis of the AEPS Fine Motor domain indicated that there was a significant gain in fine motor skills over the course of the intervention. Paired-samples t tests indicated the difference from pre-/post-test rating on the AEPS Fine Motor domain raw score was statistically significant, $t(52) = 5.51$, $p < .01$.

Gross Motor

On the Gross Motor domain, the students' entry scores ranged from a raw score of 10 to 34 and displayed a mean score of 20.65 (SD = 5.69), out of 34 potential points, 61 % domain mastery. At exit, mean score was 22.97 (SD = 6.65), 68 % domain mastery. Results of the Gross Motor domain analysis indicated that there was a significant gain in gross motor skills over the course of the intervention. Paired-samples t tests indicated the difference from pre-/post-test rating on the AEPS Gross Motor domain raw score was statistically significant, $t(51) = 3.72$, $p < .01$.

Adaptive

For the Adaptive domain, the program entry raw scores ranged from 8 to 66 raw score points (out of a possible 70) with an average of 34.63 (SD = 13.60), 49 % domain mastery. At program exit, raw scores ranged from 4 to 65 raw score points with an average of 40.46 (SD = 14.35), 58 % domain mastery. Analysis of the AEPS Adaptive domain indicated that there was a significant gain in adaptive skills over the course of the intervention. Paired-samples t tests indicated the difference from pre-/post-test rating on the AEPS Adaptive domain raw score was statistically significant, $t(58) = 3.95$, $p < .01$.

Cognitive

At program entry on the Cognitive domain, students' performance ranged from 0 to 97 raw score points (out of 108), with an average of 39.27 (SD = 29.44), 36 % domain mastery. At program exit on the Cognitive domain, students' performance ranged from 0 to 101 raw score points, with an average of 54.36 (SD = 29.94), 50 % domain mastery. Results of the Cognitive domain analysis indicated that there was a significant gain in cognitive skills over the course of the intervention. Paired-samples t tests indicated the difference from pre-/post-test rating on the AEPS Cognitive domain raw score was statistically significant, $t(60) = 4.65$, $p < .01$.

Social Communication

The raw scores ranged from 0 to 98 (maximum for the domain) on the Social Communication domain at program entry, with an average score of 39.22 (SD = 29.54), 44 % domain mastery. The raw scores ranged from 0 to 98 (maximum for the domain) on the Social Communication domain at program exit, with an average score of 52.08 (SD = 30.11), 53 % domain mastery. Analysis of the AEPS Social Communication domain indicated that there was a significant gain in social communication skills over the course of the intervention. Paired-samples *t* tests indicated that the difference from pre-/post-test rating on the AEPS Adaptive domain raw score was statistically significant, $t(51) = 4.77$, $p < .01$.

Social

At program entry on the AEPS Social domain, the DATA students' raw scores ranged from 12 to 80 (out of a possible 94), with a mean raw score of 41.82 (SD = 18.76), 48 % domain mastery. At program exit on the AEPS Social domain, the DATA students' raw scores ranged from 7 to 89, with a mean raw score of 50.36 (SD = 20.60), 54 % domain mastery. Results of the Social domain analysis indicated that there was a significant gain in social skills over the course of the intervention. Paired-samples *t* tests indicated that the difference from pre-/post-test rating on the AEPS Cognitive domain raw score was statistically significant, $t(59) = 4.18$, $p < .01$.

PKBS—Social Skills

The results of the PKBS pre-/post-test analyses relied on standard score interpretation. Standard scores have a mean of 100 and standard deviation of 15 points. On the PKBS Social Skills subscales, higher scores indicate greater levels of social adjustment, while higher scores on the Problem Behavior subscales indicate greater levels of problem behavior.

Cooperation

At program entry, the Social Skills subscale of Social Cooperation rated the students from Project DATA from a standard score of 36 to 112, with an average standard score of 75.98 (SD = 20.12). At program end, the Social Skills subscale of Social Cooperation rated the students from Project DATA from a standard score of 31 to 120, with an average standard score of 85.47 (SD = 20.83). Analysis of the Social Cooperation subscale indicated that there was a significant gain in social cooperation skills over the course of the intervention. Paired-samples *t* tests indicated that the difference from pre-/post-test rating on the PKBS Social Cooperation subscale standard score was statistically significant, $t(60) = 4.30$, $p < .01$.

Social Interaction

On the Social Skills subscale of Social Interaction at program entry, the students' ratings ranged from standard scores of 41 to 98, with a mean standard score of 61.29 (SD = 15.78). On the Social Skills subscale of Social Interaction at program exit, the students' ratings ranged from standard scores of 43 to 115, with a mean standard score of 69.80 (SD = 18.82). Results of the Social Interaction subscale analysis indicated that there was a significant gain in social interaction skills over the course of the intervention. Paired-samples *t* tests indicated that the difference from pre-/post-test rating on the PKBS Social Interaction subscale standard score was statistically significant, $t(60) = 5.307$, $p < .01$.

Independence

At program entry, the ratings on the Social Independence subscale ranged from 24 to 103, with an average standard score of 61.91 (SD = 21.01). The ratings on the Social Skills subscale of Social Independence ranged at program end from standard scores of 36 to 115, with an average standard score of 74.81 (SD = 21.34). Analysis of the Social Independence subscale indicated that there was a significant gain in social independence skills over the course of the intervention. Paired-samples *t* tests indicated that the difference from pre-/post-test rating on the PKBS Social Cooperation subscale standard score was statistically significant, $t(60) = 5.66$, $p < .01$.

Social Skills Composite

The Social Skills composite of the PKBS at program entry indicated an average standard score rating of 62.91 (SD = 18.79), with ratings ranging from a standard score of 42 to 104. This composite score would suggest that the students in Project DATA social skills at program entry were rated as averaging over 2 standard deviations below the average standard score of 100. The Social Skills composite of the PKBS at program exit indicated an average standard score rating of 73.59 (SD = 20.95), with ratings ranging from a standard score of 42 to 116. This composite score would suggest that the students in Project DATA social skills at program exit were still rated as social skills deficits. Results of the Social Skills composite subscale analysis indicated that there was a significant gain in social skills over the course of the intervention. Paired-samples *t* tests indicated that the difference from

pre-/post-test rating on the PKBS Social Interaction subscale standard score was statistically significant, $t(60) = 5.85$, $p < .01$.

PKBS–Problem Behaviors

Externalizing

On the Externalizing Problems subscale, program entry ratings ranged from a standard score of 81 to 122, with an average of 98.77 (SD = 10.36) standard score points. This finding indicated that, according to this subscale of the PKBS, students from Project DATA displayed externalizing problem behaviors at rates similar to typically developing students. On the Problem Behaviors subscales of the PKBS, the program exit ratings for the Externalizing Problems subscale ranged from a standard score of 79 to 126, with an average of 97.81 (SD = 11.45) standard score points. Analysis of the Externalizing Problems subscale of the PKBS indicated that there was no statistically significant reduction in externalizing problems over the course of the intervention. Paired-samples *t* tests indicated that the difference from pre-/post-test rating on the PKBS Externalizing Problems subscale standard score was not statistically significant, $t(60) = -1.28$, $p > 21$.

Internalizing

At program entry on the Internalizing Behaviors subscale, the students' ratings ranged from a standard score of 82 to 135, with a mean standard score of 110.29 (SD = 11.05). At program exit on the Internalizing Behaviors subscale, the students' ratings ranged from a standard score of 82 to 140, with a mean standard score of 107.33 (SD = 10.60). Results of the Internalizing Problems subscale analysis indicated that, according to the PKBS, there was no significant reduction in internalizing problems over the course of the intervention. Paired-samples *t* tests indicated that the difference from pre-/post-test rating on the PKBS Internalizing Problems subscale standard score was not statistically significant, $t(60) = -1.86$, $p > .07$.

Problem Behavior Composite

The Problem Behaviors composite on the PKBS indicated that at program entry, the students from Project DATA were rated with standard scores ranging from 60 to 125, with an average standard score of 104.36 (SD = 10.92). Taken as a whole, the results from the program entry PKBS suggest that the students from Project DATA display problem behaviors at rates similar to their typically developing peers. The Problem Behaviors composite on the PKBS indicated that at program exit, the students from Project DATA were rated with standard scores ranging from 82 to 127, with an average standard score of 102.83 (SD = 9.37). Analysis of the Problem Behaviors composite subscale of the PKBS indicated that there was no significant reduction in ratings of problem behavior over the course of the intervention. Paired-samples *t* tests indicated that the difference from pre-/post-test rating on the PKBS Externalizing Problems subscale standard score was not statistically significant, $t(60) = -1.40$, $p > .17$. Although the students' ratings revealed a small, but non-significant reduction in problem behaviors as measured by standard score growth on the PKBS, it is encouraging to note that the average post-test rating of problem behaviors in the DATA students was rated just slightly above the mean of 100, indicating that these students were rated as having problem behaviors at near the rates as the average student.

Discussion

These data indicate that Project DATA is a promising practice for providing an early intensive behavioral intervention to young students with ASD in a public school setting. Based on the descriptive statistics of the participants in Project DATA, it was evident that the population was similar to previous epidemiological studies that indicated a 4:1 male-to-female ratio in students with ASD (e.g., Ehlers & Gillberg, 1993; Kogan et al., 2009). The mean age at entry was over 50 months of age; however, many students did enter the program closer to their third birthday. The performance and/or ratings on the outcome measures varied dramatically across participants, suggesting that the students in Project DATA represented a broad range of skills and abilities, especially in the core deficit areas of ASD (social communication, social skills, and degree of ASD characteristics). Importantly, like school districts providing services to children with ASD, the only requirements to participate in this trial were a diagnosis of ASD, being of preschool age, referral by the school district, and parent consent.

Overall, the results of the univariate pre-/post-program analyses suggest support for positive change in child behavior during their participation in the Project DATA intervention. Significant differences were observed between pre- and post-program measurement on 13 of the 16 outcome measures. More importantly, the student change scores in social communication, social skills, and ratings of degree of ASD characteristics, the core deficit areas associated with ASD, improved dramatically, all achieving a medium effect strength (calculated using *Cohen's d*; range = .42 to .61). In particular, student ratings on the CARS, a measure of degree

of ASD characteristics, decreased over 5 points to a post-program mean rating of 30.39. This is a significant decrease not only described by significant raw score change, but also in degree of ASD characteristics displayed by students; a raw score of 30 is the minimum score required to be considered on the autism spectrum on the CARS. Stated in a different way, the post-program mean rating on the CARS suggests that students exited the Project DATA program with a marked reduction in degree of ASD characteristics and a student represented by the post-program mean rating would be considered in the "mildly autistic" range on the measure. Although we do not have a control group to compare these changes, it has been well documented that children with ASD do not make these types of improvements without systematic instruction.

Communication skills as measured by the PPVT-III and the AEPS Social Communication domain scores displayed significant changes as well. The PPVT-III standard score improved by almost half of a standard deviation, though the effect size was small. These results should be interpreted with the knowledge that at program entry, over 33 % of the population received the lowest possible standard score of 40 due to student inability to perform the tasks requisite to standardized administration of the PPVT-III. The fact that a large percentage of the participants could not perform on the PPVT-III implies a few things. First, the PPVT-III may not be an appropriate measure of skills in some young students with ASD. Second, this subpopulation of the students could be representative of the 40–55 % prevalence rates of intellectual disability in persons with ASD found in recent, empirically based research (Edelson, 2006). Finally, the PPVT-III raw score appears to be a more appropriate measure of change in young students with ASD due to a limitation in standard score characteristics, which automatically accounts for raw score growth over time. More clearly stated, standard scores are interpolated by an algorithm that accounts for chronological maturation. In practical application, this means that a given student must achieve a higher raw score at each future administration of an assessment to maintain the same standard score over time. As a result of this metric issue, the PPVT-III raw score displayed larger post-program gains than standard scores and the effect strength was within the medium range.

Social skills as measured by the AEPS Social Skills domain and the PKBS Social Skills subscales and composite indicated significant changes from pre- to post-program, suggesting that Project DATA intervention is effective at improving the social skills development in young students with ASD. Though the social functioning of the students who participated in Project DATA still fell within the impaired range when compared to same-aged peers, there was considerable positive change in both the AEPS Social Skills domain raw score and all of the PKBS Social Skills subscales standard scores at post-program. In particular, the PKBS Social Skills subscale of Social Cooperation fell within the very low end of the average range, with a mean standard score of 85.47. The effect sizes for all of the social skills measures fell within the medium effect range (*Cohen's d*; range = .44 to .61).

Relative to other empirical studies of the efficacy of CTMs in service delivery to young students with ASD, Project DATA chose not to use a true standardized assessment of intelligence on the participating students. A few of the reasons to not use standard intelligence batteries on the Project DATA participants included the following: time required to complete a full intelligence battery; the fact that many of our students with ASD had limited expressive repertoires rendering most standardized batteries as inappropriate measures for the population; the notorious variance in test–retest reliability with early childhood populations, especially those students with ASD; and lack of congruence with intelligence scores as predictors of success in schools for young students with ASD. Instead, cognitive skills were assessed using the AEPS Cognitive domain and the PPVT-III as a proxy measure of intelligence due to the strong positive correlations with this measure and other recognized measures of intelligence (Dunn & Dunn, 1997). Similar to other empirical studies of CTMs which utilized cognitive outcome measures as reviewed in Rogers and Vismara (2008), participants in the Project DATA intervention displayed significant post-program gains in cognitive skills as measured by the AEPS Cognitive domain and the PPVT-III standard score; students gained almost 15 raw score points on the AEPS Cognitive domain and, as mentioned previously, gained over 7 points on the PPVT-III standard score. The gain on the PPVT-III standard score created an effect size in the small effect range, though the change in raw score on the AEPS Cognitive domain resulted in an effect size within the medium range.

The only outcome measures that did not differ significantly from pre- to post-program assessment were those pertaining to problem behaviors as measured by the PKBS Problem Behavior subscales. The students in Project DATA were rated as having slightly more Internalizing problem behaviors than their same-aged peers and only modest decreases in ratings of Internalizing problem behaviors were noted post-program. More positively, the students in Project DATA were rated as exhibiting slightly fewer Externalizing behaviors than their typically developing age-mates at both pre- and post-program ratings. The Problem Behaviors composite on the PKBS, which averages the Internalizing Problems and Externalizing Problem subscales, indicated overall problem behavior ratings in the average range and suggests that the students in Project DATA displayed

problem behaviors at rates similar to their same-aged peers. This may indicate that this sample of young children with ASD did not demonstrate many challenging behaviors or that the PKBS may be inadequate to identify the presence of both externalizing and internalizing problem behaviors in young children with ASD. More research on effective and efficient strategies to assess challenging behaviors in groups of children with ASD is needed.

Limitations

There are a number of serious limitations to the data presented here and the conclusions that can be drawn from this data set. First, this quasi-experimental evaluation of child behavior during participation in the Project DATA CTM did not incorporate a control group. Though an RCT design would enact the gold standard within the research community, there are ethical and potential legal ramifications for employing this methodology and randomly assigning students to groups within the context of programs serving students in educational settings rather than clinical or home settings. Future research should focus on establishing effective, non-RCT designs to determine the efficacy of CTMs operating within educational systems.

Second, rival hypotheses could explain the changes on outcome measures. It is possible that the gains in outcome measures of the participants in the Project DATA CTM were merely the result of typical developmental maturation. On the AEPS in particular, dramatic differences in performance could be attributed to chronological maturation and associated developmental sophistication. It stands to reason that a 5-year-old student would outperform an analogous 3-year-old student on a developmental measure that spans the period from 36 months of age to 72 months of age. The significant improvements over time on the measures utilizing standard scores, which should account for developmental/chronological maturation, as well as the significant changes in the performance on measures of the core deficit areas of ASD would seemingly invalidate the rival hypothesis of growth due to typical developmental maturation, but further replication studies of the Project DATA CTM will be required to answer that question.

Although further replications of Project DATA will help to address this issue, we have burgeoning evidence from the growing number of RCTs on treatment for ASD that when young children with ASD do not receive high-quality early intensive behavioral intervention, they do not make significant improvements in development. Comparing our model to the LEAP RCT (Strain and Bovey 2011) provides more evidence to suggest that the children who participated in Project DATA would not have made significant improvements in development across domains without this intensive and specialized intervention.

Finally, no measure of procedural fidelity was utilized in these analyses. As the studies of the efficacy of various CTMs increase, a number of researchers have noted the importance of utilizing a procedural fidelity measure to ensure adherence to program methodologies and practices (Reichow et al., 2008; Reichow & Wolery, 2009). An ongoing measure of the fidelity of implementation across participants, settings, and service providers would enhance the findings of further analyses of the Project DATA CTM.

Implications for Practice

Project DATA shares some key characteristics with LEAP which is the most studied school-based CTM for young children with ASD. First, like LEAP, Project DATA is based on applied behavior analysis (Strain & Schwartz, 2009), but embed instruction into ongoing activities rather than relying on de-contextualized discrete trial training. Second, both LEAP and Project DATA begin inclusion of children with ASD into high-quality preschool classrooms on day one of intervention and view the opportunity to interact successfully with typically developing children as an important part of the intervention. Both of these programs operate on the idea that children do not need to earn their way into inclusive settings, but rather need to be supported in those settings to flourish and learn. Finally, both models view intensity not as a simple metric of the number of hours of services per week, but rather as a complex notion of how children are engaged in learning across settings, materials, peers, and adults. As the number of preschoolers with ASD continues to increase, school programs must develop effective and sustainable school-based treatment options for young children with ASD. These two projects can serve as models for school district as they work to develop local programs for young children with ASD.

The current results provide preliminary support for the Project DATA CTM tenets and practices as an effective early intensive childhood intervention program for young students with ASD. These findings further the research base on successful early intensive intervention practices for young students with ASD, particularly supporting service delivery models within public education settings that combine best practices in early childhood special education with best practices in service provision to persons with ASD. The study identifies the Project DATA CTM as a viable early intervention program appropriate for adoption in other public education early intervention settings and that the next step for the program would be conducting RCT design replication studies at other sites.

References

Autism Developmental Disabilities Monitoring Network. (2012). Prevalence of autism spectrum disorders—autism and developmental disabilities monitoring network, 14 Sites, United States 2008. *Morbidity and Mortality Weekly Report, 61*(SS03), 1–19.

Bricker, D., Pretti-Frontczak, K., Johnson, J. J., Straka, E., Capt, B., Slentz, K., et al. (2002). *Assessment, evaluation, and programming system for infants and children (second edition)* (2nd ed.). Baltimore, MD: Paul H. Brooks Publishing Co.

Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., et al. (2010). Randomized, controlled trial of an intervention for toddlers with autism: The early start Denver model. *Pediatrics, 125*, 17–23.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test* (3rd ed.). Circle Pines, MN: American Guidance Services.

Edelson, M. G. (2006). Are the majority of children with autism mentally retarded? A systematic evaluation of the data. *Focus on Autism and Other Developmental Disabilities, 21*(2), 66–83.

Ehlers, S., & Gillberg, C. (1993). The epidemiology of Asperger syndrome. A total population study. *Journal of Child Psychology and Psychiatry, 34*(8), 1327–1350.

Kogan, M. D., Blumberg, S. J., Schieve, L. A., Boyle, C. A., Perrin, J. M., Ghandour, R. M., et al. (2009). Prevalence of parent-reported diagnosis of autism spectrum disorder among children in the US, 2007. *Pediatrics, 124*(5), 1395–1403.

Levy, S., Kim, A.-H., & Olive, M. L. (2006). Interventions for young children with autism: A synthesis of the literature. *Focus on Autism and Other Developmental Disabilities, 21*(1), 55–62.

Lord, C., Bristol-Power, M., Cafiero, J. M., Filipek, P. A., Gallagher, J. J., Harris, S. L., et al. (2001). *Educating children with autism*. Washington, DC: National Academy Press.

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.

Merrell, K. W. (2003). *Preschool and Kindergarten Behavior Scales* (2nd ed.). Austin, TX: PRO-ED.

Odom, S. L., Boyd, B. A., Hall, L. J., & Hume, K. (2010). Evaluation of comprehensive treatment models for individuals with Autism spectrum disorders. *Journal of Autism and Developmental Disorders, 40*, 425–436.

OSEP. (1995–2007). Part B Child Count Data by Disability 1995–2007: OSEP Data Accountability Center.

Reichow, B., Volkmar, F. R., & Cicchetti, D. V. (2008). Development of the evaluative method for evaluating and determining evidence-based practices in autism. *Journal of Autism and Developmental Disorders, 38*, 1311–1319.

Reichow, B., & Wolery, M. (2009). Comprehensive synthesis of early intensive behavioral interventions for young children with Autism based on the UCLA Young Autism Project Model. *Journal of Autism and Developmental Disorders, 38*(1), 23–41.

Rickards, A. L., Walstab, J. E., Wright-Rossi, R. A., Simpson, J., & Reddihough, D. S. (2007). A randomized, controlled trial of a home-based Intervention program for children with autism and developmental delay. *Journal of Developmental and Behavioral Pediatrics, 28*(4), 308–316.

Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child & Adolescent Psychology, 37*(1), 8–38.

Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal on Mental Retardation, 110*, 417–438.

Schopler, E., Reichler, R. J., & Renner, B. R. (1998). *The childhood autism rating scale (CARS)*. Los Angeles: Western Psychological Services.

Schwartz, I. S., Sandall, S. R., McBride, B. J., & Boulware, G.-L. (2004). Project DATA (Developmentally Appropriate Treatment for Autism): An inclusive school-based approach to educating young children with Autism. *Topics in Early Childhood Special Education, 24*(3), 156–168.

Strain, P. S., & Bovey, E. H. (2011). Randomized, controlled trial of the LEAP model of early intervention for young children with autism spectrum disorders. *Topics in Early Childhood Special Education, 31*(3), 133–154.

Strain, P., & Schwartz, I. (2009). Positive behavior support and early intervention for children with autism: Case studies on the efficacy of proactive treatment of problem behavior. In W. Sailor, G. Dunlap, G. Sugai, & R. Hunter (Eds.), *Handbook of positive behavior support* (pp. 107–124). Baltimore, MD: Paul Brookes.