J Autism Dev Disord (2016) 46:544–560
DOI 10.1007/s10803-015-2597-0

ORIGINAL PAPER

Randomized, Controlled Trial of a Comprehensive Program for Young Students with Autism Spectrum Disorder

Helen E. Young[1] · Ruth A. Falco[1] · Makoto Hanita[2]

Published online: 5 October 2015
© Springer Science+Business Media New York 2015

Abstract This randomized, controlled trial, comparing the Comprehensive Autism Program (CAP) and business as usual programs, studied outcomes for 3–5 year old students with autism spectrum disorder (ASD). Participants included 84 teachers and 302 students with ASD and their parents. CAP utilized specialized curricula and training components to implement specific evidence-based practices both at school and home. A comprehensive set of outcome areas was studied. Hierarchical linear modeling was used to estimate the treatment impact. CAP had small positive impacts on the students' receptive language (effect size of .13) and on their social skills as rated by teachers (effect size of .19). Treatment effects were moderated by severity of ASD.

Keywords Autism spectrum disorder · Comprehensive treatment model · Evidence-based practices · Hierarchical linear modeling · Randomized controlled trial · School-based intervention

Introduction

Development and study of comprehensive treatment models (CTMs) for use by educators in schools serving the growing numbers of young children with autism spectrum disorder (ASD) forms an essential area of current educational research (Odom et al. 2010). Odom et al. defined CTMs as those addressing the core features of ASD and other domains that are fundamental to success in school, such as cognitive skills and adaptive behavior. This comprehensive approach to educational intervention was recommended by the National Research Council on Autism Spectrum Disorders (NRC 2001). Odom et al. identified the need to clearly describe components of each proposed CTM, develop measures of implementation, and provide evidence of efficacy. They contrasted CTMs with evidence-based practices (EBPs), which address specific learning needs of children with ASD (Odom et al. 2010, p. 425). The National Professional Development Center on Autism Spectrum Disorders and The National Standards Project reviewed extensive research and identified 27 EBPs meeting rigorous evidence criteria (Wong et al. 2014; Wilczynski 2010).

While research has identified recommended CTM components (NRC 2001) and specific EBPs (Wong et al. 2014) for young children with ASD, there remains a gap between this evidence and the implementation of these components and practices by educators in public school classrooms (Dingfelder and Mandel 2011; Kasari and Smith 2013; Odom et al. 2013). Only a few recent studies of CTMs investigated comprehensive outcomes for young students, including the core features of ASD (e.g., social, communication) and other curricular areas essential to performance in school (e.g., cognitive, adaptive behavior) as recommended by the NRC (2001). Additionally, only a small number of studies involved educators in school classrooms, rather than research personnel, as primary providers of the intervention for students. School-based studies, addressing comprehensive outcomes, should play a critical role in closing the research-to-practice gap for young students with ASD.

✉ Ruth A. Falco
falcor@pdx.edu

[1] Department of Special Education, Graduate School of Education, Portland State University, P.O. Box 751, Portland, OR 97207-0751, USA

[2] Education Northwest, 101 SW Main St, Suite 500, Portland, OR 97204, USA

In a quasi-experimental study of CTMs with intervention by trained educators, Boyd et al. (2014) compared outcomes for preschool classrooms utilizing LEAP (Strain and Bovey 2011); classrooms implementing TEACCH (Mesibov and Shea 2010); and high quality Non-Model-Specific (NMS) special education classrooms. Criteria for participation in the study required all classrooms have at least average performance on the *PDA Program Assessment* (Professional Development in Autism Center 2008), an observational measure of preschool classroom quality. Participation also required better than average performance on a measure of fidelity for LEAP or TEACCH for classrooms implementing those models. Examination of common components of all three classroom types found considerable overlap in the specific practices used. Above average fidelity in intervention practices of LEAP, TEACCH, or NMS resulted in substantial gains in performance on outcome measures, but statistically significant differences in outcomes across the three approaches were not identified (Boyd et al. 2014).

In another comparison of CTMs, Mandell et al. (2013) conducted a randomized controlled trial (RCT) to examine outcomes for kindergarten-through-second-grade students with ASD when teachers received training to use either the *STAR Program: Strategies for Teaching based on Autism Research* (*STAR Program*) (Arick et al. 2004) or structured teaching, based upon the TEACCH program (Mesibov and Shea 2010). Educators using the *STAR Program* (Arick et al. 2004) received significantly more hours of training (M = 58 h) than those using structured teaching (M = 16.5 h). Mandell et al. reported that the criterion for fidelity of implementation on 80 % of components was not achieved for either program during the 8 months of the study. Fidelity of implementation was achieved for an average of 57 % of components for the *STAR Program* (Arick et al. 2004) and 47 % of components for structured teaching. Although students in both groups achieved substantial cognitive gains on the *Differential Abilities Scales-Second Edition* (Elliott 1990), analysis did not identify a main effect of program when comparing the *STAR Program* (Arick et al. 2004) and structured teaching groups (Mandell et al. 2013, p. 282). Additionally, Mandell et al. noticed considerable overlap between the practices of *STAR Program* (Arick et al. 2004) and structured teaching (Mesibov and Shea 2010) approaches in the areas of classroom organization, scheduling, and transition strategies.

Strain and Bovey (2011) focused on the role of training and coaching to support fidelity in a RCT which studied outcomes for the LEAP CTM as implemented by educators in preschool classrooms. The LEAP model focused on use of a high quality early childhood curriculum and peer-mediated intervention. The model included a parent education component. Strain and Bovey compared outcomes over a two-year period for students with ASD in classrooms with teachers who received on-going training and coaching in the LEAP model to outcomes for classrooms with teachers who were given the LEAP implementation manuals and brief initial training. Student outcomes were significantly better when educators were provided with 2 years of training and coaching to fidelity in implementation of the model.

Training and coaching for educators, as shown by Strain and Bovey (2011), can play a critical role in an effective educational model for young children with ASD. Ruble et al. (2014) applied the framework of implementation science (Fixen et al. 2005) in a study of relationships between fidelity of implementation of the Collaborative Model for Promoting Competence and Success (COMPASS) with children having ASD, the fidelity of intervention practices by educators as identified in intervention plans, and the outcomes for students. Ruble et al. (2014) pointed out that implementation science provided a way to differentiate the fidelity of implementation of components of the model (e.g., the methods used to consult with and coach teachers) from the fidelity in the use of intervention practices by the teachers. This provided a way to better analyze and understand the roles of these components of COMPASS in the outcomes for students. Using RCT data, Ruble et al. concluded that fidelity to the COMPASS model was related to quality of teacher intervention practices, and intervention quality was related to child outcomes.

These studies demonstrated that young students with ASD can make important gains in classrooms where CTMs and EBPs are implemented by trained educators. The primary aim of the Comprehensive Autism Program (CAP) RCT, as reported here, expands upon this body of research by investigating outcomes for 3- to 5-year-old students, across the spectrum of ASD, served by educators in public school preschool and kindergarten classrooms implementing the CAP, a comprehensive model applying components recommended for young children with ASD by the NRC (2001). Outcomes for students in CAP schools, in core areas of ASD and other areas important for performance in school, were compared to outcomes for students in business as usual (BAU) public schools. The CAP model utilized specialized curricula designed for educators and parents of young children with ASD. CAP training and coaching components focused on preparing educators and parents to use the curricula and a specific set of EBPs (Wong et al. 2014). Building on implementation science (Dunst et al. 2013; Fixen et al. 2005), we monitored fidelity of CAP implementation components by the research team and fidelity of intervention with the EBPs by educators in treatment and control classrooms. The RCT addressed the following research question: Does implementation of the CAP model have a positive impact on outcomes for

expressive language, receptive language, social skills, cognitive abilities, and adaptive behavior for young students with ASD after one school year compared to outcomes for a control group of students with ASD receiving their Business-As-Usual program?

Methods

Participants for this cluster randomized trial were recruited from public school districts in the states of Oregon and Washington. The project was conducted over three different school years and included three cohorts. In the spring and summer prior to each school year, project staff contacted special education administrators to explain the purpose of the study and enrollment criteria. Participation in the study required that (a) the school served at least three students, ages 3–5, with ASD (had a medical diagnosis of ASD or *Childhood Autism Rating Scale* [*CARS*] [Schopler et al. 1988] score within the autistic range); (b) participants were not implementing the curricula used in the study [i.e., *STAR Program* (Arick et al. 2004) or *Autism Partnership Program: Parents and Educators Partnering to Improve Outcomes for Children and Youth with Autism (APP)* (Young et al. 2009)]; (c) teachers had not participated in the study in a previous school year; and (d) teachers and parents were willing to sign a consent form agreeing to participate if randomly assigned to the CAP condition or the BAU condition. We included in the study students across the range of ASD who were being educated in school programs. Prior to recruitment, the Institutional Review Board at the researchers' university approved the study.

Sample Size and Enrollment

Based on recommendations from the Institute of Education Sciences regarding sample sizes for RCTs (U.S. Department of Education 2003), we planned to enroll a minimum of 50 schools (at least 25 schools in the intervention group and 25 schools in the control group). Over the course of the study, administrators in 38 school districts agreed to participate in the study. The districts, located in both urban and rural areas, were diverse in size, with total enrollment ranging from 1500 to 52,000 students.

Before the school year began, administrators nominated schools that met the enrollment criteria. All teachers who served students with ASD at each nominated school were invited to participate. Teachers signed consent forms agreeing they would participate as either a CAP or BAU site, and they sent consent forms home to parents of all eligible students. These students had been evaluated by their school district and were nominated to participate because they were identified as having an ASD. Children of all consenting parents were included in the study.

Randomization

Over the course of 3 years, administrators nominated a total of 78 schools for participation: 15 in year one, 47 in year two, and 16 in year three. Each year, schools were randomly assigned to conditions using a random number generator in *Stata* (StataCorp 2009). In total, 41 schools were randomly assigned to the CAP condition, and 37 schools were randomly assigned to the BAU condition.

Description of CAP and BAU Conditions

CAP

Table 1 outlines components of CAP and compares them to components in BAU schools. The CAP model targeted use of a selected set of EBPs for all young students with ASD across environments and throughout the day. Environmental arrangement and visual supports (Hume 2013a, b) provided concrete cues to guide behavior and promote the functional use of skills at school and at home. These practices supported learning of communication, social interaction, and social behavior skills, core areas of concern for children with ASD. Naturalistic and incidental teaching applied behavioral strategies (e.g., task analysis, reinforcement, prompting, and time delay) within routines (Boyd and Wong 2013; Cox 2013a, b; Fleury 2013a, b; Kucharczyk 2013; Wong 2013a) to teach communication, social interaction, social behavior, cognitive and adaptive skills within meaningful contexts at school and at home. Pivotal response training (PRT) (Wong 2013b) is also based on applied behavior analysis and was used to enhance child-initiated communication, build spontaneous interaction, and expand skills within naturalistic situations including play. In addition to application throughout the day, educators were asked to use behavioral strategies within routines and PRT in one-to-one, 15-min, daily instructional sessions for each child with ASD. CAP also used discrete trial training (DTT) (Fleury 2013c) to focus on the acquisition of new language, cognitive and pre-academic skills. Educators were asked to provide one-to-one, 15-min, daily, DTT sessions for each student. The National Professional Development Center on Autism Spectrum Disorder provides evidence of efficacy for each of these EBPs on their website (http://autismpdc.fpg.unc.edu/). CAP staff trained educators to use these EBPs, but educators were also allowed to continue other practices they were already using. For example, some educators used the Picture Exchange Communication System (Frost and Bondy 2002) or work systems (Hume et al. 2012).

To support use of the EBPs, CAP provided each classroom teacher the *STAR Program* (Arick et al. 2004) curriculum. This curriculum included a manual that described

Table 1 Intervention practices and implementation components of CAP and BAU

CAP	BAU
Intervention practices recommended for use by educators and parents who implemented CAP:	*Intervention practices* used by educators who implemented BAU:
Environmental arrangement (used throughout the day)	Intervention practices, as observed and reported by educators included: Applied Behavior Analysis, DTT, direct instruction, environmental arrangement, Floor Time, Picture Exchange Communication System, PRT, social skills training, TEACCH work systems, visual supports
Visual supports (used throughout the day)	
Behavioral strategies within routines (used throughout the day and 15 min per school day of 1:1 instruction)	
PRT (used throughout the day and 15 min per school day of 1:1 instruction)	
Additional intervention practices for use by educators who implemented CAP:	
DTT (15 min per school day of 1:1 instruction recommended)	
Other intervention practices, as reported by educators included: Applied Behavior Analysis, Floor Time, preacademic instruction, picture communication systems, social skills training, TEACCH work systems	
Implementation components provided for CAP educators:	*Implementation components* provided by CAP for BAU educators:
One 2-day training workshop and 6–8 monthly coaching visits for use of EBPs, including:	No training workshops or coaching
Environmental arrangement	No curricula
Visual supports	
Behavioral strategies within routines	
PRT	Other for BAU educators:
DTT	No use of the *STAR Program* (Arick et al. 2004) curriculum
Use of *STAR Program* (Arick et al. 2004) materials	No limitations on use of other curricula
STAR Program curriculum, including:	No restrictions on other training
Manual	No limitations on use of training or coaching by school district
Curriculum-based assessment	
Written lesson plans for Routines, PRT, and DTT	
APP curriculum (Young et al. 2009)	
Opportunities for educators to attend parent workshops and/or participate in consultation/home visits	
Implementation components provided for CAP parents:	*Implementation components* provided by CAP for BAU parents:
Two, 2-h, group workshops regarding ASD and use of EBPs, including:	No workshops
Environmental arrangement	
Visual supports	No consultation/home visits
Behavioral strategies within routines	No use of *APP* curriculum (Young et al. 2009)
PRT	No limitations on other training or curricula provided by parents or school district
	Other for BAU parents:
Consultation/home visits offered for individual families	No restrictions on other training
APP curriculum (Young et al. 2009)	

teaching practices for functional routines (FR), PRT, and DTT. The curriculum materials included lesson plans at three levels of difficulty for use in FR, PRT, and DTT instruction, data collection forms to monitor progress, and a curriculum-based assessment used to place students in appropriate lessons. The *STAR Program* (Arick et al. 2004) does not include training materials, a coaching component, or a parent education component. These components were added for the CAP model.

The CAP research team created training and coaching guidelines and materials to prepare classroom staff for implementation of the selected EBPs. Initial workshops presented by project staff provided approximately 10 h of training across 2 days. All teachers, paraeducators, and

speech-language pathologists who served students enrolled in the study, were invited to attend the workshops. Other professionals who worked with the students, such as occupational therapists, school psychologists, and administrators, were also encouraged to participate. Workshop content addressed (a) learning strengths and challenges of children with ASD, (b) environmental arrangement and visual supports, (c) behavioral teaching strategies within routines, (d) PRT, (e) DTT, (f) procedures to place students in *STAR Program* (Arick et al. 2004) lesson plans, and (g) use of the lesson plans and related data collection procedures to guide instruction. Following the initial training, approximately 10 h of in-class training across 2 days were provided at each school site. This training assisted educators to arrange the classroom environment, develop the class schedule, provide individual visual schedules for each child, design other visual supports, assess students for placement in the lesson plans, and prepare classroom staff for implementation of the EBPs. After the initial on-site training, school staff received an average of 27 h of on-site coaching.

Coaching applied a model similar to that described by Wilson et al. (2012). Coaches used a collaborative process to observe, provide performance feedback, encourage reflection, problem-solve, and provide support and encouragement. The coaching occurred 1 day per month, for at least 6 months. While most monthly visits lasted approximately 4 h, some visits were longer and extended across two consecutive days. Longer visits were provided when coaches observed fidelity to be low or the educators identified greater needs. During each visit, CAP coaches modeled the EBPs and used checklists of guidelines for each EBP developed by project staff to observe and provide performance feedback to educators. Coaching occurred during group and one-to-one instructional sessions for students with ASD. CAP provided coaching for teachers, paraeducators, speech-language pathologists and some other related services staff, such as occupational therapists. Phone and email consultation was available to educators between coaching visits.

Parent education components of CAP focused on the EBPs of environmental arrangement and visual supports; incidental teaching to apply behavioral teaching strategies within routines; and PRT within home and community routines. These EBPs were the same strategies taught to the CAP educators, with the exception of DTT. Parent education was guided by the *APP* curriculum (Young et al. 2009). Instead of focusing on just one core area, such as communication, this curriculum, designed by both parents and educators, includes numerous evidence-based strategies that parents can successfully use in their homes and community to improve outcomes for their children. It outlines information and strategies in five areas: (a) evidence-based information and interventions for ASD; (b) visual supports in home environments; (c) teaching communication skills; (e) behavior management within home routines; and (f) teaching self-care and other home routines, building independence, and teaching play skills. Included are "parent-friendly" instructions and pictures, describing and illustrating evidence-based strategies. All participating families received a copy of the *APP* (Young et al. 2009). Each teacher also received a copy of this curriculum for use in consulting with families.

CAP coaches offered two parent workshops at each school. Workshops addressed (a) learning strengths and challenges of children with ASD; (b) an overview of the CAP study; (c) components and practices used in CAP classrooms; (d) environmental arrangement and visual supports at home; (e) methods for teaching communication during home routines; (f) methods to develop independence in home routines using effective visual supports and prompting strategies; and (g) positive behavioral support. Each workshop lasted approximately 2 h.

Parents received written invitations to the workshops through the mail, and CAP staff and teachers used additional notes, emails, and phone calls to encourage parents to participate. Teachers and other classroom staff were also invited to attend. Paraeducators received gift cards to local stores for providing childcare during each workshop. After the workshops, CAP coaches used specific procedures and forms to help parents plan for implementation of EBPs and offered parents one or two follow-up consultations in their home or in a setting at school. Follow-up consultations, which lasted from 1 to 2 h, focused on priority routines for the family, such as toilet use or mealtime. The coach helped to plan ways to organize the environment and use visual supports for home routines, use behavioral teaching strategies within the routines, and embed PRT and visual supports to elicit use of language. The coach also pointed out relevant sections of the *APP* curriculum (Young et al. 2009) to provide further information and examples, including pictures. Additionally, project staff developed individualized visual supports and social stories designed to address the child's needs for use by parents during daily routines. Parents could also contact coaches by email or telephone to ask questions and seek advice.

BAU

BAU classrooms were instructed to continue use of practices they were already implementing for children with ASD. Examples of these practices are listed in Table 1. Before or during data collection, CAP staff did not provide any curricula, materials, training, or coaching to educators or parents of children attending BAU classrooms. All communications from CAP staff to administrators and

teachers of BAU classrooms asked them to continue to implement their usual components and practices. Communications to parents of children attending BAU classrooms did not include advice or training regarding intervention for their children.

As an incentive for participation, workshops on implementation of CAP were provided to BAU schools after all data was collected at the end of the school year. At that time, BAU parents received a copy of the *APP* curriculum (Young et al. 2009), and BAU educators received copies of the *STAR Program* (Arick et al. 2004) and the *APP* curriculum.

Fidelity of Implementation and Intervention

Record-keeping procedures and data forms were developed to monitor implementation of staff training and parent education for CAP schools. Measures included sign-in sheets for participants in workshops and schedules and logs for trainers and coaches. Coaches recorded the date and time that each coaching visit began and ended in their training notebook.

Measures were developed to monitor fidelity of educators' use of the specific EBPs. Some of these measures were adapted from measures developed for monitoring fidelity of intervention practices within routines, PRT, and DTT in the Mandell et al. (2013) study. Researchers involved in the Mandell et al. study provided training for CAP research staff in the use of the tools from that study. The CAP study used the following fidelity measures: (a) Classroom Observation Form to assess environmental arrangement and visual supports; (b) Routines Fidelity Form to assess use of behavioral teaching strategies within routines; (c) PRT Fidelity Form; and (d) DTT Fidelity Form. During winter and spring of the intervention year for each cohort, a trained research assistant visited each CAP and BAU classroom to collect fidelity of intervention data. During these visits, an observation was conducted using the Classroom Observation Form. The observers also recorded video clips for use in monitoring fidelity of other EBPs (i.e., routines, PRT, DTT). Teachers identified 10-min periods on the observation day for video-taping of each of these three types of practices or similar practices. Additionally, during teacher interviews at baseline and near the end of the school year, teachers were provided with clear descriptions of these practices and asked to report the minutes of implementation per day for each of the three practices for each individual child with ASD.

Analysis of Fidelity of EBP at School

Total scores on the Classroom Observation Form were calculated in winter and in spring. These two scores were averaged to provide an overall fidelity score for each classroom. Procedures for coding of the winter and spring video clips were developed for behavioral strategies within routines, PRT, and, DTT. The coding systems were adapted from the procedures developed for Mandell et al. (2013). Research assistants were trained to use the coding systems with at least 90 % reliability before they began coding tapes for the data analysis. At least 30 % of tapes coded by each research assistant were also coded by a CAP coach and at least 90 % reliability of coding was maintained. Fidelity scores for behavioral strategies within routines, PRT, and, DTT for each classroom were based upon the average of scores for the classroom staff in winter and spring.

Student Eligibility and Assessment Measures

The CAP model was designed to improve student outcomes in core areas of autism (i.e., expressive language, receptive language, and social skills) and other curricular areas important to success in school (e.g., cognitive skills, adaptive behavior). Multiple instruments were used to collect data on these outcome areas from students, teachers, and parents. The same data was collected from both the CAP and BAU groups. The *CARS* (Schopler et al. 1988) was administered by certified teachers at baseline (summer/fall before intervention began), and it was used to determine study eligibility and severity of autism. According to the developers, special educators and other professionals "…can be trained through brief written or videotaped instructions to administer the *CARS"* (Schopler et al. 1988, p. 6). Teachers, who already had training and experience in completing behavioral observations, were provided with an explanation of how to complete the rating scale, but since *CARS* (Schopler et al. 1988) was already commonly used in public school settings, most of the educators were already familiar with the instrument.

The other seven instruments were administered at baseline and at the end of the school year. The *ASIEP-3* subtests were specifically designed for children with ASD and have successfully been used in prior studies to measure social interaction and spontaneous expressive language (Arick et al. 2003). The other instruments are all well-known measures that have been used in a number of research studies. Following is a brief description of each instrument, including reliability coefficients for children, ages 3–5, as reported by the developers.

Study Eligibility and Severity of Autism

The *CARS* (Schopler et al. 1988) was designed to identify children with ASD. The internal consistency is .94, the average inter-rater reliability is .71, and the test–retest

reliability is .88. Research by the developers found that professionals from related fields (e.g., special educators, speech language pathologists) can successfully use this instrument and their scores highly correlate (.83) with the scores of clinical directors with extensive experience (Schopler et al. 1988). To meet study eligibility criteria, students were required to have a medical diagnosis for ASD or a total *CARS* (Schopler et al. 1988) score of ≥30 (autism range). Students' baseline *CARS* (Schopler et al. 1988) scores were also used to measure the severity of the child's ASD.

Adaptive Behavior

The *Vineland Adaptive Behavior Scales—Second Edition (VAB-II)—Survey Interview Form—Daily Living Skills Domain* (Sparrow et al. 2005) measure an individual's adaptive behavior. The daily living skills domain consists of three subdomains: personal, domestic, and community. The test–retest reliability and the split-half reliability for this domain are .90; and the inter-interviewer reliability is .87.

Cognitive Ability

The *Battelle Developmental Inventory-2*—(BDI-2)—*Cognitive Domain* (Newborg 2005) assesses an individual's developmental skills. The Cognitive Domain measures conceptual skills and abilities. It includes three subdomains: Attention and Memory, Reasoning and Academic Skills, and Perceptions and Concepts. The split-half reliability average for the cognitive domain is .94, and the test–retest reliability is .87.

Expressive Language (Two Instruments)

The *Expressive One Word Picture Vocabulary Test* (EOWPVT) (Brownell 2000a) assesses an individual's speaking vocabulary. The corrected internal consistency reliability ranges from .96 to .98; and the corrected test–retest reliability ranges from .88 to .89. The *ASIEP-3 Sample of Vocal Behavior* (Krug et al. 2008) assessment was also used. During this observational assessment, 50 utterances (e.g., sounds, words) made by the student are scored to evaluate the student's spontaneous speech and measure communication expressed by vocalizations accompanied by gestures or other means. The split-half reliability is .97; and the test–retest reliability ranges from .81 to .94. Students in this study had a broad range of expressive language skills. Some spoke no words, while others expressed themselves better than the norm for their age. To accommodate this spectrum, we used both of these instruments to allow us to obtain an age-equivalency score for each student. At their baseline assessment, each student was first presented with test plates from the EOWPVT (Brownell). If the student could accurately name most of the example pictures, this instrument was used at their baseline and follow-up assessments. If the student was not able to respond to the example pictures, the *ASIEP-3 Sample of Vocal Behavior* (Krug et al. 2008) was used at baseline and at the follow-up assessment. The *ASIEP-3 Sample of Vocal Behavior* (Krug et al. 2008) could not be used for all students, because it had a test ceiling of 48 months. Approximately 50 % percent of students were assessed using the *ASIEP-3 Sample of Vocal Behavior* (Krug et al. 2008) and 50 % were assessed using the *EOWPVT* (Brownell).

Receptive Language

The *Receptive One Word Picture Vocabulary Test (ROWPVT).* (Brownell 2000b) measures an individual's hearing vocabulary. The corrected internal consistency reliability is .98; and the corrected test–retest reliability ranges from .81 to .93.

Social Skills (Observation by Research Staff)

The *Autism Screening Instrument for Educational Planning—3rd Edition (ASIEP-3): Interaction Assessment* (Krug et al. 2008) was administered by research staff. The assessor records the child's interaction with the adult, constructive play, aggressive behavior, self-stimulation, self-abuse, and absence of observable behavior during 10-s intervals of a 12-min observation period with an adult in a play setting. The observation period includes three 4-min phases: (a) adult provides active modeling, (b) adult remains passive, and (c) adult gives the child direct cues. The child's behavior is recorded and coded by a second adult during each phase. The item reliability is .86, and inter-rater agreement level is listed as 89 % from 87 raters.

Social Skills (Teacher and Parent Ratings)

The *Social Skills Rating System* (*SSRS*) (*Parent and Teacher Versions*) (Gresham and Elliot 1990) were completed by teachers and parents to rate social behaviors and problem behaviors that could affect social competence at school and home. The test–retest reliability for the Social Skills Domain is .85 for the teacher version and .87 for the parent version; and for the Problem Behavior Domain the test-reliability is .84 for the teacher version and .65 for the parent version. The internal consistency reliability for the Social Skills Domain is .94 for the teacher version and .90 for the parent version; and for the Problem Behavior Domain the test–retest reliability is .82 for the teacher version and .73 for the parent version.

Measures Developed by Research Team

The *Family Demographic Information Form* was used by parents to report baseline information for each child's family (e.g., ethnicity, home language, income level, parent education level). A *Student Demographic Information Form* collected baseline information from the teacher about the student (e.g., date of birth, gender, diagnosis, eligibility for special education). The *Student Program Questionnaire* gathered baseline and follow-up information from the teacher regarding each student's specific program (e.g., number of hours each week the student attended school, program information, and related services student received). An *Educator Experience Form* collected baseline information from teachers regarding their previous training, education, and experiences (e.g., years in teaching profession, highest educational degree held). An *Exit Survey* enabled teachers to provide feedback and to assess the social validity of the CAP intervention practices and implementation methods.

Data Collection Procedures

Measures Completed by Parents and Teachers

Forms, assessments, and surveys were mailed to participants along with a postage-paid, pre-addressed return envelope. If the form or assessment was not completed and returned by a specific due date, a follow-up phone call and email were sent to the participants. Research staff was available during all business hours to assist participants. For their participation, parents and teachers received a $50.00 gift card.

Measures Administered by Research Staff

Individual student assessments [i.e., *BDI-2—Cognitive Domain* (Newborg 2005); *ASIEP-3: Sample of Vocal Behavior* (Krug et al. 2008); *EOWPVT* (Brownell 2000a); *ROWPVT* (Brownell 2000b); and *ASIEP-3: Interaction Assessment* (Krug et al. 2008)] were conducted at each school by research staff with masters' or doctoral degrees and extensive experience working with students with ASD. To ensure assessors were blind to treatment condition, teachers were told not to identify their treatment condition if they spoke with the assessor during a visit to their schools. Additionally, another research staff member accompanied the assessor to the school and was responsible for accompanying the child to the assessment area (e.g., conference room).

Research staff also conducted baseline (fall) and follow-up (spring) phone interviews, both in English and Spanish, to collect data from teachers and parents. In phone interviews, both CAP and BAU teachers used their records (e.g., student individual education plan [IEP]) to report student program information, including daily hours spent at school each day, classroom placement (e.g., self-contained classroom), and type and amount of one-to-one instruction received (e.g., DTT, PRT). Additionally, the *VAB-II* (Sparrow et al. 2005) was conducted by phone interviews with parents. Other researchers have used this method and found *VAB-II* telephone interviews to be highly correlated (.98) with in-person interviews (Limperopulos et al. 2006).

Analysis

Due to the hierarchical structure of the data collected in this cluster RCT (students nested in schools), hierarchical linear modeling (HLM) was used to estimate the treatment impact. HLM enabled us to model the nested structure of the data appropriately and also enabled tests of the impact of moderating variables. One of the advantages of multilevel modeling using modern software is the ability to accurately handle different cell sizes and varying amounts of missing data. Several preliminary analyses were conducted prior to the impact estimate to ensure the validity of the estimate. These included descriptive analysis, attrition analysis, and baseline equivalency checks on the pre-assessment scores. Our original plan was to include only pre-assessment scores in the HLM as the covariate. However, since the descriptive analysis found group differences in the students' *CARS* (Schopler et al. 1988) scores and the days between student assessments, those variables were also considered as potential covariates in the HLM. Comparisons between HLMs with and without those additional covariates found that the model with those additional covariates had a superior fit to the data. Consequently, these two covariates were retained in the final analysis model:

$$y_{ij} = b_0 + b_1[\mathrm{TRT}]_j + b_2[\mathrm{PRE}]_{ij} + b_3[\mathrm{days\_between}]_{ij} + b_4[\mathrm{CARStotal}]_{ij} + d(k)\mathrm{Cohort}(k) + u_j + e_{ij}$$

The variables shown in the final model include: [TRT] = treatment condition; [PRE] = students' pre-assessment scores; [days_between] = days between student assessments; [CARS total] = students' baseline CARS (Schopler et al. 1988) score; and Cohort = a set of indicator variables for Cohort. The subscript $_i$ refers to an individual student $_i$, the $_j$ refers to a school $_j$, the e_{ij} refers to the student-level residual, and the u_j refers to the school-level error.

During the preliminary analyses, a group difference was also found in the mean age of the students at the first assessment (BAU mean was higher). However, the HLM with this additional covariate had better fit to the data only

for one of the outcome measures, *SSRS* problem behaviors score—teacher version (Gresham and Elliot 1990). Consequently, for this outcome only, age at first assessment was also included in the final analysis model as a covariate.

Aside from the impact estimate, statistical analyses were conducted to investigate the moderating effects of severity of ASD as measured by the students' *CARS* (Schopler et al. 1988) score. Based on their CARS score, students were placed into three categories: (a) score <30 (did not meet the CARS criteria for autism, but had a medical diagnosis of ASD); (b) score 30–37 (mild to moderate autism); and (c) score over 37 (severe autism). We then estimated treatment impact for each category, using a model in which the overall treatment effect is represented by three "Treatment x ASD Severity Category" interaction terms.

Results

Participants

Seventy-eight schools were nominated to participate in this study. Each school had one to three participating teachers. Forty-one schools were randomly assigned to the CAP condition, and 37 schools were randomly assigned to the BAU condition (see Fig. 1). Nine schools, considered to be eligible before the study began (three CAP and six BAU), were found not to meet the enrollment criteria and removed from the study (i.e., no eligible students; inadequate number of students; some staff participated in previous cohort). One CAP school withdrew due to a change in the teacher's workload. Three schools were not included in the final HLM analysis (two CAP and one BAU), because the teachers did not complete the *CARS* (Schopler et al. 1988) for any of the students in their schools. Twelve additional students were also excluded from the final analysis due to missing *CARS* (Schopler et al. 1988) scores. The *CARS* (Schopler et al. 1988) score was a covariate in the HLM model and students were required to have this score for their outcomes to be evaluated. Final analysis included 35 CAP schools (92 % retained) and 30 BAU schools (97 % retained). Even after attrition, the recommended sample size of 25 schools in each condition (U.S. Department of Education 2003) was met.

Table 2 describes the baseline characteristics of the 302 (178 CAP and 124 BAU) children and 84 teachers who participated in this study. One hundred percent of the teachers were female. The majority of the teachers had both special education and general education licenses. There were no statistically significant differences found in the baseline characteristics of the teachers in the CAP and BAU groups. Most of the students in the study were white, not Hispanic, (56 % CAP and 61 % BAU) and male (82 % CAP and 78 % BAU). Students in both groups spent approximately 16 h a week at school. Their primary school placements included (a) self-contained classrooms with all students with ASD or students with mixed disabilities or (b) inclusive classrooms with typical peers. There were no statistically significant differences found when examining their primary classroom placements or student-to-teacher ratio for their classrooms. There were statistically significant differences found in age-in-months at their first assessment (51.71 CAP and 53.96 BAU) and in their baseline *CARS* (Schopler et al. 1988) scores (38.25 CAP and 34.71 BAU). These last two variables were included as covariates in the HLM models used in the impact analysis.

Baseline equivalence was assessed (see Table 3), using data for students who were originally assigned to condition and had a post-score for each outcome measure. No student outcome suffered from the combination of high attrition and baseline non-equivalence exceeding the effect size of 0.25. The analytic sample was defined for each outcome, and its size differed from one outcome to another. The *n* was higher for assessments completed by the research team and teachers; *n* was lower for assessments completed by parents (i.e., parents had lower return rate).

Fidelity of Implementation and Intervention

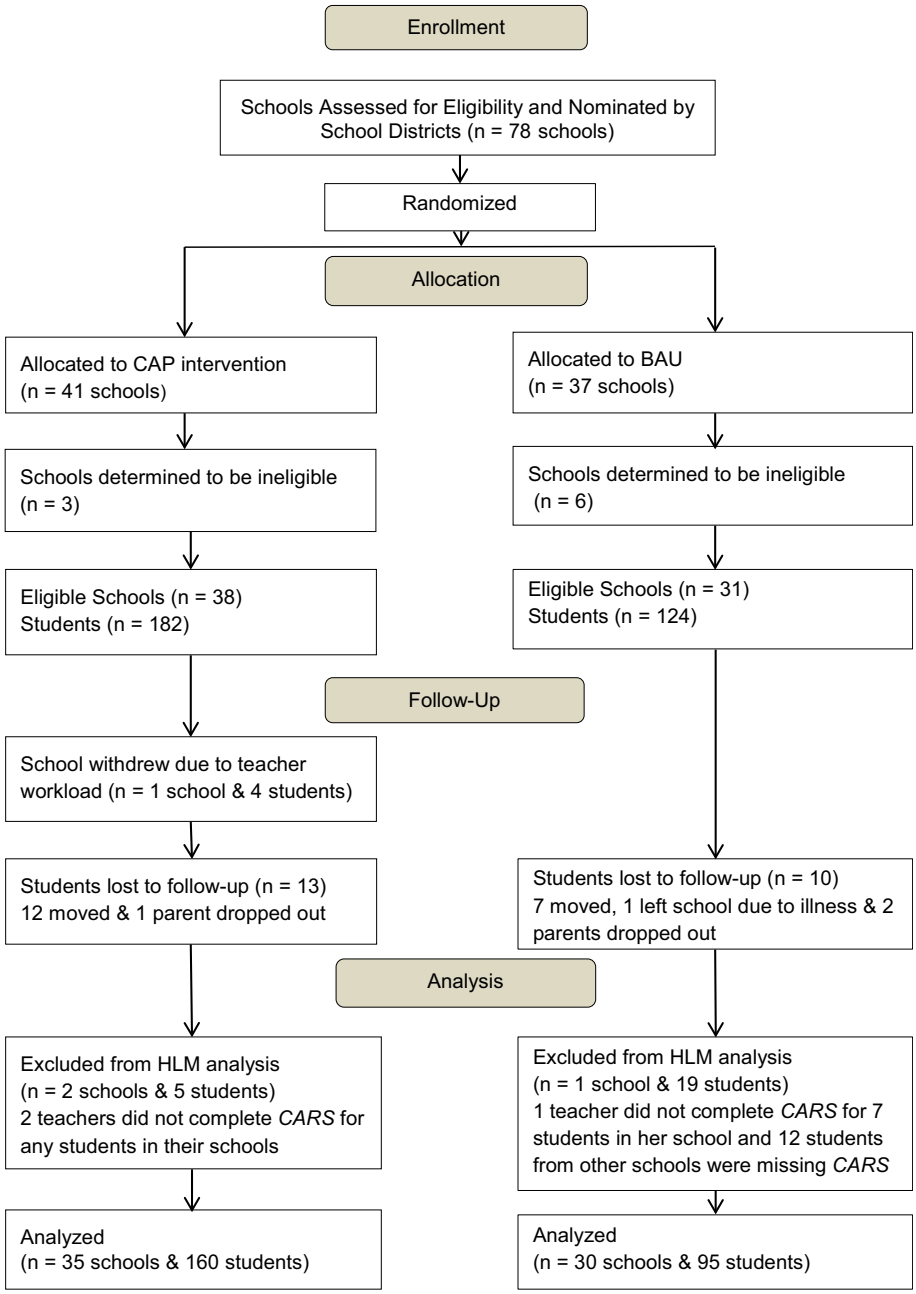
Implementation of Training for Staff and Parents of CAP Schools

Ninety-eight percent of teachers and at least one paraprofessional from each school attended initial workshops. A speech language pathologist from 70 % of schools also attended. The two teachers who did not attend workshops, due to illness, were provided with extra training during or following the initial 2 days of in-class assistance for their school. All teachers and at least one paraprofessional participated in the initial days of in-class training and at least six monthly follow-up visits. Seventy percent of the families participated in at least one parent workshop, and approximately 69 % of those attendees participated in at least one follow-up consultation in their home or at their child's school.

Fidelity of Intervention by Educators in CAP and BAU Schools

Table 4 presents classroom fidelity data on EBPs collected by project staff during fidelity visits and minutes per day of implementation of EBPs with each student as reported in post-interviews by teachers in CAP and BAU classes. Statistically significant differences between CAP and BAU

Fig. 1 Flow of participants



educators were seen in fidelity of intervention for environmental arrangement and visual supports, PRT, and DTT. Educators in CAP classrooms used behaviors identified on the fidelity checklists for these practices with greater consistency than educators in BAU classrooms. No significant difference was found between CAP and BAU educators in fidelity of behavioral strategies within routines. Students in CAP classrooms were provided with significantly more minutes of behavioral teaching strategies within routines and PRT. No significant difference was found between the CAP and BAU classrooms in the minutes of DTT provided to students.

Impact Analysis

Theoretically in a RCT, the difference in the treatment and the control group means represents the treatment impact. However, in this study inclusion of covariates was necessary for adjusting the non-equivalence between the two groups and for improving the precision of impact estimate. Table 5 presents the estimated impact as the difference between CAP and BAU students' estimate of covariate-adjusted posttest scores. The standard error of estimate was adjusted for the nesting of students within schools. CAP had a positive impact on the students' receptive language

Table 2 Baseline teacher and student characteristics

Teacher characteristics	CAP		BAU		χ^2	*p*
	n = 44 Teachers		*n* = 40 Teachers			
	n = 178 Students		*n* = 124 Students			
	n	%	*n*	%		
Highest educational degree					1.798	.18
Bachelor's degree	12	27.3	16	40		
Master's degree	31	70.4	22	55		
Missing	1	2.3	2	5		
Licensure					.464	.79
Special education only	18	40.9	14	35		
General education only	2	4.5	3	7.5		
Both special and general education	23	52.3	19	47.5		
Missing	1	2.3	4	10		

	n	*M/SD*	*n*	*M/SD*	*t*	*p*
Years in teaching profession	41	13.06/10.30	37	14.65/8.60	−.73	.47
Missing	3		3			

Student characteristics	*n*	%	*n*	%	χ^2	*p*
Gender					.480	.49
Male	145	81.5	97	78.2		
Female	33	18.5	27	21.8		
Ethnicity					.869	.97
African American	13	7.3	10	8.1		
American Indian or Alaskan Native	1	.6	1	.8		
Asian or Other Pacific Islander	20	11.2	11	8.9		
Hispanic	21	11.8	12	9.7		
White (Not Hispanic)	99	55.6	75	60.5		
Other	22	12.4	15	12.1		
Missing	2	1.1	0	0		
Primary classroom placement					2.61	.27
Self-contained—students with only ASD	79	44.4	44	35.5		
Self-contained—students with mixed disabilities	40	22.5	35	28.2		
Classroom with typical peers	59	33.1	45	36.3		

	n	*M/SD*	*n*	*M/SD*	*t*	*p*
Student to teacher ratio—primary placement						
Self-contained—students with only ASD	79	2.18/.81	44	2.08/.69	.62	.53
Self-contained—students with mixed disabilities	40	2.68/.89	35	2.99/1.07	−1.41	.16
Classroom with typical peers	59	2.91/.91	45	4.29/4.80	−1.90	.06
Hours per week child attends school	178	15.62/6.79	123	16.11/6.63	−.62	.53
Missing	0		1			
Age in months at first assessment	171	51.71/9.16	114	53.96/8.46	−2.09	.04*
Missing	7		10			
Baseline CARS score	172	38.25/9.11	103	34.71/9.86	3.03	.003*
Missing	6		21			

* $p \leq .05$

Table 3 Assessment of baseline equivalence

Measure	CAP		BAU		*Hedges' g*
	n	*M/SD*	*n*	*M/SD*	
ASIEP-3 Interaction Assessment[e,f]	155	51.61[f]/17.32	100	46.89/17.70	0.27
Battelle Developmental Inventory (BDI-2): Cognitive subdomain[c]	154	6.36/5.43	102	7.28/7.69	−0.14
Expressive One-Word Vocabulary Test (EOWPVT) or ASIEP-3 Sample or Vocal Behavior Test[b]	152	27.07/13.86	100	31.55/16.36	−0.30
Receptive One-Word Picture Vocabulary Test[d] (RCWPVT)	153	16.29/18.28	98	23.49/19.87	−0.38
Social Skills Rating Scale (SSRS) Problem Behaviors: Teacher Form[a,g]	152	106.86/11.88	101	107.46/9.89	−0.05
Social Skills Rating Scale (SSRS) Social Skills: Teacher Form[a]	152	64.49/16.47	101	71.05/17.97	−0.38
Social Skills Rating Scale (SSRS) Problem Behaviors: Parent Form[a,g]	87	105.07/12.62	53	107.51/15.10	−0.18
Social Skills Rating Scale (SSRS)	89	67.11/13.97	53	70.17/16.02	−0.21
Social Skills: Parent Form[a]					
Vineland Adaptive Behavior Scales(VAB-II)	118	76.75/15.83	76	76.38/15.24	0.02
Daily Living Skills Subdomain[a]					

[a] Standard score; [b] language age equivalency score; [c] scaled score; [d] raw score; [e] total score; [f] lower score indicates more interaction with adult; [g] lower score indicates fewer problem behaviors reported

Table 4 Classroom fidelity and daily use of EBPs with students

Classroom fidelity of EBPs	CAP		BAU		*t*	*p*
	n	*M/SD*	*n*	*M/SD*		
Environment and visuals	44	1.41/0.41	38	0.98/0.39	3.82[a]	<.001[a]
Range 0 (clearly not present)–2 (clearly present)						
PRT	40	3.68/0.59	23	2.98/0.74	3.20[a]	.002[a]
Range: 0 (teacher does not implement) to 5 (teacher implements throughout session)						
DTT	40	3.89/0.62	27	2.99/0.58	5.27[a]	<.001[a]
Range: 0 (teacher does not implement) to 5 (teacher implements throughout session)						
Behavioral strategies within routines	43	3.49/0.64	37	3.35/0.72	0.79[a]	.43[a]
Range: 0 (teacher does not implement) to 5 (teacher implements throughout session)						

Daily minutes of one-to-one EBPs with students	*n*	*M/SD*	*n*	*M/SD*	*t*	*p*
Behavioral strategies within routines	142	12.08/22.44	86	2.06/6.61	4.97	<.001
PRT	142	10.60/12.66	86	.52/2.77	9.13	<.001
DTT	142	13.64/11.96	86	12.48/22.25	.448	.66

[a] Test statistic and *p* value are adjusted for clustering teachers to schools

(effect size of .13), as measured by the *ROWPVT* (Brownell 2000b) and on their social skills (effect size of .19), as measured by the teachers' rating on the *SSRS—Social Skills Domain* (Gresham and Elliot 1990). There were no positive impacts on students' social interaction as measured by *ASIEP-3 Interaction Assessment* (Krug et al. 2008); cognitive skills as measured by the *BDI-2—Cognitive Domain* (Newborg 2005); expressive language as measured by the *EOWPVT* (Brownell 2000a) or the *ASIEP-3 Sample of Vocal Behavior* (Krug et al. 2008); social skills as measured by the *SSRS—Social Skills Domain—Parent Version* (Gresham and Elliot 1990); problem behaviors as measured by the *SSRS Problem Behavior Domain—Parent and Teacher versions* (Gresham and Elliott) or on adaptive behaviors as measured by the *VAB-II—Daily Living Domain* (Sparrow et al. 2005).

Moderator Analysis

Initial statistical analyses suggested that severity of ASD at baseline might influence the effectiveness of the treatment. Further analysis indicated that the treatment effect was moderated by severity of ASD as measured by the *CARS* (Schopler et al. 1988) score. The total *CARS* (Schopler

Table 5 Impact of CAP adjusted for pretest score, CARS and days between

Outcome measure	CAP *mean (n)*	BAU *mean (n)*	Impact: CAP – BAU difference *(SE)*	Test statistic *z* *p*	Effect size Hedges' *g*
ASIEP-3 Interaction Assessment	42.11 (151)	41.56 (85)	0.55 (1.82)	$z = 0.30$ $p = 0.76$	0.04
BDI-2: Cognitive Domain (log-transformed)	1.63 (150)	1.69 (85)	−0.06 (0.06)	$z = -0.97$ $p = 0.33$	−0.08
EOWPVT or ASIEP-3 Sample of Vocal Behavior	38.05 (147)	38.58 (82)	−0.53 (1.40)	$z = -0.38$ $p = 0.71$	−0.03
ROWPVT	31.45 (149)	28.31 (83)	3.14 (1.46)	$z = 2.15$ $p = 0.03$	0.13
SSRS: Problem Behavior (Teacher)*	105.12 (143)	105.14 (79)	−0.02 (1.41)	$z = -0.01$ $p = 0.99$	−0.002
SSRS: Social Skills (Teacher)	79.943 (148)	76.302 (85)	3.6401 (1.89)	$z = 1.93$ $p = 0.05$	0.19
SSRS: Problem Behavior (Parent)	106.34 (86)	108.15 (43)	−1.82 (2.02)	$z = -0.90$ $p = 0.37$	−0.13
SSRS: Social Skills (Parent)	73.00 (88)	72.98 (43)	0.02 (2.16)	$z = 0.01$ $p = 0.99$	0.001
VAB-II: Daily Living Skills Subdomain	78.73 (116)	78.37 (63)	0.36 (1.67)	$z = 0.21$ $p = 0.83$	0.02

The covariate-adjusted outcome represents the estimated outcome score of a student in Cohort 2 who had a baseline outcome score, a *CARS* score, and days between assessment score, all at the grand mean. Cohort 2 was used as the referent cohort in the analysis, as it was the largest cohort

* *SSRS: Problem Behavior* (*Teacher*) was also adjusted for students' age at first assessment

et al. 1988) score may range from 15 (behaviors rated within normal limits on all 15 items scored) to a high of 60 (severely impacted by ASD). The positive impact of CAP on receptive language, as measured by the *ROWPVT* (Brownell 2000b), was mostly driven by CAP having a strong impact on students with *CARS* (Schopler et al. 1988) scores from 30 to 37. The positive impact on social skills, as measured by the teachers' ratings on the *SSRS* (Gresham and Elliot 1990) was mostly driven by CAP having a strong impact on students with scores <30. These students had a medical diagnosis of ASD, but fell into the non-autistic range on the *CARS* (Schopler et al. 1988).

Social Validity

At the end of the school year, CAP teachers were asked to anonymously complete an exit survey to report on the social validity of CAP. Eighty-five percent of the teachers completed the survey. One hundred percent of respondents agreed or strongly agreed they learned new skills to help their students make progress and their students benefited from participation in the project. Ninety-one percent of the teachers planned to continue using the model the following year.

Discussion

This study compared outcomes for 3–5-year-old students with ASD in classrooms implementing the CAP model with those in classrooms implementing BAU. Students in both the CAP and BAU made improvements in all outcome areas, except problem behaviors as reported by parents, which essentially stayed the same for both groups. HLM analysis revealed CAP had a positive impact on students' outcomes for receptive language and social skills at school compared to students' outcomes in BAU schools. Treatment effects were moderated by severity of ASD. Strongest positive impact on receptive language was for students with mild to moderate ASD. The CAP's strongest impact on students' social skills was for those who had medical diagnoses of ASD but fell into the non-autistic range on the *CARS* (Schopler et al. 1988).

Positive impacts on the CAP students' receptive language and social skills were achieved after 6–8 months of implementation. Due to projected changes in students' school placements, it was not possible to continue the study with the same students for multiple years. Further investigation is needed to address this time limitation and identify the effects of CAP when implemented over a

longer period of time. Strain and Bovey (2011), suggested that at least 2 years of training and coaching for classroom staff in the implementation of the LEAP Model for young children with ASD was needed to achieve significant group differences across a comprehensive set of curricular areas.

Observational measures found statistically significant differences between CAP classroom staff and BAU classroom staff in fidelity of intervention using EBPs for environmental arrangement and visual supports, PRT, and DTT. Additionally, CAP educators reported using EBPs for behavioral strategies within routines and PRT in one-to-one instruction with students with ASD for significantly more minutes per day than educators in BAU classrooms. These differences in fidelity and minutes of EBPs may be due to use of the *STAR Program* (Arick et al. 2004) lesson plans and the training and coaching provided by CAP staff.

There were no statistically significant differences in minutes of one-to-one DTT instruction. This reflected the fact that some of the BAU teachers were using a substantial number of minutes of DTT in their classrooms. While these teachers reported providing a substantial number of minutes of DTT per day, observations found lower fidelity to quality practices for DTT for BAU educators than that observed in CAP schools. This suggests that teacher-report of use of this EBP may not necessarily reflect use which meets standards for fidelity. Additionally, BAU teachers were often not using DTT in combination with other naturalistic interventions, such as incidental teaching and PRT within routines, as recommended by Smith (2001). Our observations also supported the findings of Boyd et al. (2014) and comments by Mandell et al. (2013) regarding the importance of other indicators of classroom quality, such as organization of the classroom environment and scheduling.

There was overlap between practices used in CAP and those used in BAU classrooms. Differences between CAP and BAU were primarily in the fidelity to quality guidelines and extent of daily, one-to-one use of the EBPs, rather than in the types of practices used. Additionally, while fidelity of intervention for the EBPs of environmental arrangement and visual supports, PRT and DTT for CAP educators was significantly greater than fidelity to these practices for BAU educators, fidelity of CAP educators ranged from 70 % (environment, visuals, and behavioral teaching strategies within routines) to 74 % (PRT) to 78 % (DTT). These ratings are somewhat less than the standard of 80 % fidelity referred to by Mandell et al. (2013). More frequent, intensive training and coaching over a longer period of time may be needed for CAP educators to reach higher levels of fidelity.

Using the *CARS* (Schopler et al. 1988) scores, we found that treatment effects were moderated by severity of ASD. Other studies also identified differences in response to intervention related to severity of ASD (e.g., Ben-Itzchak and Zachor 2007; Sandall et al. 2011). Additional research is needed, using a stronger measure such as the *Autism Diagnostic Assessment Schedule* (Lord et al. 2012), to better understand ways to match intervention strategies to the different profiles of strengths and weaknesses experienced by students with ASD. Further studies of moderating effects of the impact of ASD on outcomes of CTMs may demonstrate that comprehensive models need the flexibility to adapt the selection of curricula and EBPs for students according to their profile of ASD symptoms and skill levels.

No positive impacts for CAP, compared to BAU, were found for students' adaptive behavior, cognitive abilities, expressive language, social skills as rated by parents, or social interaction with an adult in a play setting. These differences in impact may indicate the need for longer implementation of the CAP intervention to see positive effects in these areas. Additionally, other curricula or EBPs to address these areas may be needed. For example, we observed that some students needed instruction in academic skills not included in the *STAR Program* (Arick et al. 2004). Different academic curricula or evidence-based practices for adapting the general education curriculum for students with ASD may be needed for these students. Similarly, while CAP had a positive impact on social skills, this impact was driven by its effects on students with less severe impact of ASD. Other curricula or EBPs may be needed to address social skills for students most impacted by ASD. Our results support further study of approaches involving differentiated curricula and EBPs, based upon the effects of ASD and skill needs for specific students.

A potential strength of the CAP is the component of parent education, including a curriculum and on-going coaching. Our original plan was to include educators in the home visits, but due to contractual or logistical issues, that was not a possibility in most of the school districts. This resulted in CAP coaches conducting home visits alone and limited the number of visits. Parents' comments in follow-up surveys indicated that CAP parents were using EBPs (e.g., visual supports and PRT within routines) more often than BAU parents. However, without observation data, we have no way to confirm the frequency or fidelity of this use.

While 100 % of the CAP parents received the *APP* (Young et al. 2009) curriculum, not all attended workshops or participated in home visits. An extensive effort was made to inform parents of the workshops (i.e., phone calls, emails, letters, flyers). Child care was also provided at each workshop to enable parents to attend. However, only 70 % of the parents participated in the trainings and 69 % of those attendees elected to have a follow-up visit. Limited participation made it difficult for us to interpret whether the

parent education component affected the student outcomes. In future studies of CAP, consent from districts for teachers to participate in home visits, adequate funds for on-going coaching of parents, and measures of fidelity in the home are needed. Additional research that focuses on the parent component to study its impact on student outcomes is also needed.

Results of the CAP study support a model that may be generalizable to many real-world educational settings. CAP classrooms were located in diverse urban and rural public schools and served students across the range of severity of ASD. Given initial training and monthly coaching for use of EBPs, educators in CAP schools implemented environmental arrangement and visual supports, PRT, and DTT with better fidelity and implemented PRT and behavioral strategies within routines for more time per day than educators in BAU classrooms. Implementation of EBPs was carried out by classroom teams, including teachers, paraeducators, speech-language pathologists, and a few occupational therapists. Paraeducators played a large role in implementation of these EBPs with students and demonstrated their valuable contribution in a comprehensive program. Positive impacts in two core areas of ASD (i.e., receptive language and social skills) were also achieved within schools that provided an average of 16 h of service per week for students with ASD, in contrast to the 20–40 h per week often recommended (NRC 2001). This number of hours was similar to the 17 h of instruction reported by Strain and Bovey (2011).

This study adds to the small but growing number of school-based RCTs investigating CTMs for young students with ASD. The results reported here indicate the importance of further study of the CAP model. We plan further analysis of the relationship between fidelity of implementation of the CAP components by project staff and fidelity of intervention practices by educators in these CAP study outcomes. Additionally, further analysis of follow-up surveys regarding teachers' perceptions of social validity and sustainability of CAP should add to our understanding of the roles of implementation components such as use of a curriculum, training and coaching practices. Because of the moderating effect of severity of ASD on outcomes, we also plan further analyses to investigate differences in response to the CAP intervention for students with differences in impact of ASD. Future studies, with full implementation of the CAP model, including the parent education component, for a longer period of time, are needed to replicate and expand our results.

Acknowledgments We thank project staff, who served as trainers, assessors, and office personnel: Jocelyn Bray, Mindy Campbell, Mikal' Davis, David Krug, Laurel Oliver-Gilmore, Jennifer Reinmuth, Tiffany Sanford, Corrine Young, and numerous student research assistants. We also thank our consultants, Richard Simpson, Aubyn Stahmer, and Michael Coe, for their thoughtful advice. Finally, we thank the numerous school districts, educators, children, and families who participated in this project.

Author Contributions HEY led implementation and coordination of the study, participated in the statistical analysis and interpretation of the data, and drafted major sections of the manuscript; RAF participated in implementation of the study, contributed to interpretation of the data, and drafted major sections of the manuscript; MH led the statistical analysis, contributed to interpretation of the data and writing of the manuscript. All authors read and approved the final manuscript.

Funding This study was funded by the U.S. Department of Education, Institute of Education Sciences, Award Number R324A090094.

Compliance with Ethical Standards

Conflict of interest Helen E. Young declares no conflict of interest. Ruth A. Falco is a joint-author of the *STAR Program* (Arick et al. 2004); she receives royalties from the publisher, PRO-ED. Makoto Hanita declares no conflict of interest.

References

Arick, J. R., Loos, L., Falco, R., & Krug, D. A. (2004). *The STAR program: Strategies for teaching based on autism research*. Austin, TX: Pro-Ed.

Arick, J. R., Young, H. E., Falco, R. A., Loos, L. M., Krug, D. A., Gense, M. H., & Johnson, S. B. (2003). Designing an outcome study to monitor the progress of students with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 18*(2), 74–86.

Ben-Itzchak, E., & Zachor, D. A. (2007). The effects of intellectual functioning and autism severity on outcome of early behavioral intervention for children with autism. *Research in Developmental Disabilities, 28*, 287–303. doi:10.1016/j.ridd.2006.03.002.

Boyd, B. A., Hume, K., McBee, M. T., Alessandri, M., Gutierrez, A., Johnson, L., & Odom, S. L. (2014). Comparative efficacy of LEAP, TEACCH and non-model-specific special education programs for preschoolers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44*(2), 366–380. doi:10.1007/s10803-013-1877-9.

Boyd, B., & Wong, C. (2013). *Response interruption/redirection (RIR) fact sheet*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, The National Professional Development Center on Autism Spectrum Disorders.

Brownell, R. (2000a). *Expressive one-word picture vocabulary test*. Novato, CA: Academic Therapy Publications.

Brownell, R. (2000b). *Receptive one word picture vocabulary test*. Novato, CA: Academic Therapy Publications.

Cox, A. W. (2013a). *Modeling fact sheet*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, The National Professional Development Center on Autism Spectrum Disorders.

Cox, A. W. (2013b). *Prompting (PP) fact sheet*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, The National Professional Development Center on Autism Spectrum Disorders.

Dingfelder, H. E., & Mandel, D. S. (2011). Bridging the research-to-practice gap in autism intervention: An application of diffusion of innovation theory. *Journal of Autism and Developmental Disorders, 41*, 597–609. doi:10.1007/s10803-010-1081-0.

Dunst, C. J., Trivette, C. M., & Raab, M. (2013). An implementation science framework for conceptualizing and operationalizing fidelity in early childhood intervention studies. *Journal of Early Intervention, 35*(2), 85–101. doi:10.1177/1053815113502235.

Elliott, C. (1990). *Differential ability scales: Administration and scoring manual* (2nd ed.). San Antonio, TX: Psychological Corp.

Fixen, D.L., Naoom, S.F., Blase, K.A., Friedman, R.W., & Wallace, F. (2005). *Implementation research: A synthesis of the literature.* Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231). Retrieved from http://nirn.fpg.unc.edu/sitesunc.edu/files/resources/NIRN-MonographFull-01-2005.pdf

Fleury, V. P. (2013a). *Task analysis (TA) fact sheet*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, The National Professional Development Center on Autism Spectrum Disorders.

Fleury, V. P. (2013b). *Time delay (TD) fact sheet*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, The National Professional Development Center on Autism Spectrum Disorders.

Fleury, V. P. (2013c). *Discrete trial training (DTT) fact sheet*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, The National Professional Development Center on Autism Spectrum Disorders.

Frost, L., & Bondy, A. (2002). *The picture exchange communication system training manual* (2nd ed.). Cherry Hill, NJ: Pyramid Educational Consultants.

Gresham, F. M., & Elliot, S. N. (1990). *Social skills rating system*. Circle Pines, MN: American Guidance Service.

Hume, K. (2013a). *Antecedent-based intervention (ABI) fact sheet*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, The National Professional Development Center on Autism Spectrum Disorders.

Hume, K. (2013b). *Visual supports (VS) fact sheet*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, The National Professional Development Center on Autism Spectrum Disorders.

Hume, K., Pavnik, J., & Odom, S. L. (2012). Promoting task accuracy and independence in students with autism across educational setting through the use of individual work systems. *Journal of Autism and Developmental Disorders, 42*, 2084–2099. doi:10.1007/s10803-012-1457-4.

Kasari, C., & Smith, T. (2013). Interventions in schools for children with autism spectrum disorders: Methods and recommendations. *Autism, 17*(3), 254–267. doi:10.1177/1362361312470496.

Krug, D. A., Arick, J. R., & Almond, P. (2008). *Autism screening instrument for educational planning*. Austin, TX: Pro-ed.

Kucharczyk, S. (2013). *Reinforcement (R+) fact sheet*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, The National Professional Development Center on Autism Spectrum Disorders.

Limperopulos, C., Majnemer, A., Steinbach, C. L., & Shevell, M. I. (2006). Equivalence reliability of the Vineland Adaptive Behavior Scale between in-person and telephone administration. *Physical and Occupational Therapy in Pediatrics, 26*(1/2), 115–127. doi:10.1300/J006v26n01_08.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic assessment schedule (ADOS-2)* (2nd ed.). Torrace, CA: Western Psychological Services.

Mandell, D. S., Stahmer, A. C., Shin, S., Xie, M., Reisinger, E., & Marcus, S. C. (2013). The role of treatment fidelity on outcomes during a randomized field trial of an autism intervention. *Autism, 17*(3), 281–295. doi:10.1177/1362361312473666.

Mesibov, G. B., & Shea, V. (2010). The TEACCH Program in the era of evidence-based practice. *Journal of Autism and Developmental Disorders, 40*, 570–579. doi:10.1007/s10803-009-0901-6.

National Research Council. (2001). Educating children with autism: Committee on educational interventions for children with autism. In Catherine Lord & J. P. McGee (Eds.), *Division of behavioral and social sciences and education*. Washington, DC: National Academy Press.

Newborg, J. (2005). *Battelle developmental inventory* (2nd ed.). Itasca, IL: Riverside.

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. doi:10.1007/s10803-009-0825-1.

Odom, S. L., Cox, A. W., Brock, M. E., & National Professional Development Center on Autism Spectrum Disorders. (2013). Implementation science, professional development, and autism spectrum disorders. *Exceptional Children, 79*, 233–251. doi:10.1177/001440291307900207.

Professional Development in Autism Center. (2008). *PDA program assessment*. Seattle, WA: PDA Center at University of Washington.

Ruble, L. A., McGrew, J. H., & Toland, M. D. (2014). Mechanisms of change in COMPASS consultation for students with autism. *Journal of Early Intervention, 35*(4), 378–396.

Sandall, S. R., Ashmun, J. W., Schwartz, I. S., Davis, C. A., Williams, P., & McBride, B. J. (2011). Differential response to a school-based program for children with ASD. *Topics in Early Childhood Special Education, 31*(3), 166–177. doi:10.1177/0271121411403166.

Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *The childhood autism rating scale*. Los Angeles: Western Psychological Services.

Smith, T. (2001). Discrete trial training in the treatment of autism. *Focus on Autism and Other Developmental Disabilities, 16*(2), 86–92. doi:10.1177/108835760101600204.

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland adaptive behavior scales (Vineland II)*. Bloomington, MN: Pearson.

StataCorp. (2009). *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP.

Strain, P. S., & Bovey, E. H, I. I. (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. doi:10.1177/0271121411408740.

U.S. Department of Education. (2003). *Identifying and implementing educational practices supported by rigorous evidence: A user friendly guide*. Washington, D.C: U.S. Department of Education.

Wilczynski, S. M. (2010). Evidence-based practice and autism spectrum disorders: The National Standards Project. *Communique: The Newspaper of the National Association of School Psychologists 38*(5), 24–25. Retrieved from http://www.naspponline.org.

Wilson, K. P., Dykstra, J. R., Watson, L. R., Boyd, B. A., & Crais, E. R. (2012). Coaching in early education classrooms serving children with autism: A pilot study. *Early Childhood Education Journal, 40*, 97–105. doi:10.1007/s10643-011-0493-6.

Wong, C. (2013a). *Naturalistic intervention (NI) fact sheet*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, The National Professional Development Center on Autism Spectrum Disorders.

Wong, C. (2013b). *Pivotal response training (PRT) fact sheet*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, The National Professional Development Center on Autism Spectrum Disorders.

Wong, C., Odom, S. L., Hume, K., Cox, A. W., Fettig, A., Kucharczyk, S., & Schultz, T. R. (2014). *Evidence-based practices for*

children, youth, and young adults with Autism Spectrum Disorder. Chapel Hill: The University of North Carolina.

Young, H., Rowan, J., Pardew, M., Sanford, T., Falco, R., & Reinmuth, J. (2009). *The autism partnership program: Parents & educators partnering to improve outcomes for children and youth with autism (APP)*. Unpublished manuscript, Autism Training & Research Center, Portland State University, Portland, OR.