Journal of Autism and Developmental Disorders
https://doi.org/10.1007/s10803-018-3740-5

ORIGINAL PAPER

Enhancing Low-Intensity Coaching in Parent Implemented Early Start Denver Model Intervention for Early Autism: A Randomized Comparison Treatment Trial

S. J. Rogers[1] · A. Estes[2] · L. Vismara[1,5] · J. Munson[2] · C. Zierhut[1] · J. Greenson[2] · G. Dawson[3] · M. Rocha[1] · C. Sugar[4] · D. Senturk[4] · F. Whelan[4] · M. Talbott[1]

Abstract
Short-term low intensity parent implemented intervention studies for toddlers with autism spectrum disorder (ASD) have found it difficult to demonstrate significantly improved developmental scores or autism severity compared to community treatment. We conducted a randomized comparative intent-to-treat study of a parent implemented intervention to (1) test the effects of an enhanced version on parent and child learning, and (2) evaluate the sensitivity to change of proximal versus distal measures of child behavior. We randomized 45 children with ASD, 12–30 months of age, into one of two versions of parent-implemented Early Start Denver Model (P-ESDM), the basic model, in which we delivered 1.5 h of clinic-based parent coaching weekly, and an enhanced version that contained three additions: motivational interviewing, multimodal learning tools, and a weekly 1.5-h home visit. We delivered the intervention for 12 weeks and measured child and parent change frequently in multiple settings. We found a time-by-group interaction: parents in the enhanced group demonstrated significantly greater gains in interaction skills than did parents in the non-enhanced group. Both interventions were associated with significant developmental acceleration; however, child outcomes did not differ by group. We found a significant relationship between degree of change in parental interaction skill and rate of children's improvement on our proximal measure. Parents in both groups reported satisfaction with the intervention. These findings suggest that parent skills improved more in the enhanced group than the comparison group. Children in the two groups showed similar improvements. Rate of individual parent learning was associated with greater individual child progress on a measure quite proximal to the treatment, though not on standardized assessments.

Keywords Early intervention · Parent-implemented intervention · Autism · Toddlers · ESDM

Parents of young children with autism spectrum disorder (ASD), toddlers under 36 months of age, often receive the diagnosis of ASD for their children without simultaneous entrance into empirically based treatments for their children. Start of treatment may be delayed by many factors, including the wide disparities in public delivery of early intervention that exist across states, separation of diagnosis and intervention in systems of care, waiting lists, time delays while third party payments are being worked out, time delays while parents research various options, and lack of public system provision of efficacious treatments for children under 36 months of age. Although involving parents in developing goals and embedding learning opportunities into daily routines is considered best practice in the field of early intervention (McWilliam 2010; Hanft et al. 2004), efficacious interventions for ASD were largely developed for somewhat older children and designed to be delivered by therapists and may inadvertently leave parents relatively uninvolved. This is unfortunate, because providing parents with empirically-validated interventions they can embed within their ongoing daily routines with toddlers provides them with

✉ S. J. Rogers
sjrogers@ucdavis.edu

1 University of California Davis MIND Institute, Sacramento, CA, USA

2 University of Washington, Seattle, WA, USA

3 Duke University, Durham, NC, USA

4 University of California, Los Angeles, Los Angeles, CA, USA

5 Present Address: Emory University, Atlanta, GA, USA

Published online: 10 September 2018

both (1) immediate access to effective intervention, and (2) activities they themselves can do as part of their parenting routines that meet their desire to parent and help their child. Providing parents with effective techniques supports their experience of competence as parents and confidence in their children as learners (Estes et al. 2014) and may buffer them from feelings of inadequacy from dependence on others to help their children.

Additionally, and of importance for public systems' delivery of care, parent-implemented interventions involve much lower costs than do intensive therapist-delivered interventions, and may still provide the opportunity for children to receive a high intensity of learning opportunities from daily activities with parents. Currently, therapist-delivered, comprehensive, high-intensity interventions are considered the most efficacious interventions for early autism and, they have provided evidence that significant gains in child IQ, child expressive and receptive language, changes in adaptive behavior, and decreases in autism symptoms can be achieved (McEachin et al. 1993; Dawson et al. 2010). Because IQ, autism severity, and spoken language ability in the preschool period are the best predictors of adult outcomes (Magiati et al. 2014), improving developmental, language, and social functioning of young children with autism are arguably crucial outcomes for defining an efficacious comprehensive early intervention approach. Targeted treatments, on the other hand, designed to improve specific behaviors rather than overall functioning, often do not measure effects on these types of outcomes (e.g. Bradshaw et al. 2017; Brian et al. 2016), although there are certainly exceptions (notably the randomized controlled trial of Pivotal Response Training: Hardan et al. 2015; Gengoux et al. 2015).

In an effort to determine efficacy of parent-implemented interventions for toddler outcomes, many studies have been carried out over the past decade. The most robust evidence of efficacy from studies of children under three involves changes in parent–child interactions involving parental synchrony and shared attention. This is theoretically important because of the relationship between these skills and later language acquisition or IQ gain in children with ASD demonstrated in previous studies (Siller and Sigman 2002). Several recent RCTs have demonstrated efficacious effects for improving parent interactions with their toddlers with ASD as well as targeted child behaviors, such as joint attention (Schertz et al. 2013; Kasari et al. 2010, 2014, 2015). In a few cases, studies have demonstrated direct links between changes in parent behavior and child change (Shire et al. 2016). However, very few studies report the kinds of gains in developmental skills and expressive language reported by the intensive therapist-delivered interventions. Wetherby's (2014) study stands alone in reporting effects of a parent-implemented intervention on standardized measures involving language, adaptive behavior, and social communication. However, this study documented parent delivery of a relatively intense amount of intervention for a longer time period (i.e. 20 h a week spent delivering the intervention) than other studies to date.

In an effort to identify possible reasons for the lack of developmental changes demonstrated in most studies of parent-delivered intervention (Rogers et al. 2012a), we considered three possible problems that could be examined within such a study. The first reason involves treatment dosage, in that parents need time to learn the interventions before they can be delivering them at the ended "dosage". Parents are learning at the same time that children are learning, and until parents reach and sustain fidelity of implementation, children are not receiving full dosage of the intervention. Thus, the effects of the changes in parent–child interaction that are being accomplished via parent coaching may take longer to affect child developmental and symptom changes than can be measured in short-term studies.

The second reason involves a measurement sensitivity. When standardized developmental, language, and autism diagnostic measures are used pre and post treatment to assess short-term change, sensitivity of such measures to detect smaller, specific changes in child function may be limited by the wide range of functioning of the children, the individual patterns of behavior over time, and the need for child generalization of newly learned skills from home routines to the standardized assessment content and context, quite distal from child learning context. Standardized test items are neither proximal to the treatment nor necessarily sensitive to small changes in social communication skills that may be occurring as treatment progresses, and detection of change may be enhanced by using frequent observational measures of social communication during treatment (MacDonald et al. 2014). An additional measurement issue involves use of parent-based measures as primary outcomes because parents are not blind to assignment group.

The third complication involves the lack of experimental control of treatment exposure that occurs when the comparison group receives community treatment, given that early intervention is mandated for all young children with ASD.

The current study carried out a randomized comparative intent-to-treat study of a parent implemented short-term intervention for toddlers with ASD designed to address these methodological concerns in several ways: (1) by randomly assigning both groups to a specific intervention rather than using a treatment-as-usual group; (2) by using a primary child

outcome measure that was proximal to the intervention content and administered by naïve assessors; and (3) by increasing dosage of the intervention in one treatment groups by increasing amount of parent coaching. We did this by adding three enhancements: an additional weekly parent coaching session at home (to increase ease of generalization to everyday home activities), use of multi-modal intervention materials with parents to accommodate individual learning styles, and the use of motivational assessment and interviewing techniques to enhance parental motivation. The current study tests whether, in combination, these design enhancements significantly enhanced parent fidelity of implementation and child progress. It also examines parent satisfaction with these interventions, given possible concerns that parent-delivered interventions add stress to parents' lives.

Primary Hypothesis There will be a time-by-group interaction such that parents in the enhanced P-ESDM group will demonstrate greater sensitivity and skill in supporting their children's social communicative development as measured by the increase in fidelity of implementation skills after 12 weeks of intervention than will parents in the control group: the standard P-ESDM group.

Secondary Hypothesis 1 There will be a significant positive relationship between improvements in parent fidelity of implementation skills and child social-communication change scores on the ESDM curriculum checklist.

Secondary Hypothesis 2 There will be a time-by-group interaction such that children in the enhanced P-ESDM group will demonstrate greater gains on a proximal measure of social-communication, the revised ESDM checklist after 12 weeks of intervention, than will children in the control group.

Ancillary Hypothesis 1 There will be a time-by-group interaction such that children in the enhanced P-ESDM group will demonstrate greater gains on scores from MSEL, ADOS, and Vineland social and communication scores, and on associated symptoms (sleep, eating, tantrums, self-injury) reported on the CBCL, than will children in the control group.

Ancillary Hypothesis 2 Parents will rate both treatments as satisfactory but will find twice a week intervention less satisfactory than once a week intervention, given its demands on their time.

This report focusses on these five hypotheses. Analyses of other measures and hypotheses will be reported elsewhere.

Method

IRB Oversight

All aspects of the study were conducted under IRB approval and parental consent at both universities. No adverse effects of the study occurred at either site.

Study Design

The trial was conducted at two universities: University of Washington and University of California Davis MIND Institute, by teams with extensive experience conducting multi-site studies of the ESDM treatment model. Families with a child with a diagnosis or risk features of ASD aged 12–30 months were recruited via website announcements and fliers to community pediatric care and service sites, given a standard telephone interview to screen for inclusion/exclusion criteria, and administered two ASD screeners. Those children who met all three screening criteria and consented then received an eligibility assessment. Three more weekly visits were then carried out for eligible children: (1) a clinic visit to randomize to groups and administer two measures: the PATH curriculum checklist and video-recorded parent–child interaction, (2) a home visit to provide a video camera, gather another parent–child interaction video, complete a home environmental survey, and discuss home routines, (3) and a clinic visit to administer a second baseline PATH curriculum checklist and a third baseline parent–child interaction video. 12 weeks of intervention were then delivered. Child performance was measured twice more, once at the end of the 12 weeks of treatment, and at a third point 12 weeks later, originally planned as a maintenance period. (There were no limitations on family choices regarding outside treatment during any phase). However, design problems in the management of the maintenance period (different maintenance regimens administered to the two groups) ruled out the use of the third data collection point for group comparisons.

Sample Size

To determine the necessary sample size for the study we conducted a power analysis on a sample of 48 children (24 from UCD and 24 from UW) from a previous parent coaching study, taking into account the shared variance among the primary developmental measures: MSEL DQ and Vineland. Assuming a correlation of .25 across measures between the first and second measurement (based on data from Dawson et al. 2010) and solving for the effect size that a given sample size has .80 power to detect, determined that we had .80

power to detect a small to medium effect size of f^2 = .06 (Cohen's small and medium effect sizes are f^2 = .02 and 0.15, respectively) on the Vineland and MSEL. Assuming a higher correlation of .5 across measures and solving for the effect size that a given sample size has .80 power to detect determined that we had .80 power to detect a small to medium effect size of f^2 = .04. Thus, a sample of 48 participants was determined to have sufficient power to detect true differences with effect sizes of only .06, and we randomized 45 children and families into the study.

Randomization and Allocation Concealment Strategies

The participants were randomized by a statistician in an independent data center, using a dynamic allocation procedure based on minimization of the variance of the treatment effect in the linear model relating the outcome to three stratification variables: age (cutoff 20 months), Mullen developmental quotient (DQ) score (cutoff 60) and gender (Begg and Iglewicz 1980). Stratification cutoffs were used to balance cells with respect to age, DQ, and gender.

After the eligibility assessment was completed and consent was obtained, the project coordinator at a site entered the child's age, DQ score and gender into the Data Coordinating Center (DCC) online data system and then notified the DCC statistician to request a treatment assignment. The statistician logged into the system, completed the randomization and notified the coordinator of the assignment, who then notified the parents.

All assessors and data coders were kept unaware of treatment allocation; however treatment assignment could not be masked from families, children, and therapists. Strict physical separation was kept between assessment and treatment data and between assessors and therapists.

Inclusion/Exclusion Criteria

We excluded children with: (1) any identifiable genetic condition associated with autism or intellectual disability (2) neurological disease or injury (e.g., epilepsy) (3) significant sensory or motor impairment (e.g., cerebral palsy), (4) birth weight < 2500 g and/or gestational age < 36 weeks, (5) prenatal exposure to neurotoxins (including alcohol, drugs), (6) current substance abuse, bipolar disorder, or psychosis in caretaking parent, (7) home located greater than a specified distance from the clinic, (8) English not read fluently and spoken in the home on a daily basis; (9) previous ESDM treatment or 8 h or more weekly of 1:1 autism treatment; (10) DQ below 35 and (11) not yet walking due to requirements of the autism assessment measure, the Autism Diagnostic Observation Schedule for Toddlers (ADOS-T; Lord et al. 2009). Children were not excluded once they had been assigned to a group regardless of a change in their criteria (i.e. number of treatment hours) due to the ITT design. Some children were excluded at the time of screening due existing exclusionary criteria (e.g. gestational age) while other were excluded after the eligibility assessment because of resulting exclusionary test scores resulting from the assessment.

We included children between 12 and 30 months at enrollment who met full criteria for ASD both by ADOS-T cutoff scores and by two independent clinicians' clinical judgment, and whose parents agreed to a weekly home visit and clinic visit, and who met no exclusion characteristics.

Participant Flow

The consort table, shown in Fig. 1, illustrates the flow of children through the study. As shown, 72 children were screened, 34 at UCD and 38 at UW. Of those, 18 did not proceed further, 8 because they did not screen positive or meet inclusion criteria (4 at UCD, 4 at UW) and 5 because the parents chose to discontinue (1 at UCD, 4 at UW). Of the 63 who proceeded to be assessed for eligibility, 10 did not meet inclusion criteria (4 at UCD, 6 at UW) and 8 declined to continue (3 at UCD, 5 at UW). Of the 45 that were included and randomized (26 at UCD and 19 at UW), 24 were allocated to P-ESDM (15 at UCD, 9 at UW) and 21 to P-ESDM++ (11 at UCD, 10 at UW). 11 of the 24 allocated to P-ESDM declined 25% or more of the intervention hours, all at UCD, and 10 of these declined or were lost to follow-up assessment (all at UCD). In the P-ESDM++ group, 7 declined 25% or more of their scheduled intervention hours (6 at UCD, 1 at UW), and 3 of these declined or were lost to follow-up assessment 2 at UCD, 1 at UW. Thus, there was differential dropout with respect to treatment group; those assigned to P-ESDM rather than P-ESDM++ had higher dropout rates at UCD, and most of it happened in the first 30 days. This was primarily due to family choices related to the greater availability of intensive treatment services via public services to families in the UCD than in the UW site. Given the intent-to-treat longitudinal design, all children were included in all assessments and all collected data from all three time points were used in the analyses regardless of the amount of treatment they received and their course in the study once enrolled.

Management of Attrition

Because dropout occurred early in study and these parents chose not return for later assessments, little longitudinal information was available to compare outcome trajectories of dropout versus completing participants. Since most of the dropout was at UCD, we studied dropout effects by including two and three-way interactions between site, time in treatment, and treatment group in the models for the analyses of

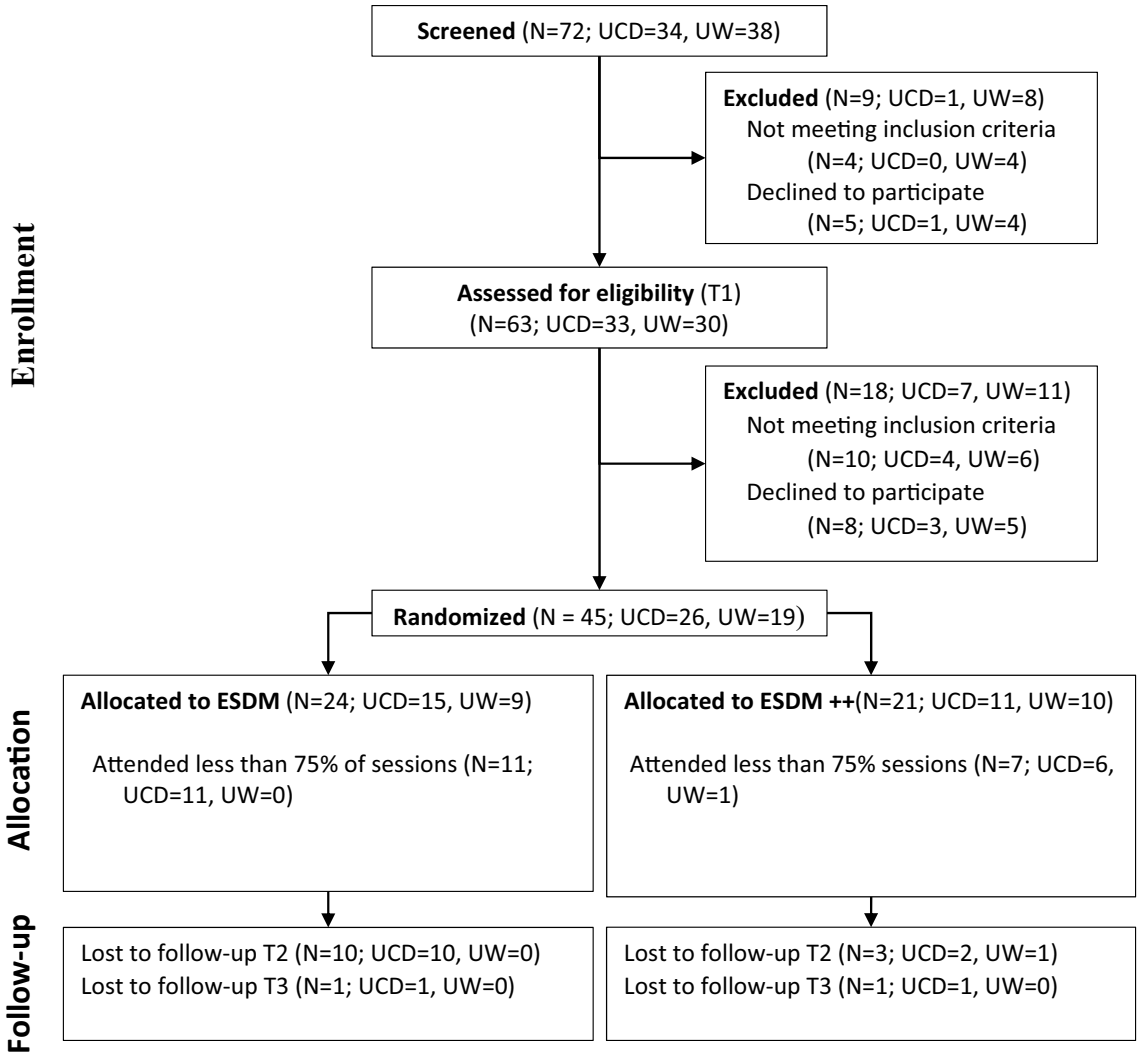

Fig. 1 Consort table

the primary and secondary hypotheses. There were no significant three-way interactions between site, time in treatment, and group. Hence none of the group differences reported in the "Results" section was due to differential dropout at the sites and the final models did not include site. Note however that given the amount of attrition that occurred, the analyses are based on a much smaller sample than was determined to be needed by the power analysis. Nevertheless, the design still had ample power on the two primary measures of parent fidelity and curriculum checklist due to significantly more repetitions obtained on these measures throughout the study (at eight time points).

Initial Equivalency of Groups

A variety of baseline demographic variables involving both child and parent were compared between treatment groups to assess the success of randomization using T tests and chi squares, shown in Table 1. There were no significant differences in the distribution of gender, race, Hispanic ethnicity, chronological age, mother's education and outside intervention hours between the treatment groups. Effect size in Cohen's d for group differences for outside intervention hours was .44.

Implementation of Interventions

Overview

Two versions of the parent-implemented Early Start Denver Model (P-ESDM; Rogers et al. 2012a) were used for this study. One group received the P-ESDM as it was described in Rogers et al. (2012a) and the parent treatment manual, *An*

Table 1 Baseline description of subjects and family characteristics: child age, ethnicity, race, parent age, maternal education, treatment hours received outside of the study

	All mean ± sd	P-ESDM N = 24	P-ESDM++ N = 21	χ^2 or T	p Value
Child age years	2.1 ± 0.4	2.1 ± 0.4	2.1 ± 0.4	0.34	.73
Child hours of outside treatment	4.0 ± 6.6	2.7 ± 3.3	5.5 ± 8.7	1.34	.17
Mother educ					
≤ HS	4	3	1		
BS	23	11	12		
Grad sch	11	4	7	1.8	.4
No data	6				
Ethnicity					
Hispanic	4	3	1	1.5	.47
Not	24	11	13		
No data	17	10	7		
Race					
Black	3	0	3	1.8	.77
Asian	2	1	1		
Multiple	7	3	4		
White	20	12	8		
No data	15	8	7		
Gender					
Male		16	15		
Female		7	6	0.02	.9

Early Start for your Child with Autism (Rogers et al. 2012a), which was given to all parents. This group—the P-ESDM group—received 12 consecutive 1.5 h sessions in weekly clinic visits. The second group—the enhanced P-ESDM (P-ESDM++) group—received the same treatment but with three enhancements added: (1) a second 1.5- h weekly home session; (2) materials provided in multiple modalities according to parent preferences (e.g. paper materials as requested including sketches, worksheets, cue cards and a website containing video examples, narrated and print condensed text for each topic); (3) and use of motivational interviewing (MI) techniques (Miller and Rollnick 2012).

Therapists received training in MI techniques through continuing education activities online and in real time and meet weekly for peer supervision. Therapists began MI dialogues in P-ESDM++ sessions in which parents reported motivational problems on written and/or verbal reports of current motivational state, assessed weekly via administration of a brief (seven questions, 1 min) self-assessment questionnaire modeled from Prochaska et al. (1997) to identify motivational stage. Therapists used these parent reports to focus their dialogue at the start and end of the sessions and to tailor their focus in terms of opening and closing dialogues, goals for the session, and review versus teaching new skills.

Data were examined from 201 weekly parent questionnaires gathered from 21 of the 22 families. Two families provided only one or fewer data points. Across the sessions, 19 parents reported being actively engaged in learning and using with positive motivation for learning, either in (1) preparation, ready to learn and working with the therapist on preparing by setting up their house, learning from the therapist, manual, and other materials, (2) action phase of active learning and practice, for whom we focused on coaching them in the P-ESDM techniques as per the manual, or (3) maintenance, in which they reported that their new skills now felt natural and were embedded into everyday life. Since we collected data weekly there was often movement from one of these stages to the next. For them, sessions focused on continuing the use and refining techniques while using reflection and self-evaluation for review, generalization, and problem solving. Since we collected data weekly there was often movement from one of these stages to the next over time. Of these 19 parents, 12 reported no slippage or lapse in motivation across the 12 weeks. Seven reporting some lapses in motivation at some point during the 12 weeks. Only one family ended the intervention in ongoing motivational lapse. For families with lapsing motivation, the therapist listened carefully to parent reports of motivational struggles during the week, used MI techniques to ascertain which stage they were now describing, and worked with them to establish a goal for their interactions with their child for the coming week that helped them move a step closer to the motivation needed for preparation, action, and eventually maintenance.

Treatment Sessions

P-ESDM intervention sessions were carried out weekly in the clinics. P-ESDM++ interventions were carried out twice per week, once in the clinic and once in children's homes. Both treatments conformed fully to the detailed parent training manual (Rogers et al. 2012a), curriculum, and parent fidelity of implementation measure (Rogers et al. 2012a; Rogers and Dawson 2010), a specific coaching intervention method specified in a parent coaching manual and defined by a fidelity of implementation measure for therapists (Rogers and Vismara 2013), and 12–15 individualized written treatment objectives developed for each child by their therapist from their initial assessments, on which progress data were gathered during each session.

The intervention was delivered for 12 weeks and the data were analyzed at the end of treatment. For both groups, the first treatment session was devoted to developing children's learning objectives. In sessions occurring in weeks 2–10 parents were coached in each of ten intervention topics from the parent treatment manual, a copy of which was given to

each family. In each session one new topic was covered and topics taught earlier were reviewed. The ten topics were: (a) increasing child's attention and motivation; (b) using sensory social routines; (c) promoting dyadic engagement and joint activity routines; (d) enhancing nonverbal communication; (e) building imitation skills; (f) facilitating joint attention; (g) promoting speech development; (h) using antecedent-behavior-consequence relationships ("ABC's of learning"); (i) employing prompting, shaping, and fading techniques; and (j) conducting functional assessment of behavior to develop new interventions. Weeks 11–12 involved review of learned skills.

Sessions 2–12 followed a specific structure, beginning with 5 min of brief chat about the week and the parent focus, and then proceeding to a 5-min warm-up: a parent-chosen parent–child activity based on the parent focus of the past week. In the next 10 min, therapists introduced a new topic each week through verbal description and written materials from the manual and from previous observations of child and parent. In the next 5 min the parent practiced the technique in an activity with the child while the therapist provided coaching, encouragement, and feedback on technique use. A 5–10-min parent–therapist reflection followed. Two more activities occurred following the same format of parent–child activity with coaching followed by reflection. Activities varied to represent the typical interaction routines in a young child's life (books, feeding, dressing or changing, water activities, toy play, social play). Each session concluded with a discussion involving generalization of the new topic to home and community, as well as any other topics the parent wanted to cover. The parent left with abbreviated written materials on the new topic and a plan for practicing the new topic in various activities across the week.

Each session was videotaped for clinical supervision and fidelity coding of parent and therapist. All parents videotaped themselves monthly at home carrying out play routines with and without toys as per a written set of instructions. The project supplied tiny cameras that were worn on the body. Parent recordings were uploaded and viewed when they came to clinic sessions. Any un-codable videos resulted in a request to the family to send another video.

Intervention Staff

The treatments were administered by highly experienced, ESDM-certified professional staff/faculty at both sites: two psychologists, a speech/language therapist, three behavior analysts, and a family therapist. All had been providing ESDM parent coaching as well as direct ESDM and other treatments for several years before this study began. All completed the required P-ESDM training programs as well as web-based and peer-based training on the enhancements involved in the P-ESDM++ group. All met fidelity of implementation standards (85% or greater on multiple consecutive measurements) on both the child treatment measure and parent coaching measure before the study began and maintained it throughout the study. Staff at both sites video conferenced monthly to review videos, code and check agreement, provide peer supervision, and guard against site drift. All treatment was supervised by two authors, SR and LV, who were instrumental in developing the P-ESDM approach. Therapist fidelity of implementation was examined repeatedly across the trial and any slippage resulted in immediate retraining and re-assessment.

Assessment Staff

All child and parent measures were administered and coded by highly trained professionals who had completed a final degree in their fields, were experienced with assessing young children with ASD, and who met and maintained assessment fidelity of implementation measures and inter-rater reliability standards. The assessment team at each site was supervised by a senior team member. The assessment teams from both sites met monthly to review and code videos, provide peer supervision, and ensure consistency across and within sites.

Measurement

Practices

Following authentic assessment principles (Bagnato et al. 2010), we frequently collected data from multiple individuals and from multiple settings on child and parent mastery of skills being taught as well as using distal measures of child change. Every 4 weeks, naïve examiners coded videos of (1) parent mastery of treatment skills from parent–child videos collected at home and in the clinic, and of (2) child performance on the social communication skills being targeted by the parent, administered by naive assessors using the revised ESDM curriculum tool. Finally, we assessed social validity of the intervention at the end of treatment using a parent satisfaction measure and a parent-therapist working alliance scale.

Screening Measures

Each child received two of three telephone screening measures according to age.

Early Screening of Autistic Traits Questionnaire (*ESAT*) (Swinkels et al. 2006) is an autism risk parent interview previously tested in several studies, including a population-based study of 31,000 14–15 month-old children. This was used for children 12–15 months.

Infant Toddler Checklist (*ITC*) (Wetherby and Prizant 2002) is a broadband parent questionnaire developed to determine risk for communication disorders including ASD. The ITC is a standardized tool that has both screening cutoffs and standard scores at monthly intervals from 6 to 24 months based on a normative sample of over 2188 children. This screener was used with all children in the study.

The Modified Checklist for Autism in Toddlers (*M-CHAT*) (Robins et al. 2001) is a 23-item parent questionnaire developed to screen for autism symptoms in children ages 16–30 months. The M-CHAT has been validated in two large validation studies. This was used for children 16–24 months.

Primary Outcome Measures

Two primary outcome measures were gathered monthly, one to capture parent change, and one to capture child change.

Parent change was examined using the *ESDM Fidelity Rating System* (Rogers and Dawson 2010), a Likert-based, 5-point rating system of 13 adult behaviors. Parent fidelity was rated monthly from videotapes recorded by parents in the home and from videotapes recorded by study staff in the clinic. Because the focus of the intervention is on incorporating ESDM strategies into all types of interactions with their children throughout typical daily routines, we asked parents to film themselves playing with their child as they typically would, in order for us to determine what techniques had generalized into their ordinary interactions. We asked them to carry out one activity with toys and one activity without toys. In the clinic, play utilized different materials than were used in treatment sessions. In the home, play materials were familiar to the child.

Videotapes were reviewed and activities to be coded were selected by research staff naïve to treatment assignment according to strict definitions about what behaviors began an activity, ended an activity, and defined an activity and the required length of an activity (2–4 min). The unit of analysis was the activity itself, so codes were not influenced by the length of an activity. The selected activities were edited and uploaded to a protected web-based file, assigned a dummy ID, and coded by trained raters (ESDM certified therapists) naïve to subjects and group assignments. Raters established agreement with an expert coder before beginning to code and 10% of videos were rated for agreement. Total scores range from 14 to 60, with higher scores reflecting greater use of ESDM practices. Inter-rater reliability for parent fidelity was ICC of 0.47, which reveals only fair agreement, based on 80 video clips from 29 children.

Child change was examined using the *PATH Curriculum Checklist* (PATH CC; Rogers et al. 2013). This tool, rather than a standardized developmental assessment tool, was chosen as the primary outcome measure because it was proximal rather than distal to the actual skills that children were being taught. Additionally, the play-based administration method was inherently more variable than a standardized assessment and thus made it less likely that children would learn the test over the course of the frequent assessments that were conducted. The PATH CC is a criterion-based measure of child development derived from the Early Start Denver Model curriculum checklist (Rogers and Dawson 2010). The PATH CC consists of 136 items organized in nine developmental domains most affected by ASD in early development: receptive understanding of gestures and words, expressive use of gestures and words, joint attention, social interaction with adults, imitation, cognition, and play skills, both functional and symbolic. The items span the developmental range from 8 to 30 months.

The PATH CC was administered by certified ESDM therapist who was not otherwise involved in the study and was thus naïve to group, child and parent. The assessment was carried out in a 90-min timeslot using standard set of materials. Parents were present throughout and joined in as needed to help the child feel comfortable. Parent report was used to corroborate examiner observations. The assessor used child-preferred play materials selected from the standard set to engage the child in appropriate play with materials and with social games and scored each CC item based on the child's consistent demonstration of each skill in two or more activities. Examiners established a basal level of three consecutive passes in each domain and continued to probe and score items until the child failed three consecutive items in a domain. Item descriptions specified what was required to pass each item. The session was videotaped to aid with scoring. Inter-rater reliability was scored from the videos by a trained rater naïve to all aspects of the study. The overall scores were derived as follows. Because each child was only scored on items that fell within the child's own basal and ceiling levels, children received differing numbers of items. In order to account for this, we summed the number of passed items and added to this the number of items that fell below the child's basal score for each domain, and then summed these for all domains, which yielded the total number of passes an individual child performed on the CC. This sum was divided by the total number of items on the scale minus any skipped items, yielding a percentage pass score for each child.

Psychometric analysis was based on 262 observations gathered from a total of 40 children. Inter-rater reliability from ten children for the total percent administered summary score yielded ICC of .92. Individual subscale ICCs were as follows: receptive communication/gestures = .52, receptive communication speech = .65, expressive communication gestures = .71, expressive communication vocal/verbal = .93, imitation = .49, joint attention = .42, play = .78, social skills dyadic = .82, cognition = .93. Test–retest reliability based on 213 samples of the total scores was calculated using

Spearman's r which yielded .90. Two samples were generated for this analysis. Sample 1 consisted of all observations in session 1–7, which was considered as time 1. Sample 2 consisted of all observations in session 2–8, which was considered as time 2.

Construct and concurrent validity was examined by testing correlations between the PATH CC, the MSEL, and the ADOS 2 social affect score. Spearman's r between PATH CC total percent administered score and MSEL total was $r = .90$, ($p < .001$) and domain scores were $r = .82$, .86, and .83 for expressive language, receptive language, visual reception, respectively ($p < .0001$ for all). Given that the PATH CC (like the ESDM CC) was built to reflect a profile specific to early autism involving specific difficulties with social communication and play, we also examined the correspondence between the children's PATH CC scores and the ADOS 2 social affect scores. There were significant negative correlations ($p < .05$) between all PATH CC domains and the ADOS 2 social affect scores, ranging from $-.48$ for Path CC cognition skills to $-.67$ for Path CC expressive communication/speech. Thus, increases in PATH CC scores were significantly related to decreases in ADOS 2 social affect scores and increases in child MSEL scores.

Secondary Outcome Measures

The Autism Diagnostic Observation Scale for Toddlers (ADOS-T; Lord et al. 2012) was used at the eligibility and both post-treatment assessments for all children, as were all other secondary child outcome measures. The ADOS-T uses a semi-structured observational assessment that provides a number of opportunities for interaction (e.g., play, turn-taking games, looking at books, etc.) and measures social and communicative behaviors, as well as repetitive behaviors diagnostic of autism on a scale of 0–3. The full range of scores was used, including 3's, to maximize the range of each item. Subtotals were computed separately for the two domains, social affect and restricted, repetitive behaviors. Severity scores were computed according to procedures described by Gotham et al. (2009). Lab personnel were trained to 85% reliability on the full range of scores. Inter-observer reliability at individual sites was assessed on at least 20% of interviews, and any deviation from standard led to retraining.

Mullen Scales of Early Learning (Mullen 1995), is a standardized, normed developmental assessment for children aged birth through 68 months. A more detailed set of administration instructions and training materials were created to train study staff and reinforcers for child attention and cooperation were used as needed to support children's motivation and cooperation. Because of floor effects using the standardized t scores, we created three DQ scores to use in all analyses from the subscale developmental age (DA) equivalents ($DQ = DA/CA \times 100$) as measures of intervention efficacy: a verbal DQ comprised of receptive and expressive language scales; a nonverbal DQ comprised of the visual reception and fine motor scales and an overall DQ created from these four subscales. 20% of the assessments were scored by two raters for administration fidelity.

Vineland Adaptive Behavior Scales, Second Edition (*VABS II*), (Sparrow et al. 2005) examines four domains of adaptive behavior: communication, self-care, social, and motor skills. The VABS-II has excellent psychometric properties. It was administered to the primary caretaker by telephone at both assessment points. Age equivalent scores and the DQs for the four domains were used to assess treatment outcome.

Child Behavior Checklist for 1 ½–5 Years (CBCL: Achenbach and Rescorla 2000). The CBCL is a norm-referenced questionnaire that obtains parent ratings of 99 behavior items and descriptions of parents' concerns, descriptions of problems, and strengths of the child. It provides factor analytic derived profiles of internalizing, externalizing, and total problems scales and a stress problems scale. Instrument derived scores were used for each domain.

Child Intervention History (*Version 6-10-13*), adapted from the CPEA Network Intervention History form (Dawson et al. 2003, unpublished), see (Rogers et al. 2012b for a full description). We gathered information on every hour of every treatment children received from start to finish of the study. Intervention types included: ABA, OT, PT, speech/language, TEACCH, DIR, play therapy, and Part C infant-toddler services, excluding P-ESDM hours. The tool was administered twice: at enrollment and at the end of the treatment phase, by telephone interview. We calculated the total number of treatment hours as described in Rogers et al. (2012a). As shown in Table 1, there was no significant difference in the amount of average treatment hours received by children in the two groups [P-ESDM mean = 2.7 (SD = 3.3), P-ESDM++ mean = 5.5 (8.7), $p = .17$] during the span of the study.

Parent Satisfaction Survey

The Intervention Evaluation Form for Parents (University of Washington, unpublished) is a Likert-based scale of 14 questions administered at the end of the treatment period that asked parents to rate their perception of the utility of the intervention for promoting child change in language, rate of learning, and overall progress, its ease of use, the quality of relationships, responsivity, professionalism, respects, warmth, and competence of staff members, their

overall satisfaction with the intervention, as well as barriers to participating in the study.

Statistical Methods Used

T-tests and Chi square tests were used for baseline comparisons between treatment groups (to assess the success of randomization). Generalized linear mixed models (GLMM) with main effects of treatment (standard ESDM versus enhanced treatment) and time (baseline, end of treatment at 90 days), treatment-by-time interactions and subject level random intercepts and slopes were used to model the longitudinal trajectories of the outcomes. The primary measures: the PATH CC and parent fidelity of implementation were measured more frequently than the secondary measures, eight times for the CC and four times each for social play and toy play which were combined and analyzed providing eight repeated measurements, until the end of the study. Time was modeled such that the rate of improvement (slope) over the treatment phase (baseline to end of treatment) was allowed to differ from that over the follow-up phase (end of treatment to the 3-month follow-up, broken line model). GLMM account for correlations between repeated measures within subjects and easily allow for different measurement time points across subjects and variables. Hence, all available observations from each subject were utilized in modeling via the GLMM. For the secondary analysis involving the hypothesized association between change during treatment in the primary parent and child outcomes (parent fidelity and PATH CC scores), we regressed the predicted subject-specific PATH CC slope on the predicted subject-specific parent fidelity slopes. The subject-specific slopes were predicted using the GLMM described above in modeling the two main parent and child outcomes, separately. Direct measures of change scores or slopes were not available due to different measurement times across subjects.

Results

Outcomes and Estimation

Primary Hypothesis: Change on Parent Fidelity Measure

The goal for parent achievement of fidelity is 80% on a 5 point Likert based scale, or an average score of 4.0 when measured in the lab setting. Fidelity scores recorded from videos of social play and toy play were analyzed as a single outcome, since scores from both activities were highly correlated ($r = .56$, p value $= .0001$). Additional covariates of mother's education level, age of subject at start of treatment, baseline ADOS score, baseline Mullen DQ, average hours of outside intervention/services, along with their interaction with the time were added to the main GLMM in modeling fidelity but were not found significant and thus were omitted from the final models. There was a significant interaction effect between the treatment group and time ($F(1,166) = 7.90$, $p = .0056$) with the P-ESDM++ group exhibiting greater improvement (estimated mean baseline: 3.40 and estimated mean at end of treatment: 3.80) than the standard P-ESDM group (estimated mean baseline: 3.39 and estimated mean at end of treatment: 3.18, a nonsignificant change).

Secondary Hypothesis: Change on PATH CC Score

Both the P-ESDM (estimated mean baseline: 0.37 and estimated mean at end of treatment: 0.50) and the P-ESDM++ (estimated mean baseline: 0.35 and estimated mean at end of treatment: 0.51) groups combined showed improvements in PATH CC score ($F(1,218) = 145.9$; $p < .0001$) during the treatment period. The group by time interaction was not found significant ($F(1,216) = 2.02$; $p = .157$). Among the additional covariates of mother's education level, age of subject at start of treatment, baseline ADOS 2 score, baseline MSEL IQ, average hours of outside intervention/services, only baseline ADOS 2 ($\beta = -.056 \pm .012$, $F = 21.44$, $p < .0001$) and Mullen IQ ($\beta = .0062 \pm .0011$, $F = 32.60$, $p < .0001$) were found to be significant. Since both baseline autism severity and intellectual delay contribute to child learning potential and are highly correlated, it may be that overall severity of child impairments is limiting progress. This finding may also demonstrate that PATH-CC scores are reflecting both child development and autism symptoms.

Secondary Hypothesis: Relationship Between Parent and Child Change

The predicted subject-specific PATH CC slopes were regressed on the predicted subject-specific parent fidelity slopes as obtained from the respective GLMMs described above. A significant association was detected ($\beta = .057 \pm .007$, $F = 73.24$, $p < .0001$), corresponding to an additional 2.6% change in the checklist scores by the end of treatment for every 0.45 point change in the fidelity scores (on a 1–5 point scale) at the end of treatment. This significant association was largely due to the fact that improving parent fidelity scores (seen only in the P-ESDM++ group) corresponded to larger improvements in the PATH CC scores. Conversely, lack of improvement in parent fidelity scores (seen in the P-ESDM group) corresponded to smaller improvements in child checklist scores. There was not an association between parent and child improvement *within* either the P-ESDM or the P-ESDM++ groups (see Fig. 2). These findings demonstrate a threshold, rather than a continuous, effect of parent fidelity improvement on child

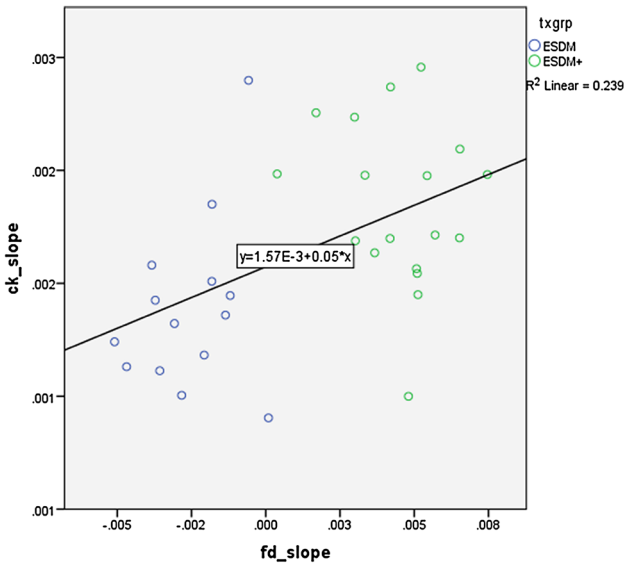

Fig. 2 Relationship of rate of parent gain estimates on fidelity scores to rate of child gain estimates on PATH curriculum checklist

developmental gains. The children of parents who showed improvement over time on fidelity scores showed greater developmental gain (i.e., improvements on PATH CC scores) over time than did children whose parents did not show fidelity improvement over time. Note that all data in Fig. 2 are based on slopes for parents and children estimated from the GLMM modelling of these outcomes rather than the group means reported in Table 2.

Ancillary Hypothesis 1: Modeling Change on ADOS, Mullen and Vineland Scores

Both the P-ESDM and P-ESDM++ groups combined showed improvement in MSEL DQ score ($F(1,57) = 31.55$; $p < .0001$) during the treatment period, where there were improvements in both P-ESDM++ group (estimated mean baseline: 67.9 and estimated mean at the end of treatment: 76.2) and the standard P-ESDM group (estimated mean baseline: 69.8 and estimated mean at end of treatment: 76.1). Group by time interaction was not found to be significant ($F(1,55) = .53$; $p = .47$).

The P-ESDM and P-ESDM++ groups did not show significant change across time in the Vineland Adaptive Scale ($F(1,56) = 1.64$; $p = .2$). There were no improvements in either group (P-ESDM++; estimated mean baseline: 80.6 and estimated mean at the end of treatment: 80.9; P-ESDM estimated mean baseline: 79.5 and estimated mean at end of treatment: 82.3). The group-by-time interaction term was not statistically significant $F(1,54) = 1.12$; $p = .29$.

Both the P-ESDM and P-ESDM++ groups combined showed a significant decrease in ADOS from baseline to end of treatment ($F(1,60) = 19.62$; $p < .0001$): P-ESDM++ group (estimated mean baseline 7.7 and estimated mean at the end of treatment 6.5); P-ESDM group (estimated mean baseline 8.0 and estimated mean at end of treatment 6.7). Group by time interaction was not significant ($F(1,58) = .05$; $p = .82$).

Ancillary Hypothesis 2: Parent Satisfaction

Twenty-seven of 30 parents completed the Parent Satisfaction Survey. Parents were highly satisfied, with 50%

Table 2 Baseline and end of treatment mean scores and standard errors of the mean based on model estimates for all independent and dependent variables

	Txgrp 1 ESDM n = 24	Txgrp2 ESDM ++ n = 21	Effect size f^2	F	p Value
Parent fidelity					
Baseline	3.39 ± 0.13 (n = 14)	3.40 ± 0.12 (n = 16)			
End of treat	3.18 ± 0.13 (n = 12)	3.80 ± 0.10 (n = 16)	.05	7.90	.0056
Curriculum checklist total percent (administered scales)					
Baseline	0.37 ± 0.03 (n = 23)	0.35 ± 0.04 (n = 19)			
End of treat	0.50 ± 0.04 (n = 14)	0.51 ± 0.04 (n = 18)	.01	2.01	.16
Mullen composite IQ					
Baseline	69.75 ± 3.74 (n = 24)	67.90 ± 3.89 (n = 21)			
End of treat	76.10 ± 4.28 (n = 13)	76.20 ± 4.25 (n = 17)	.01	.53	.47
ADOS2 comparison score					
Baseline	8.0 ± 0.33 (n = 24)	7.69 ± 0.33 (n = 21)			
End of treat	6.71 ± 0.51 (n = 14)	6.53 ± 0.46 (n = 18)	.001	.05	.82
Vineland adaptive behavior composite standard score					
Baseline	79.50 ± 2.32 (n = 23)	80.55 ± 2.42 (n = 20)			
End of treat	82.24 ± 2.74 (n = 14)	80.90 ± 2.64 (n = 18)	.02	1.12	.29

scoring all items as a 5. There were no statistically significant effects of treatment group on the mean scores of the groups (P-ESDM: n = 11, Mean 4.47, sd = .42; P-ESDM++: n = 16, Mean 4.63, sd = .30, t(25) = − 1.20, p = .24); however, the ES of = 0.47 indicates a medium effect favoring the P-ESDM++ group. Parents in the P-ESDM++ group rated the item "I feel that the program provided adequate training to the intervention staff working with us" significantly higher than the P-ESDM group (P-ESDM++ Mean = 4.94 (sd = .25, n = 16); P-ESDM Mean = 4.45 (sd = .82, n = 11); t(25) = 2.23, p = .035, ES = 0.81). There was a trend towards a significant difference and a moderate effect size on a second item, "How easy is it for you to conduct the PATH intervention with your child?", P-ESDM++ group Mean = 3.88 (sd = .72, n = 16), P-ESDM group Mean 3.27 (sd = .90, n = 11; t(25) = − 1.93, p = .0655, ES = 0.72).

Discussion

This study evaluated an enhanced version of a parent-implemented intervention that was designed to circumvent three difficulties that have occurred in past studies that may limit the effects of short-term low intensity treatments: (1) insensitivity of measurement tools to detect small, rapid changes in the children; (2) dosage problems in child delivery due to parent gradual learning rates of the intervention; and finally, (3) the heterogeneity and variability of treatment experiences and effects in the comparison group when using a treatment-as-usual comparison group. We addressed these three difficulties in the present study by conducting a randomized comparative treatment study in which we delivered two version of the same treatment—one the standard version and one an enhanced version to test the effects on parent and child learning of the enhancements, using both proximal and distal measures of child change.

In order for child change to occur in 12 weeks, parent learning and implementation of the intervention need to be rapid and generalized. We tested an enhancement involving the intervention dosage with the goal of increasing the rate of parental learning, which might also affect rate of child learning, since the causal model attributes treatment-related child change to improvements in frequency and sensitivity of parent–child interactions. We enhanced dosage of the parent intervention in two ways supported by the adult learning literature (Knowles 1973): (1) by creating learning materials in multiple modalities (auditory, visual, video modeling) to provide parents with choices for their preferred way to learn, and (2) by providing a second session each week in the context in which the parent was expected to use intervention (the home).

Our primary hypothesis involving the effects of the enhancements on rate of parental learning of the intervention was supported. There was a time-by-group interaction such that at the end of 12 weeks of low intensity parent coaching, parents in the P-ESDM++ group demonstrated significantly increased sensitivity and skill in supporting child social communicative development measured by increases in parent fidelity of implementation scores compared to the parents in the P-ESDM group. Based on coded video samples gathered frequently using parent-filmed videos at home in two different types of activities, parents in the P-ESDM++ group showed significantly greater increases in their skills during the 12 weeks of instruction than did the P-ESDM group. However, this finding is tempered by moderate to low inter-rater reliability. While neither group achieved our internal standard of 4.0 for fidelity of implementation, we considered the mean score of the enhanced group, 3.8, to reflect skilled use of the ESDM intervention techniques. The filming situation in this study: parent-filmed home videos of daily home routines using materials at hand without a project staff member present required much more generalization on the parents' part than in studies involving videos taken in the lab, a setting that provides considerable contextual support for skills learned in the lab.

We then examined the effects of the enhanced intervention on child progress, as measured by both standardized measures and on the PATH CC, a measure proximal to the skills targeted in the children's treatment. The enhanced treatment was not associated with greater gains on either the proximal or the standard measures of children's DQ, autism symptoms, adaptive behavior, or associated symptoms (sleep, eating, tantrums, self-injury). Thus, while there were significant gains for both groups over time, there were no group differences on degree of improvement in children's skills after 12 weeks of intervention. General maturation cannot be ruled out as the cause of the gains children made, on standardized tools as well as the CC. However, significant IQ gains and autism symptom decreases are not typical of untreated children with autism in early childhood (Lord et al. 2006; Sigman et al. 1999). Thus, the data suggest that the intervention may have resulted in the improvements seen in both groups of children. The parent skill levels in the standard group combined with the curriculum may have been enough to stimulate child improvement. Without a true no-treatment group (which is not feasible or ethical to conduct), this question cannot be answered.

As hypothesized, we found a significant positive relationship between degree of improvement in parental fidelity of implementation and increases in child social-communication and decreases in autism symptoms on the proximal measure of change, the PATH curriculum checklist, though not on the standardized measures. This is interesting, given the very high correlation between the PATH CC and the standard measures, and it may demonstrate that our proximal measure of change, paralleling the treatment goals very closely

and given very frequently, is more sensitive to child change than the standard measures given at longer intervals, as we hypothesized, as suggested by MacDonald et al. 2014.

Both groups of children demonstrated decreases in their ADOS social affect scores and significant acceleration of developmental rates on multiple measures, including one standard deviation gain (15 points) in their Mullen scores in 12 weeks. This mirrors the degree of improvement in cognitive ability found in our 2012 RCT of the P-ESDM and is consistent with the 18-point gain over 24 months found in the intensive ESDM treatment group reported in the original 2010 RCT by Dawson. In both the 2012 and the present P-ESDM studies, children receiving P-ESDM showed significant decreases on the ADOS social affect scores.

Given the association between positive parent change and child improvement, why was there no treatment group difference in proximal or distal outcomes? It may be due to within-group parent variability. Even though the P-ESDM++ was associated with significantly greater improvements in parents' ability to implement the intervention, as seen in Fig. 2, there are P-ESDM parents who show rapid positive changes and P-ESDM++ parents who do not change, and this within-group variability may be sufficient to limit the differential effects of group on child progress. It may also be that the difference between the groups on parent fidelity levels at the end of treatment, while statistically significant, may not be large enough to differentially affect child progress in a 12-week period.

A final point involves parent satisfaction with the interventions being delivered. Parents were extremely satisfied with the intervention that they received, and this is important in allaying concerns about parent implemented interventions and their potential for increasing parent stress. Interventions that are designed to be incorporated into everyday routines that parents already carry out may fit better into family life than those that require specific blocks of time to conduct. It may also be that parents of very young children are accustomed to spending a significant amount of time in child care routines, and their lives may already be organized around a higher intensity of parent–child interaction than occurs with older children. It is also interesting that the parents who had two visits per week do not report any less satisfaction than those who had only once visit, and they also report that they found the intervention easier to carry out, which may be due to more coaching, and/or coaching in the context in which they will deliver the intervention.

Strengths of the current study include the measurement strategy involving authentic assessment principles with multiple, frequent, blinded measures taken in various settings, the control gained from delivering both treatments rather than using a treatment-as-usual group, the use of both proximal and distal measures of child outcome, the high level of parent satisfaction and the complex modelling used in the analytic approach.

However, the findings from this study should be considered tentative, in light of several weaknesses that limit generalization. The primary weakness involves lower than desired intraclass intercorrelation coefficients for the primary outcome measure, the ESDM fidelity instrument, which showed only fair agreement between raters. Given that we have previously found acceptable inter-rater reliability on this measure (Vismara et al. 2009, 2012), it may be that various inconsistencies occurring when parents filmed themselves at home made the tool more difficult for trained coders to rate. Closer monitoring of rater agreement on this measure is needed, and, if necessary, further work on this instrument, to improve its psychometric qualities. A second weakness involves substantial attrition during the study. This resulted in smaller groups than were planned and may have compromised power to detect some additional group differences. Nevertheless, observation on the primary measures of fidelity and CC at eight time points (significantly more than that obtained on the other measures) helped with power to detect group differences on the main measures. A third weakness involves the confounding of the three enhancements provided to the P-ESDM++ group: a weekly home visit, MI techniques, and multimodal learning materials. Any of these three enhancements may have resulted in the enhanced performance of parents in the P-ESDM++ group. However, parent responses to one of the questions on the parent satisfaction survey suggest that home-based sessions may have been particularly helpful. This would be consistent with what is known about contextually based adult learning (Knowles 1973). However, as this study was not designed to evaluate "active ingredients" of intervention, it is only possible to speculate. Finally, the lack of follow-up data prevents the ability to determine the extent to which the treatment resulted in stable changes in parent delivery after contact with the therapist ended.

Given that this was an efficacy study rather than an effectiveness study, we cannot determine whether the results found here are generalizable to community settings. As often occurs in university-based efficacy studies of treatments, the study (1) recruited volunteer families at major university autism centers, provided diagnosis, reports, and immediate treatment, (2) most families rated themselves as highly motivated and committed to treatment throughout the study, and (3) a large majority of mothers had some college education. The children were selected to be free of known biological risk markers of ASD and the DQ range was restricted, though in practice almost no children were excluded due to DQ under 35. The treatment staffs at both sites were highly trained, highly skilled, experienced with the age group and the disability, and received expert ongoing supervision.

To examine the applicability of this intervention to community settings, we are currently conducting a community implementation study to test the P-ESDM intervention in low resource, low income areas using community workers in very low intensity public Part C intervention systems. While community implementation studies present many challenges, they are critical because they address the basic goal of autism early intervention studies: to develop interventions that can be used for the population of young children and families who need help for their child's symptoms across a nation and across the globe. The findings presented here move us one step closer to the goal.

Acknowledgments This study was supported by Autism Speaks Grants #8089 (Rogers) #8376 (Estes) and by the Children's Miracle Network (Rogers). Conflicts of interest for Rogers, Dawson, and Vismara involve royalties from publications and lectures related to the Early Start Denver Model. Thanks to children and families involved and to Gayna Guidici, Diane Larzelere, and Ehoua Vue for formatting the manuscript.

Author Contributions SJR, LV, AE, and GD conceived of the study, participated in its design and coordination and drafted the manuscript; CZ, LV, JG, MR, and SJR managed the treatment arm of the study and provide the interventions; JM, JG, and MT participated in the design of the study, and performed the measurement and data management; JM, DS, CS, and FW participated in the design of the study and performed the statistical analysis; All authors helped to draft the manuscript. All authors read and approved the final manuscript.

References

Achenbach, T. M., & Rescorla, L. (2000). *Manual for ASEBA preschool forms & profiles*. Burlington: Research Center for Children, Youth, and Families, University of Vermont.

Begg, C. B., & Iglewicz, B. (1980). A treatment allocation procedure for sequential clinical trials. *Biometrics*, 81–90.

Brian, J. A., Smith, I. M., Zwaigenbaum, L., Rogerts, W., & Brysin, S. (2016). The Social ABCs's caregiver-mediated intervention for toddlers with autism spectrum disorder: Feasibiltiy, acceptability, and evidence of promise from a multisite study. *Autism Research, 9*(8), 899–912.

Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., … Varley, J. (2010). Randomized controlled trial of the Early Start Denver Model: A developmental behavioral intervention for toddlers with autism: Effects on IQ, adaptive behavior, and autism diagnosis. *Pediatrics, 125*(1), e17–e23.

Dawson, G., Sigman, M. D., & Rogers, S. J. (2003). *CPEA Network Early Intervention Interview* (**unpublished manuscript**).

Estes, A., Vismara, L., Mercado, C., Fitzpatrick, A., Elder, L., Greenson, J., … Dawson, G. (2014). The impact of parent-delivered intervention on parents of very young children with autism. *Journal of Autism and Developmental Disorders, 44*(2), 353–365.

Gengoux, G. W., Berquist, K. B., Salzman, E., Schapp, S., Phillips, J. M., Frazier, T. W., … Hardan, A. Y. (2015). Pivotal response treatment parent training for autism: Findings from a 3-month follow-up evaluation. *Journal of Autism and Developmental Disorders, 45*, 2889–2898.

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(5), 693–705.

Hanft, B. E., Rush, D. D., & Shelden, M. L. L. (2004). *Coaching families and colleagues in early childhood*. Baltimore: Paul H. Brookes.

Hardan, A. Y., Gengoux, G. W., Berquist, K. L., Libove, R. A., Ardel, C. M., Phillips, J., … Minjarez, M. B. (2015). A randomised controlled trial of Pivotal Response Treatment Group for parents of children with autism. *Journal of Child Psychology and Psychiatry, 56*(8), 884–892.

Kasari, C., Gulsrud, A., Paparella, T., Hellemann, G., & Berry, K. (2015). Randomized comparative efficacy study of parent-mediated interventions for toddlers with autism. *Journal of Consulting and Clinical Psychology, 83*(3), 554.

Kasari, C., Gulsrud, A. C., Wong, C., Kwon, S., & Locke, J. (2010). Randomized controlled caregiver mediated joint engagement intervention for toddlers with autism. *Journal of Autism and Developmental Disorders, 40*(9), 1045–1056.

Kasari, C., Siller, M., Huynh, L. N., Shih, W., Swanson, M., Hellemann, G. S., & Sugar, C. A. (2014). Randomized controlled trial of parental responsiveness intervention for toddlers at high risk for autism. *Infant Behavior and Development, 37*(4), 711–721.

Knowles, M. (1973). *The adult learner: A neglected species*. Houston: Gulf.

Lord, C., Luyster, R., Gotham, K., & Guthrie, W. (2012). *Autism Diagnostic Observation Schedule—Toddler Module. Los Angeles: Western Psychological Services*.

Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychaitry, 63*, 694–701.

MacDonald, R., Parry-Cruwys, D., Dupere, S., & Ahearn, W. (2014). Assessing progress and outcome of early intensive behavioral intervention for toddlers with autism. *Research in Developmental Disabilites, 35*, 3632–3644.

Magiati, I., Tay, X. W., & Howlin, P. (2014). Cognitive, language, social and behavioural outcomes in adults with autism spectrum disorders: A systematic review of longitudinal follow-up studies in adulthood. *Clinical Psychology Review, 34*(1), 73–86.

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

McWilliam, R. A. (2010). Routines-based early intervention. In *Supporting young children and their families*. Baltimore: Paul H. Brookes.

Miller, W. R., & Rollnick, S. (2012). *Motivational interviewing: Helping people change*. New York: Guilford Press.

Mullen, E. M. (1995). *Mullen scales of early learning*. Circle Pines: AGS.

Prochaska, J. O., Redding, C. A., & Evers, K. E. (1997). The transtheoretical model and stages of change. In K. Glanz, F. M. Lewis & B. K. Rimer (Eds.), *Health behavior and health education* (pp. 60–84).

Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The Modified Checklist for Autism in Toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 31*(2), 131–144.

Rogers, S. J., & Dawson, G. (2010). *Early Start Denver Model for young children with autism: Promoting language, learning, and engagement*. New York: Guilford Press.

Rogers, S. J., Dawson, G., & Vismara, L. A. (2012a). *An early start for your child with autism: Using everyday activities to help kids connect, communicate, and learn*. New York: Guilford Press.

Rogers, S. J., Dawson, G., Zierhut, C., Winter, J., McCormick, C., & Holly, E. (2013). *The PATH curriculum checklist for young children with autism*. (**unpublished document**).

Rogers, S. J., Estes, A., Lord, C., Vismara, L., Winter, J., Fitzpatrick, A., … Dawson, G. (2012b). Effects of a brief Early Start Denver Model (ESDM)–based parent intervention on toddlers at risk for autism spectrum disorders: A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(10), 1052–1065.

Rogers, S. J., & Vismara, L. A. (2013). *P-ESDM parent coaching fidelity measure* (**unpublished document**).

Schertz, H. H., Odom, S. L., Baggett, K. M., & Sideris, J. H. (2013). Effects of joint attention mediated learning for toddlers with autism spectrum disorders: An initial randomized controlled study. *Early Childhood Research Quarterly, 28*(2), 249–258.

Sigman, M., Ruskin, E., Arbeile, S., Coronoa, R., Dissanayake, C., Esponosa, M., … Zierhut, C. (1999). Continuity and change in the social competence of children with autism, Down syndrome, and developmental delays. *Monogrpah of the Society for Research in Child Development, 64*(1), 1–114.

Siller, M., & Sigman, M. (2002). The behaviors of parents of children with autism predict the subsequent development of their children's communication. *Journal of Autism and Developmental Disorders, 32*(2), 77–89.

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland adaptive behavior scales: (Vineland II), survey interview form/caregiver rating form*. Livonia, MN: Pearson Assessments.

Swinkels, S. H., Dietz, C., van Daalen, E., Kerkhof, I. H., van Engeland, H., & Buitelaar, J. K. (2006). Screening for autistic spectrum in children aged 14 to 15 months. I: The development of the Early Screening of Autistic Traits Questionnaire (ESAT). *Journal of Autism and Developmental Disorders, 36*(6), 723–732.

Vismara, L., Colombi, C., & Rogers, S. J. (2009). Can one hour per week of therapy lead to lasting changes in young chidlren with autism? *Autism, 13*, 93–115.

Vismara, L., Young, G. S., & Rogers, S. J. (2012). Telehealth for expanding the reach of early autism training to parents. *Autism Research and Treatment*. https://doi.org/10.1155/2012/121878.

Wetherby, A. M., Guthrie, W., Woods, J., Schatschneider, C., Holland, R. D., Morgan, L., & Lord, C. (2014). Parent-implemented social intervention for toddlers with autism: An RCT. *Pediatrics, 134*(6), 1084–1093.

Wetherby, A. M., & Prizant, B. M. (2002). *Communication and symbolic behavior scales: Developmental profile*. Baltimore: Paul H. Brookes.