Journal of Autism and Developmental Disorders (2021) 51:4534–4550
https://doi.org/10.1007/s10803-021-04900-5

ORIGINAL PAPER

# A Longitudinal Study of Language Trajectories and Treatment Outcomes of Early Intensive Behavioral Intervention for Autism

Thomas W. Frazier[1,2,3] · Eric W. Klingemier[4] · Christa J. Anderson[5] · Grace W. Gengoux[6] · Eric A. Youngstrom[7] · Antonio Y. Hardan[6]

Accepted: 27 January 2021 / Published online: 8 February 2021


## Abstract

The present study examined language trajectories and placement outcomes for children with autism spectrum disorder (ASD) receiving early intensive behavioral intervention (EIBI). Language measures were collected at baseline and 6, 12, 18, 24, and 36 months or until exit from EIBI in 131 children with ASD. Growth models estimated overall and subgroup language trajectories. Overall, children receiving EIBI showed substantial increases in language relative to normative expectations. Earlier age at EIBI start, higher baseline cognitive function, and lower baseline ASD severity predicted better language trajectories. Although there was significant variability in language trajectories and educational outcomes, most children showed significant increases in language scores, relative to normative expectations. Additional research, in more representative samples, is needed to understand this variability.

**Keywords** Applied behavior analysis · Language · Outcomes · Education · Growth model · Prediction

An accumulation of evidence from early developmental and behavioral interventions in children with Autism Spectrum Disorder (ASD) has demonstrated efficacy in enhancing cognitive development, reducing symptomology, and increasing adaptive function (Dawson et al. 2010; Eldevik et al. 2009; Granpeesheh et al. 2009; Green et al. 2017; Hardan et al. 2015; Howard et al. 2005; Howlin et al. 2009; Kasari et al. 2008; Mohammadzaheri et al. 2014; National Research Council & Division of Behavioral and Social Sciences Education 2001; Peters-Scheffer et al. 2011), in addition to providing long-term cost-effectiveness (Peters-Scheffer et al. 2012). However, reservations over the methodological rigor and overall strength of the evidence has resulted in ambiguity over the efficacy of early behavioral interventions for ASD (Ospina et al. 2008; Warren et al. 2011). This ambiguity has prompted additional funding (e.g. Department of Defense Autism Research Program and Autism Care Demonstration) to better understand the efficacy of early behavioral interventions.

✉ Thomas W. Frazier
tfrazier@jcu.edu

Eric W. Klingemier
klingee@ccf.org

Christa J. Anderson
andersonchrist@missouri.edu

Grace W. Gengoux
ggengoux@stanford.edu

Eric A. Youngstrom
eay@unc.edu

Antonio Y. Hardan
hardanay@stanford.edu

1 Center for Autism and Genomic Medicine Institute, Cleveland Clinic, Cleveland, OH, USA

2 Autism Speaks, New York, NY, USA

3 Department of Psychology, John Carroll University, Cleveland, OH, USA

4 Center for Autism and Caregiver Experience, Cleveland Clinic, Cleveland, OH, USA

5 Department of Health Psychology, University of Missouri, Columbia, MO, USA

6 Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA

7 Departments of Psychology and Neuroscience and Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

## Structured Behavioral Interventions

Many early intervention packages are based on Applied Behavior Analysis (ABA), which is a learning theory associated with a set of principles and techniques, underlying both structured and naturalistic intervention approaches (Schreibman et al. 2015), intended to increase the frequency and strength of desired behaviors and decrease or extinguish undesired behaviors (Cooper et al. 2007). In early intensive behavioral intervention (EIBI) programs, the structured application of ABA relies heavily on the use of discrete trials and positive reinforcement to alter stimulus–response contingencies (Landa 2007). ABA-based EIBI utilizes patient-specific functional analyses of behaviors to develop individualized intervention targets where complex behaviors are broken down into specific steps and reinforced to gradually approximate the desired behavior and ultimately enhance learning (Mohammadzaheri et al. 2014). Throughout this process, discrete trials are repeatedly administered using target stimuli to strengthen the relationship between a specific stimulus-behavior pairing (learning trials) until the target behavior/response is achieved. This ideally requires strict control over the discrete trial process and stimuli, with minimal distractors in the teaching environment. During early learning trials, primary reinforcements (e.g., edibles) are often paired with secondary reinforcements (e.g., verbal praise); with the rationale that more naturally occurring secondary reinforcers will eventually replace primary reinforcers to help promote generalization.

## Naturalistic Developmental and Behavioral Intervention

In contrast to the structured application of ABA, naturalistic developmental and behavioral interventions (NDBIs) merge ABA principles with developmental science (Schreibman et al. 2015) by emphasizing developmental approaches to teaching through the acquisition of prerequisite skills before more advanced behaviors are addressed. This is accomplished through shared control of therapeutic focus and goals (Schreibman et al. 2015) through child-led activities (Koegel et al. 1987), whereby the child's spontaneous behaviors and interests guide therapeutic responses (Mahoney and Perales 2003) resulting in learning and maintenance task variation (Dunlap 1984). Throughout this process, child engagement and motivation is maximized (Koegel et al. 1999) enabling natural reinforcement strategies to be used (Koegel and Williams 1980). Furthermore, NDBI's approach addresses concerns that structured ABA therapies that can lead to prompt-dependence, limited generalization, and child avoidance of strictly prescribed activities that do not provide them with natural motivation and attention (Schreibman 2005); while common features of both structured and naturalistic EIBI programs are the early application (before age 3–4) and intensive nature of treatment (20–40 h per week).

## Early Intensive Behavioral Intervention Outcome Measures

Over the last two decades, structured and naturalistic EIBI variants have demonstrated improvements across a range of developmental outcomes (Cohen et al. 2006; Dawson et al. 2010; Eikeseth et al. 2002; Eldevik et al. 2009; Howard et al. 2005; Mohammadzaheri et al. 2014; Smith et al. 2000). These results have suggested high rates of mainstream school placement (Howard et al. 2005; Howlin et al. 2009; Lovaas 1987), positive long-term outcomes for a proportion of individuals (Landa and Kalb 2012; McEachin et al. 1993), and increased language and communication skills (Dawson et al. 2010; Eikeseth et al. 2002). However, study reviews have emphasized the high variability in EIBI treatment response (Howlin et al. 2009; Warren et al. 2011). Many of these studies have utilized small sample sizes and lack details regarding longitudinal measurements acquired throughout intervention, which is likely due to the high time-related demands from therapists and cost to implement these procedures. As a result, there remains a gap in our understanding of individual differences in the neurobehavioral changes that occur during EIBI programs.

*Language* is an important response target for many EIBI programs (Almirall et al. 2016; Dawson et al. 2010; Kasari et al. 2008; Koegel et al. 1987), as young children with ASD often show difficulties that range from limited receptive and expressive language to mild pragmatic difficulties in highly fluent individuals with ASD (Tager-Flusberg et al. 2005). In fact, studies have demonstrated EIBI-related gains (from medium to very large) in verbal IQ and other language assessments (Dawson et al. 2010; Hardan et al. 2015; Howlin et al. 2009) that include receptive and expressive measures. However, many of these studies have reported outcomes from short-term or less intensive interventions (Green et al. 2017; Hardan et al. 2015; Kasari et al. 2008), or from a treatment midpoint and/or endpoint measure (Dawson et al. 2010; Howard et al. 2005). In order to maximize outcomes, a better understanding of the comprehensive developmental trajectory of receptive and expressive language during *long-term* EIBI, as well as predictors of language trajectories during intervention, is necessary. Predictive information could provide early indication of individual language trajectories (positive or negative) to allow for early intervention modification and appropriate allocation of resources.

Baseline cognitive ability (global IQ, verbal IQ, nonverbal IQ, receptive language) and adaptive behaviors have been consistent predictors of more positive change during EIBI treatment (Ben-Itzchak et al. 2014; Ben-Itzchak and Zachor 2007; Howlin et al. 2009; Sallows and Graupner 2005; Smith et al. 2015) and, in at least a few cases, higher cognitive ability (non-verbal ability or developmental quotient) has predicted greater EIBI efficacy within a randomized controlled trial (Hardan et al. 2015; Rogers et al. 2019). Lower ASD severity (Smith et al. 2000), improved social engagement (Ingersoll et al. 2001; Smith et al. 2015), joint attention (Mundy and Crowson 1997; Mundy et al. 1990), imitation (Sallows and Graupner 2005), and play skills (Ingersoll 2010; Kasari et al. 2008; Sherer and Schreibman 2005) have also been associated with positive outcomes in the context of treatment, but with much less consistency than cognitive ability. Further, evaluation of these predictors has predominantly occurred in studies using a single treated group where prediction of efficacy or effectiveness is conflated with prediction of natural development. In addition, age of treatment initiation has not been a consistent predictor of better outcomes in treated groups, despite the well-known impetus to initiate intervention as early as possible (Howlin et al. 2009; Smith et al. 2015); however, these findings may be influenced by methodology where small samples and restricted age ranges were used to assess early intervention age. Finally, outside of EIBI studies, early motor skills (Bedford et al. 2016) and other behavioral profiles (Bopp et al. 2009) have been linked with differences in developmental outcomes. Thus, prediction of developmental outcomes either prior to or early into EIBI could be used to guide treatment targets, track progress, and help to plan for transitions to educational environments.

## The Present Study

The current study seeks to address methodological issues in EIBI outcome research, including small sample sizes, lack of repeated measurements throughout intervention, and inadequate detail of procedures, which have previously resulted in inconsistencies and ambiguity of outcomes. Our approach addresses these limitations by reporting on the largest clinically ascertained sample of children with ASD receiving EIBI to date (to our knowledge) with planned longitudinal language data collection and post-treatment educational placement outcomes. The inclusion of clinically-ascertained participants is advantageous, as prior studies have been criticized for recruiting samples with relatively high cognitive ability and functional levels, which inadequately represents real-world clinical EIBI patients (Vivanti et al. 2014), restricting the usefulness of the outcome data.

We have addressed four specific aims in the current study. First, we sought to estimate overall language trajectories from EIBI initiation to exit; we expected that children receiving EIBI with ASD would show large overall increases in language and vocabulary scores relative to age norms (standard scores $\geq$ 10 points by 36 months of treatment), but that substantial individual variability would be observed. The expectation of a large overall increases in language scores is based on EIBI clinical trials and systematic reviews suggesting medium-to-large increases in language and/or cognitive ability (Dawson et al. 2010; Eldevik et al. 2009; Howlin et al. 2009; Peters-Scheffer et al. 2011; Rogers et al. 2019). Our second aim was to evaluate demographic and clinical predictors of overall language trajectories. Third, given high variability in overall trajectories, we explored whether sub-groups of language trajectories and their correlates could be empirically-identified. Based on prior research (Bopp et al. 2009; Fountain et al. 2012; Lord et al. 2012a, b; Smith et al. 2015) and the clinical observation of fast, moderate, and slow learning patterns during EIBI treatment, we expected to observe at least 3 language trajectory sub-groups and expected that younger age at treatment initiation, lower autism symptom severity, and higher baseline language and cognitive ability would be significant predictors of more positive trajectories. Finally, we aimed to use information easily acquired at baseline and early within the intervention period to predict post-EIBI educational placement outcomes. Post-treatment educational placement has been emphasized as an important endpoint for understanding the benefits of EIBI (Howlin et al. 2009; Lovaas 1987) and is essential for caregivers as they examine future logistical and financial considerations for child placement needs. Based on existing literature, we anticipated that the earlier age of treatment initiation, lower ASD symptom severity scores, higher baseline cognitive/language skills, and early treatment gains in language would predict more positive language trajectories and placement outcomes.

## Method

### Participants

Additional description of the methods is included in the supplemental online resources. The local IRB reviewed and approved the clinical data collection procedure for the current project. Participants consisted of children with a confirmed ASD diagnosis ($n = 131$; 114 male; ages 1.4–5.9) who were enrolled in a single-site EIBI program at an academic medical center between 08/2002 and 03/2014. Inclusion/exclusion criteria was not based on symptom, cognitive, or functional severity, which is consistent with typical clinical EIBI implementation and effectiveness research designs (Gartlehner et al. 2006). Data were collected as part of routine care, which included (a) prospectively planned language data, (b) age in years at program start (e.g. 2.5, 3.2) and race/ethnicity (coded as white Non-Hispanic versus other race/ethnicity) retrieved from medical records, (c) median household income, estimated from residential zip code (Berkowitz et al. 2015), and (d) other data elements that were coded post-hoc from available routine clinically-acquired information. Calendar year of entry (e.g., 2006, 2007, etc.) was recorded and explored as a predictor of language trajectories and outcomes.

*Consensus clinical DSM-IV-TR or DSM-5 ASD diagnosis* was made by a multi-disciplinary evaluation team that included (a) parent interview by a clinical psychologist, (b) developmental evaluation by a physician, and (c) cognitive/language testing by a trained psychology aide or speech/language pathologist. DSM-IV-TR diagnoses were reviewed and all children met criteria for DSM-5.

## Clinical Assessments

When feasible, language measures were collected at baseline and 6, 12, 18, 24, and 36 months post program entry through exit. These language measures consisted of *receptive and expressive measures* taken from the Preschool Language Scales—Fourth or Fifth Edition (PLS) (Zimmerman et al. 2011), *receptive and expressive one-word vocabulary* obtained from the Receptive One-Word Picture Vocabulary Test—Third Edition (ROWPVT) or the Peabody Picture Vocabulary Test—Third or Fourth Edition (PPVT) (Dunn and Dunn 2007 1997), and the Expressive One-Word Picture Vocabulary Test—Third Edition (EOWPVT) or the Expressive Vocabulary Test—First or Second Edition (EVT) (Williams 1997, 2007). Verbal and non-verbal IQ estimates were obtained from either the Mullen Receptive Language and Visual Reception subtests (Mullen 1995) or the Wechsler Preschool and Primary Scales of Intelligence—Third or Fourth Edition (WPPSI) (Wechsler 2002, 2012), based on age and ability of the child. All of these measures have previously shown good internal consistency reliability and construct validity across the relevant ages in this sample. Verbal IQ estimates from well-established standardized measures were prioritized, however, if these measures were not available for a participant, then all available verbal test information was averaged across (a) PLS Total Language, (b) PPVT, (c) EVT, and (d) Mullen Receptive and Expressive Language or WPPSI subtest scores.

Other measures included: (a) cognitive testing (Mullen or WPPSI) to estimate *intellectual disability*, (b) Vineland Adaptive Behavior Scales—Caregiver Rating Form to determine *adaptive function* later in treatment, and (c) educational and/or medical record descriptions to document additional behaviors and/or medical comorbidities. ASD symptomology and traits were also measured using both clinician-rated and parent-report measures. Clinician-rated measurements included calibrated severity scores from the Autism Diagnostic Observation Schedule—First or Second Edition (ADOS) (Lord et al. 2002, 2012a; b) and raw scores from the Childhood Autism Rating Scale (CARS) (Esler et al. 2015; Gotham et al. 2009; Lord et al. 2012a, b; Luyster et al. 2009; Schopler et al. 2010). Parent-report ASD trait measures utilized age-adjusted Total T-scores from the Social Responsiveness Scale—Second Edition (SRS-2; (Constantino and Gruber 2012). Each of these measures has been shown to have strong internal consistency and/or test–retest reliability (Constantino and Gruber 2012; Garfin et al. 1988; Lord et al. 2012a, b; Schopler et al. 2010), good screening or diagnostic accuracy/agreement (Chlebowski et al. 2010; Constantino and Gruber 2012; Molloy et al. 2011; Risi et al. 2006; Schopler et al. 2010; Ventola et al. 2006a), and prior research has found good convergent validity between these measures (Bolte et al. 2008; Ventola et al. 2006b). To address longitudinal inconsistencies in the collection of ASD symptom measures, a clinical judgment of ASD severity, using *mild* (= 1) or *moderate to severe* (= 2) categories, was created using the following interpretation guidelines, *mild* was scored as ADOS total severity score 4–5, CARS raw score 30–36.5, and SRS Total T-score 60–75; and *moderate to severe* was scored as ADOS total severity score 6 +, CARS raw score 37 +, and SRS Total T-score 76 +. In six cases, none of these ASD severity measures were available, therefore severity was categorized using clinical diagnostic team judgement with the use of all available baseline clinical information that they had obtained for those children. All clinical measures were administered according to standardized procedures.

## Early Intensive Behavioral Intervention Procedures

Each child's EIBI program initially utilized a structured ABA approach with heavy emphasis on 1:1 discrete trial teaching and primary reinforcement to acquire initial compliance-based skills (ex. sit in chair, look at me, hands down, etc.). Once acquired, discrete trials teaching continued to ensure compliance skill maintenance and more naturalistic child-led activities using secondary reinforcements were included. For example, children participated in group activities using a *leader-prompter model* in preschool environments. The leader-prompter model has a teacher at the front with students sitting around the teacher in a semi-circle with therapists behind each student acting as prompters. Various topics can be presented by the teacher ranging from social skills to new academic concepts and appropriate behavioral responding is reinforced by the prompter. Weekly classroom-based intervention times varied by age of the child; children < 3 received up to 30 h, and those ≥ 3 up to 32.5 h. In addition, all children received 1 h per week of 1:1 speech-language therapy that focused on building basic receptive and expressive language, and advanced sentence-level production and pragmatic language skills for more advanced students. Nearing the end of the intensive treatment phase children received a therapist who shadowed them to help facilitate transition to the new educational environment.

To support placement decisions, post-EIBI outcomes were obtained from the treatment team and ordinally coded as follows: *1-mainstream placement* with no special education supports; *2-minimal educational supports* within a typical classroom placement (no aide) that included an IEP or 504 plan to specify accommodations and therapies needed (speech/language or occupational therapy); *3-substantial educational supports* either within special education classrooms that have a low student:teacher ratio, or within a typical classroom with a 1:1 aide without ABA or behaviorally-based program support; and *4 = intensive behavioral intervention support*, with low ratio 1:1 or 2:1 ABA or specialty behaviorally-based placement. Then, the treatment team shared these outcome recommendations with the parents, who then made a final decision regarding educational placement for their child.

## Statistical Analyses

Descriptive statistics (M, SD, range) were used to characterize demographic and clinical features of the sample. No univariate or bivariate outliers were detected. Analyses with and without very low and very high language scores did not influence the pattern of language trajectories. Therefore, all cases were included in subsequent analyses. Missing data at follow-up time points were examined by estimating differences between cases with and without data on baseline *demographic* (age at EIBI start and sex) and *clinical factors* (ASD severity, time in program [exit minus entry in months], and verbal and non-verbal IQ).

### Aim 1a (Overall Trajectories)

To examine overall language trajectories, we computed random intercept and slope growth curve models separately for each of the five language measures (PLS total language, receptive language, and expressive language; ROWPVT or PPVT receptive vocabulary; and EOWPVT or EVT expressive vocabulary). Then, language measures obtained at baseline and 6, 12, 18, 24, and 36 months post-EIBI initiation were used to characterize language development throughout intervention using growth models. Maximum likelihood estimation was used and model fit was evaluated using the Comparative Fit Index (CFI) and Bayesian Information Criterion (BIC) (Bentler 1988). Comparisons were made between models with and without quadratic parameters to determine whether including this parameter enhanced description of the growth process. In all cases, adding the quadratic parameter substantially increased fit and, therefore, models are presented with linear and quadratic slopes. We also investigated whether the calendar year of entry into the EIBI program (e.g., 2006, 2007, etc.), race, and median household income were correlated with intercept and slope parameters from growth models and educational placement outcomes.

### Aim 2 (Predictors of Overall Language Trajectories)

To examine what factors may be important correlates of language ability at entry and during the course of EIBI, overall growth models from Aim 1a were re-estimated adding demographic (age at EIBI start, sex) and clinical predictors (ASD severity, time in program (months from entry to exit), verbal IQ, and non-verbal IQ) of baseline language ability (intercept) and change in language ability during intervention (linear slope). An early reviewer of this work suggested examining race/ethnicity, median household income, and the calendar year of entry into the EIBI program. Therefore, these were included in post-hoc analyses. Calendar year of entry into the EIBI program evaluates whether program improvements over time, particularly a gradual shift toward increased use of naturalistic teaching strategies as the field become more aware of their value, might predict positive language trajectories. It is important to note that, in these models, estimates of individual change can be highly variable. For example, a child with a low rate of change early in treatment can have the same end point as a child with a high rate of change early in treatment because the children may differ in their later patterns of change reflected in the quadratic slope. Furthermore, predictors of linear change in this model are relevant to early change patterns and not to later differences in the pattern of change.

### Aim 3 (Sub-group Trajectories)

When growth models identify significant variability in overall longitudinal trajectories, it is possible to estimate sub-group trajectories that represent distinct patterns of growth. To estimate sub-group trajectories, growth mixture models were computed using the same underlying growth process as described above and estimated across an increasing number of latent classes (1–4). Model fit was evaluated using the BIC (Bentler 1988; Nylund et al. 2007) and changes in BIC (Raftery 1995), with smaller BIC values and decreases in BIC > 5 indicative of better fit. Entropy was also included to describe whether more complex class structure resulted in substantially improved classifications. Model classifications were saved, and class descriptions were computed using analysis of variance and chi-square statistics.

#### Aim 4 (Predictors of Educational Placement Outcomes)

To examine potential predictors of educational placement support at EIBI program exit, we first computed bivariate Spearman's Rho correlations between ordinal educational placement support code (1–4) and demographic information (age at EIBI start, sex), baseline clinical characteristics (ASD severity, verbal IQ, non-verbal IQ), baseline language measures, and 6-month change in language measures. These measures were chosen a priori from the available data because they are commonly used in clinical practice, are often available to clinicians prior to treatment initiation, and existing data suggest they may be useful in predicting outcomes (Hardan et al. 2015; Howlin et al. 2009; Sallows and Graupner 2005). Inclusion of 6-month changes was chosen a priori to determine whether this information might be useful to assist clinicians in providing information to caregivers regarding future placement. The choice was supported by overall growth models showing that most language change occurred within the first 6–12 months. Based on the results of bivariate correlations, two separate prediction models of educational support at EIBI exit were computed using probit regression. The first model included age, ASD severity, total language at baseline, and 6-month change in total language as predictors. The second model included age, ASD severity, and non-verbal IQ. A model with only non-verbal IQ was chosen a priori because baseline and 6-month language follow-up may not be available in all clinical environments and non-verbal IQ has been found to be a strong predictor of outcome in prior research (Hardan et al. 2015; Howlin et al. 2009). Model results were internally cross-validated using training and test subsamples (70% and 30% of cases, respectively). Total language was the only language measure included in these models because it was more strongly correlated with educational support scores than other measures.

SPSS version 25 was used to compute descriptive statistics, chi-square, and ANOVA models (IBM Corp 2018). MPlus version 7.3 was used to compute all growth, growth mixture, and probit regression models (Muthén and Muthén 1998–2012). *Online Resource 1* provides additional methodological details including information on participants, diagnoses, clinical assessments, definitions for key descriptive and predictive variables, EIBI procedures, statistical analyses, and statistical power calculations. *Online Resource 2* presents the STROBE checklist for cohort studies (von Elm et al. 2007).

## Results

### Participant Characteristics

131 children (114 male; ages 1.4–5.9) were enrolled in the EIBI program and had language data (Table 1), six children were excluded due to not have any language assessments at any timepoint. While the modal time for EIBI program enrollment was 2 years, individual data showed high variability among length of enrollment with (*a*) 2% ($n=3$) being enrolled for < 6 months; (*b*) 19% ($n=25$) between 6 and 12 months; (*c*) 17% ($n=22$) between 13 and 18 months; (*d*) 18% ($n=38$) between 19 and 24 months); (*e*) 29% ($n=38$) between 25 and 36 months; and (*f*) 15% ($n=20$) for 37 + months. Age at entry was moderately negatively correlated with time in program ($r=-0.37$, $p<0.001$). Upon EIBI completion, approximately 38% of the children exited to a *no support* or *minimal support* educational placement, while 28% required significant *special education* support without *intensive behavioral intervention*, with 34% requiring ongoing *intensive behavioral intervention.* ASD symptom severity ranged from *moderate* to *severe*, and as such *baseline language* and *cognitive* scores were highly variable and fell, on average, at the lower end of the borderline range (Standard Score (SS) ~ 70). However, only 28.2% of the sample was identified as having an intellectual disability, which may be under-reported in the medical records for children in this age range, in addition, the average *adaptive behavior* composite score was in the low range. *Online Resource 3* presents age at EIBI entry and exit, length of follow-up, and missing and available data at relevant ages and timepoints.

### Overall Changes in Language with EIBI

In line with *Aim 1*, growth models estimated a substantial increase in all 5 language measures (linear slope = 1.09–1.37, SE = 0.14–0.17, all $p<0.001$), with increases from baseline to 36 months of ~ 10 standard score points (Fig. 1). Most growth occurred within the first 12 months, with continued gradual growth to 24 months, followed by a slight downturn from 24 to 36 months relative to expected trajectories (quadratic slope = − 0.02 to − 0.03, $p<0.001$). The overall pattern was highly consistent across measures.

### Predictors of Overall Language Trajectories

Earlier age at program start, lower autism severity, higher verbal and non-verbal ability estimates predicted better trajectories across language and vocabulary measures. Younger

**Table 1** Sample characteristics

| | M (SD, range) |
|---|---|
| Demographics | |
| N | 131 |
| Age at program start | 3.3 (1.0, 1.4–5.9) |
| Sex (n, %) | |
| Male | 114 (87.0%) |
| Female | 17 (13.0%) |
| Race/ethnicity (n, %) | |
| White non-Hispanic | 103 (78.6%) |
| Other race/ethnicity | 28 (21.4%) |
| Household Income ($) | $72,165 (22,001, 29,787–150,625) |
| Time in program (months) | 24.0 (11.7, 2–49) |
| Educational support at exit (n, %) | |
| No support/mainstream | 22 (17%) |
| Minimal support | 27 (21%) |
| Significant support | 37 (28%) |
| 1:1 IBI support | 45 (34%) |
| Autism symptoms | |
| ASD severity (n, %) | |
| Mild | 52 (39.7%) |
| Moderate/severe | 79 (60.3%) |
| CARS (n = 68) | 36.7 (7.2, 25–56) |
| ADOS severity score (range 1–10) (n = 30) | 7.3 (2.2. 2–10) |
| SRS total T-score (n = 44) | 88.3 (17.3, 55–131) |
| Language | |
| Total language (SS) | 68.2 (20.4, 50–150) |
| Receptive language (SS) | 69.0 (21.2, 50–150) |
| Expressive language (SS) | 70.7 (19.2, 50–150) |
| Receptive vocabulary (SS) | 70.0 (23.3, 50–148) |
| Expressive vocabulary (SS) | 70.0 (23.1, 50–144) |
| Cognitive and adaptive | |
| Verbal IQ | 70.2 (21.3, 45–150) |
| Non-verbal IQ | 67.9 (21.5, 45–115) |
| Adaptive function composite (SS) | 56.7 (24.5, 13–108) |
| Intellectual disability (n, %) | 37 (28.2%) |

Time in Program (months) represents the number of months from the child starting in the program to their exit. *SS* Standard Score, *CARS* Childhood Autism Rating Scale, *ADOS* Autism Diagnostic Observation Schedule, *SRS* Social Responsiveness Scale. Total Language was measured using the PLS Total Language Standard Score. Receptive Language was measured using the PLS Auditory Comprehension Standard Score. Expressive Language was measured using the PLS Expressive Communication Standard Score. Receptive Vocabulary was measured using the Receptive One-Word Picture Vocabulary Test—Third Edition or Peabody Picture Vocabulary Test—Fourth or Fifth Edition. Expressive Vocabulary was measured using the Expressive One-Word Picture Vocabulary Test—Third Edition or the Expressive Vocabulary Test—First or Second Edition. Non-Verbal IQ n = 44. Adaptive Function Composite n = 55 was collected using the Vineland Adaptive Behavior Scales-Caregiver Rating Form, but was collected at variable times during program enrollment, with most coming later in the EIBI program. Intellectual Disability diagnosis was only coded later during treatment using all available data. For these reasons, adaptive function composite and intellectual disability are included only for descriptive purposes and not as potential predictors of outcome

age at program start was associated with longer time in EIBI treatment ($r = -0.37$, $p < 0.001$). This likely reflects the fact that children who receive an earlier ASD diagnosis are often more impaired, have a longer time available for EIBI prior to school-age, and require additional treatment prior to placement. Longer treatment (time in program) was also a significant independent predictor of worse language trajectories. This suggests that staying longer in the program is indicative of worse language trajectory, regardless of the age at program start, and reflects the need for ongoing intervention.

Post-hoc analyses indicated that race/ethnicity and household income were not associated with intercept or slope estimates from these models (all $p > 0.100$), but calendar year of entry into the EIBI program was significantly positively correlated with greater improvements in total language ($r = 0.20$, $p = 0.019$), expressive language ($r = 0.21$, $p = 0.013$), and expressive vocabulary ($r = 0.27$, $p = 0.001$), reflecting program improvements over time. Online Resource 4 presents a priori predictors of intercept and slope for all growth models. Fit for all models was adequate to good (CFI = 0.84–0.92), but suggested that further improvements might be possible by estimating sub-group (class) structure.

## Sub-Groups of Language Development During EIBI

A growth mixture model fit indicated that a 3-class structure was optimal for all language measures except receptive one-word vocabulary, where a 2-class structure was fit best (*Online Resource 5*). In general, larger numbers of classes (4 +) tended to have a worse fit, producing at least one class with a very small base rate (< 3%) without substantial improvements in ASD classification accuracy. For comparison across measures, we estimated 3-class structure for all measures and saved classifications for subsequent analyses. However, it should be noted that, for receptive vocabulary, where a 2-class structure fit slightly better than a 3-class structure, the two classes with higher baseline receptive vocabulary scores and large increases in receptive vocabulary over time are probably best represented as a single class.

All five language measures had a class with low scores at baseline and very little improvement in standard scores over time, but the size of this class varied across measures (32–62%; Fig. 2a–e). This class was largest for receptive language and smallest for receptive one-word vocabulary, possibly reflecting problems with more complex language structures in a subset of children with ASD. Each measure also had one class which started with low or borderline scores and showed a substantial increase over time, with the biggest increases being for receptive language and expressive vocabulary. Additionally, each measure had a class with average initial language scores, but different patterns of change over time across measures. Class overlap across measures was

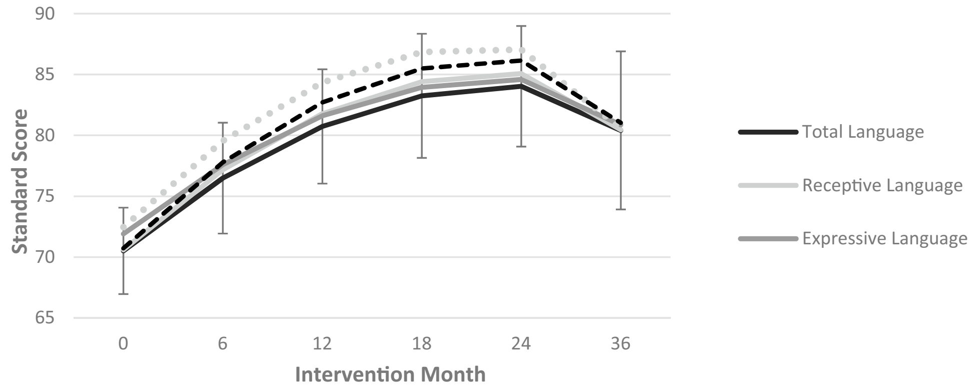


**Fig. 1** Growth model results for all language measures from baseline to 36 months. Confidence intervals are included only for total language as these intervals were highly consistent in magnitude across different measures and including for all measures obscured the point estimates

moderate to strong (class agreement 48.5–83.2%; weighted K = 0.32–0.79).

Age at EIBI program start was significantly associated with growth model classifications for total, receptive, and expressive language measures (Table 2). Children who were older at entry tended to fall in the low (Class 1) or borderline (Class 2) initial language classes, which is consistent with older age of EIBI entry being associated with worse language trajectories in the overall analysis. However, age at EIBI program entry was not significantly associated with vocabulary trajectory classifications. Not surprisingly, for all measures, classes with better language trajectories spent less time in the EIBI program. Moderate to severe ASD symptoms and intellectual disability were more prevalent in classes with worse language trajectories. Finally, sex was not associated with latent class structure.

### Prediction of Educational Placement Support at Exit

Younger age at EIBI start, lower ASD symptom severity, higher verbal and non-verbal IQ, higher baseline language, and larger 6-month increases in language measures were all significantly correlated with less educational support required at exit (Online Resource 6). Non-verbal IQ and 6-month changes in total language were the strongest associations ($r = -0.64$ for both measures). Given the observed predictive overlap between verbal and non-verbal IQ and the potential that some programs will only collect language data, we estimated two separate probit regression models. The first model indicated significant independent contributions from age at EIBI start, ASD symptom severity, and baseline and 6-month changes in total language (Fig. 3), accounting for 55% of the variance in educational placement support outcomes. The second model indicated significant independent predictive contributions from ASD symptom severity and non-verbal IQ, accounting for 53% of the variance in educational placement outcomes. Age at EIBI start was not significant in this model but trended toward significance and, for this reason, was retained as a predictor. Both models showed good internal cross-validation (Online Resource 7). Chronological year of entry into the EIBI program ($r < 0.01$, $p = 0.997$), race ($r = -0.01$, $p = 0.889$), and household income ($r = 0.15$, $p = 0.111$) were not significantly associated with placement outcomes.

## Discussion

To our knowledge, this study included the largest clinically-ascertained sample to date of ASD-affected children receiving EIBI with planned collection of language measures throughout treatment and recording of well-defined educational placement outcomes. From these data, three major findings emerged. First, generally positive, but highly variable, language trajectories were observed during EIBI. Second, three latent trajectories (sub-groups) were identified (albeit with different class sizes and slightly different trajectories across measures) that explained much of the individual-level variability in language trajectories. Finally, baseline and early treatment characteristics strongly predicted language trajectories and educational placement outcomes.

### Language Trajectories

Careful examination of individual language trajectories identified that most of the children with ASD had meaningful increases in language scores over the course of their EIBI treatment, even when compared to normative trajectories. For example, the present results identified that half of children receiving EIBI had large increases in total language (≥ 12 SS points) after twelve months of intervention. Even more impressive, nearly 3/4 of children had large improvements (> 10 SS points) in receptive (ROWPVT/PPVT) and expressive (EOWPVT/EVT) vocabulary after 1 year of EIBI. The large positive gains seen in this study, relative to the active arm of a recent three-site, randomized controlled trial of the Early Start Denver Model (Rogers et al. 2019), may reflect greater intensity of treatment, a variable

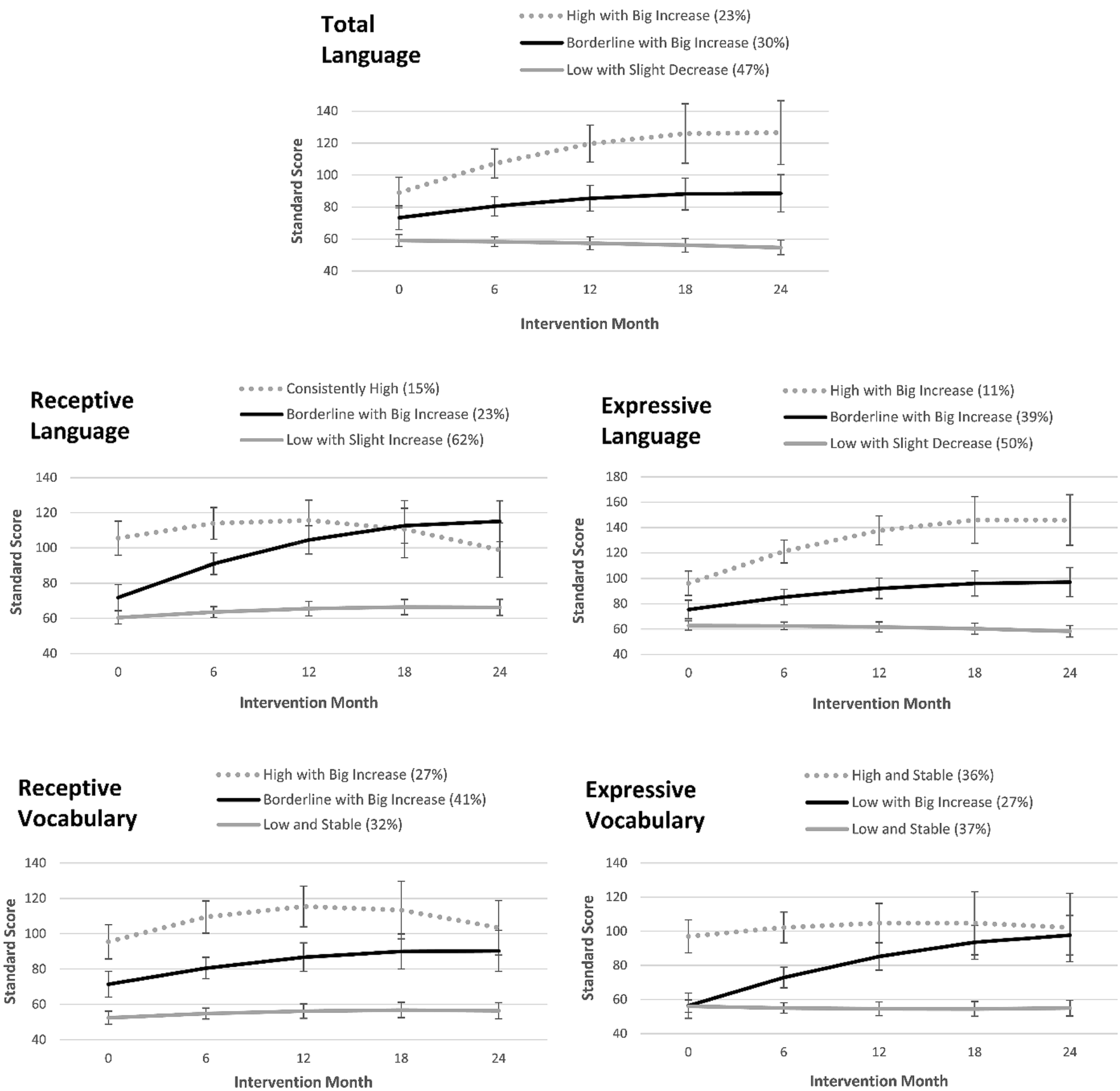


**Fig. 2** Growth mixture model results, separately for each language measure. Results are presented only for the first 24 months of intervention, as some sub-group trajectories become unstable due to attrition before 36 months

but generally longer treatment duration, and structured teaching of language as a primary intervention technique in the present study. Regardless, the sizeable proportion of children showing improved trajectories in these studies, and the magnitude of gains, suggests that EIBI may facilitate substantial language improvements in young children with ASD. Furthermore, these preliminary findings suggest that intensity, duration, and target emphasis across EIBI program delivery models may be important considerations for future controlled investigations. Although the current results are impressive, it is important to note that these results were not obtained using a randomized controlled trial, and therefore these gains could have been influenced by regression to the mean, parent and clinician expectations, practice effects, and/or natural development, resulting in language gains that may be equivalent to natural development.

Overall, most of the language gains occurred within the first 18 months of EIBI intervention, with the strongest initial gains (~ 11 SS points) occurring within the first 12 months. This implies that, for many children receiving

**Table 2** Latent growth class descriptions for each language measure

| | Total language (PLS) | | | |
|---|---|---|---|---|
| | Class 1<br>Low with slight decrease | Class 2<br>Borderline with big increase | Class 3<br>High with big increase | $X^2$/F (p) |
| N | 61 | 40 | 30 | |
| Age at Program Start (M, SD) | 3.6 (1.0) | 3.3 (0.9) | 2.9 (0.9) | 6.29 (.002) |
| Time in Program (M, SD) | 27.2 (12.3) | 25.4 (10.8) | 15.6 (7.1) | 11.90 (< .001) |
| Educational Support after Exit (n, %) | | | | |
| No Support/Mainstream | 0 (0.0%) | 4 (10.0%) | 18 (60.0%) | 75.20 (< .001) |
| Minimal Support | 11 (18.0%) | 11 (27.5%) | 5 (16.7%) | |
| Significant Support | 13 (21.3%) | 17 (42.5%) | 7 (23.3%) | |
| 1:1 IBI Support | 37 (60.7%) | 8 (20.0%) | 0 (0.0%) | |
| ASD Severity (n, %) | | | | |
| Mild | 10 (16.4%) | 20 (50.0%) | 22 (73.3%) | 29.79 (< .001) |
| Moderate/Severe | 51 (83.6%) | 20 (50.0%) | 8 (26.7%) | |
| Intellectual Disability (n, %) | 36 (59.1%) | 1 (2.5%) | 0 (0.0%) | 53.39 (< .001) |
| | Receptive Language (PLS) | | | |
| | Class 1<br>Low with Slight Increase | Class 2<br>Borderline with Big Increase | Class 3<br>High and Stable | $X^2$/F (p) |
| N | 85 | 27 | 19 | |
| Age at Program Start (M, SD) | 3.5 (1.0) | 2.9 (0.7) | 3.1 (1.0) | 5.37 (.006) |
| Time in Program (M, SD) | 27.4 (11.8) | 19.3 (8.9) | 15.2 (7.7) | 13.22 (< .001) |
| Educational Support after Exit (n, %) | | | | |
| No Support/Mainstream | 1 (4.5%) | 11 (40.7%) | 10 (52.6%) | 56.73 (< .001) |
| Minimal Support | 18 (21.2%) | 4 (14.8%) | 5 (26.3%) | |
| Significant Support | 23 (27.1%) | 10 (37.0%) | 4 (21.1%) | |
| 1:1 IBI Support | 43 (50.6%) | 2 (7.5%) | 0 (0.0%) | |
| ASD Severity (n, %) | | | | |
| Mild | 20 (23.5%) | 14 (51.9%) | 18 (94.7%) | 34.99 (< .001) |
| Moderate/Severe | 65 (76.5%) | 13 (48.1%) | 1 (5.3%) | |
| Intellectual Disability (n, %) | 36 (42.4%) | 1 (3.7%) | 0 (0.0%) | 23.85 (< .001) |
| | Expressive language (PLS) | | | |
| | Class 1<br>Low with slight decrease | Class 2<br>Borderline with big increase | Class 3<br>High with big increase | $X^2$/F (p) |
| N | 66 | 52 | 13 | |
| Age at Program Start (M, SD) | 3.6 (1.0) | 3.2 (1.0) | 2.8 (0.8) | 4.92 (.009) |
| Time in Program (M, SD) | 27.7 (12.2) | 21.8 (10.1) | 13.8 (6.6) | 10.41 (< .001) |
| Educational Support after Exit (n, %) | | | | |
| No Support/Mainstream | 0 (0.0%) | 15 (28.8%) | 7 (53.8%) | 57.80 (< .001) |
| Minimal Support | 13 (19.7%) | 11 (21.2%) | 3 (23.1%) | |
| Significant Support | 13 (19.7%) | 21 (40.4%) | 3 (23.1%) | |
| 1:1 IBI Support | 40 (60.6%) | 5 (9.6%) | 0 (0.0%) | |
| ASD Severity (n, %) | | | | |
| Mild | 13 (19.7%) | 28 (53.8%) | 11 (84.6%) | 26.34 (< .001) |
| Moderate/Severe | 53 (80.3%) | 24 (46.2%) | 2 (15.4%) | |
| Intellectual Disability (n, %) | 36 (54.5%) | 1 (1.9%) | 0 (0.0%) | 45.42 (< .001) |

**Table 2** (continued)

| | Receptive vocabulary (ROWPVT or PPVT) | | | |
|---|---|---|---|---|
| | Class 1<br>Low and stable | Class 2<br>Borderline with big increase | Class 3<br>High with big increase | $X^2$/F (p) |
| N | 43 | 49 | 37 | |
| Age at Program Start (M, SD) | 3.4 (1.0) | 3.4 (1.0) | 3.1 (1.0) | 1.31 (.274) |
| Time in Program (M, SD) | 27.2 (12.4) | 26.0 (11.7) | 17.9 (8.7) | 8.03 (.001) |
| Educational Support after Exit (n, %) | | | | |
| No Support/Mainstream | 0 (0.0%) | 6 (12.2%) | 16 (43.2%) | 47.35 (<.001) |
| Minimal Support | 7 (16.3%) | 13 (26.5%) | 7 (18.9%) | |
| Significant Support | 9 (20.9%) | 15 (30.6%) | 13 (35.1%) | |
| 1:1 IBI Support | 27 (62.8%) | 55 (30.6%) | 1 (2.7%) | |
| ASD Severity (n, %) | | | | |
| Mild | 7 (16.3%) | 17 (34.7%) | 28 (75.7%) | 30.20 (<.001) |
| Moderate/Severe | 36 (83.7%) | 32 (65.3%) | 9 (24.3%) | |
| Intellectual Disability (n, %) | 28 (65.1%) | 8 (16.3%) | 0 (0.0%) | 47.18 (<.001) |
| | Expressive vocabulary (EOWPVT or EVT) | | | |
| | Class 1 Low and stable | Class 2 Low with big increase | Class 3 High and stable | $X^2$/F (p) |
| N | 47 | 35 | 46 | |
| Age at Program Start (M, SD) | 3.4 (1.0) | 3.3 (1.0) | 3.3 (1.1) | 0.04 (.966) |
| Time in Program (M, SD) | 26.6 (13.4) | 20.2 (10.5) | 25.3 (9.7) | 3.92 (.022) |
| Educational Support after Exit (n, %) | | | | |
| No Support/Mainstream | 0 (0.0%) | 6 (17.1%) | 16 (34.8%) | 44.17 (<.001) |
| Minimal Support | 7 (14.9%) | 7 (20.0%) | 13 (28.3%) | |
| Significant Support | 10 (21.3%) | 13 (37.1%) | 14 (30.4%) | |
| 1:1 IBI Support | 30 (63.8%) | 9 (25.7%) | 3 (6.5%) | |
| ASD Severity (n, %) | | | | |
| Mild | 6 (12.8%) | 13 (37.1%) | 33 (71.7%) | 33.76 (<.001) |
| Moderate/Severe | 41 (87.2%) | 22 (62.9%) | 13 (28.3%) | |
| Intellectual Disability (n, %) | 24 (51.1%) | 7 (5.5%) | 4 (8.7%) | 22.31 (<.001) |

Time in Program represents the number of months from the child starting in the program to their exit

*IBI* Intensive Behavioral Intervention. Total, receptive, and expressive language N = 131. Receptive vocabulary N = 129. Expressive vocabulary N = 128

EIBI, the majority of gains can be expected early in treatment. While additional data and replication are needed, this finding suggests that some children can exit the program, receive decreased intensity of treatment, or shift the focus of treatment to other skills after 12–18 months of intervention, with plans for maintaining language gains while promoting integration into the next educational setting. Furthermore, other children may be able to shift from a heavy language focus to more intensive teaching of social skills, adaptive function, and/or other skills after acquiring these early improvements in language.

A non-trivial minority of children (~40% sub-group size across different measures) began with limited language/vocabulary skills and showed language gains comparable to, but not greater than, normative growth trajectories. It should be noted, however, that these children still showed *absolute increases* in language/vocabulary over time, even though they did not gain ground relative to norms. Across measures, these subgroups of children had lower initial language/vocabulary and verbal/non-verbal IQ scores, along with more severe ASD symptoms at baseline. Upon EIBI exit, these participants tended to require more intensive supports and were more likely to receive a comorbid intellectual disability diagnosis. This observation suggests that, in these subgroups with flat language/vocabulary trajectories, EIBI may promote slow, steady progress in language development and prevent loss of ground but that dramatic improvements are not generally observed. Maintaining a normative rate of growth in language development for these children may

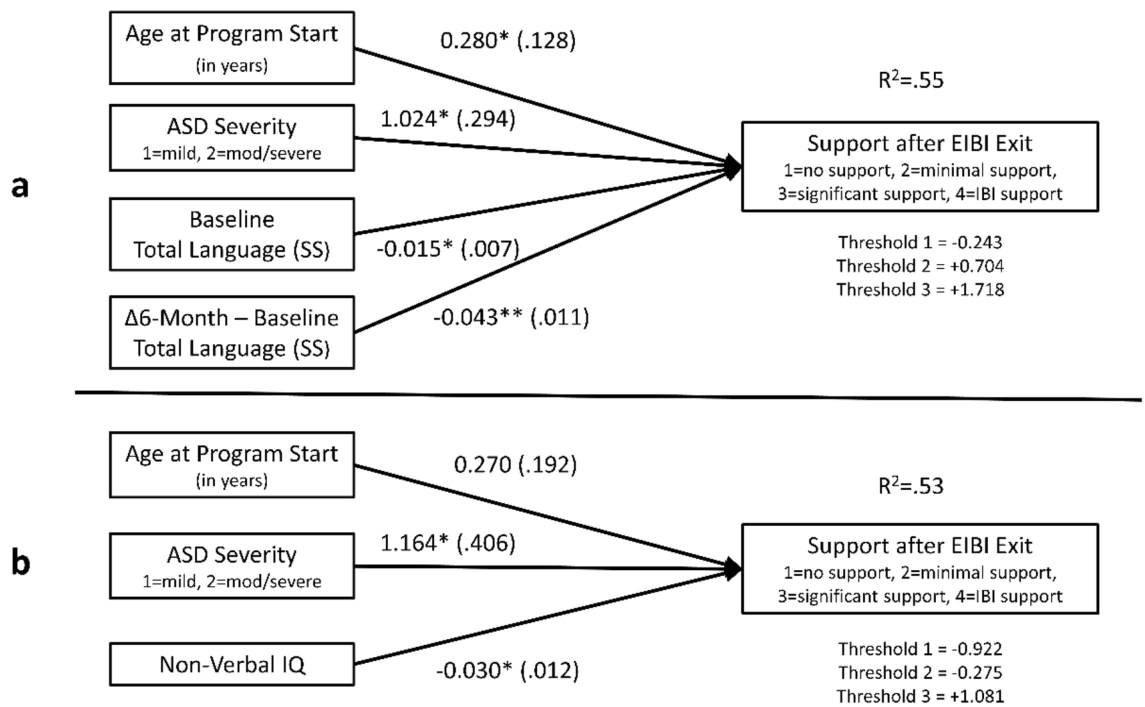


**Fig. 3** Probit regression model results (estimate and SE) for baseline plus 6-month changes in total language (**a**) and baseline characteristics only (**b**)

be considered an optimal outcome (Georgiades and Kasari 2018).

For each measure, there was another subgroup of children with borderline language scores at baseline, showed significant and rapid language gains, particularly with expressive language. This finding is consistent with the anecdotal clinical observation that some children with ASD, who have not received substantial intervention, do not speak because it is difficult for them to understand the social and functional value of language (Bradshaw et al. 2017; Tager-Flusberg et al. 2005). Early EIBI programs that utilize structured ABA stimulus–response pairings, with progression into naturalistic teaching methods when appropriate (Mohammadzaheri et al. 2014), may help these children to understand that language can be used to achieve desired outcomes, rather than employing aberrant behaviors that do not achieve the desired results. Thus, language, rather than tantrums or other dysfunctional behavior, becomes the primary mechanism for interacting with others.

## EIBI Exit Support

A large minority of children (38%) that exit EIBI required less intensive or no support in their future educational placements. Specifically, higher initial language scores, improved language early in treatment, earlier age of EIBI initiation, and mild ASD symptoms are indicative of more positive educational placement outcomes. Unfortunately, a similar proportion of children (34%) will need intensive behavioral intervention support. This highlights the wide array of outcomes from EIBI, the need to improve prediction of treatment response to maximize resources (Etscheidt 2003), and the imperative to tailor EIBI to produce optimal outcomes for all children (Georgiades and Kasari 2018).

By examining predictors of placement outcomes, we can begin to build more accurate models that inform and guide clinical practice. Combining age at EIBI start, baseline autism severity, initial language ability, and early changes in language accounted for more than 50% of the variance in placement outcomes. Age at EIBI start, ASD severity, and non-verbal IQ (as a group) achieved similar predictive value to the model using age at EIBI start, ASD severity, and baseline and early change in language. Individually, the predictive value of these factors is not surprising, as prior studies have found similar predictive relationships (Hardan et al. 2015; Howlin et al. 2009; Smith et al. 2015). However, taken together, these models, if externally replicated, could help clinicians in predicting outcomes for future children, allowing better intervention by therapists and life planning by families. Online Resource 7 provides additional information for using these models in future research. Given the small size of the test sample, it will be important to use larger samples coupled with train-test-validate approaches to increase the probability that models will replicate and generalize to new samples and to provide more precise estimates of model parameters.

### Changes in Vocabulary

Although less complex and predictive of EIBI outcome, one-word receptive (ROWPVT/PPVT) and expressive (EOWPVT/EVT) vocabulary tests showed interesting patterns. First, a substantial proportion of children with ASD receiving EIBI scored high on receptive vocabulary at program entry, identifying a strength prior to intervention that can be built on. This is consistent with research identifying early receptive vocabulary as an important predictor of receptive language growth over time (Yoder et al. 2015). Second, a subgroup of approximately 1/4 of participants who initially had low expressive vocabulary scores made substantial gains (≥ 18 SS points) in receptive and expressive one-word vocabulary across the first year of EIBI, improving rapidly from low to average scores. This pattern may reflect the use of discrete trial teaching and appropriate prompting of expressive language during regular EIBI therapy and SLP-conducted sessions. Directed attention toward a speaker has been shown to partially explain variances in receptive-expressive vocabulary in children with ASD, which could explain these rapid and substantial increases in language scores, providing further support for the inclusion of direct language instruction in EIBI (McDaniel et al. 2018). Future research focusing on the specific components of EIBI that may be generating these improvements will be important for individualized EIBI programming.

### Age of Intervention Start

Although prior studies have found inconsistent relationships between age at EIBI entry and outcomes (Howlin et al. 2009; Smith et al. 2000, 2015), data from the current study found early age of entry in EIBI to significantly predict a better language trajectory, especially for expressive language, and a better educational placement outcome. This highlights the importance of utilizing expressive language as an important initial treatment target, and supports the need to improve early ASD identification efforts to enable early and rapid enrollment in EIBI programs (Robins et al. 2016; Rogers et al. 2014). The present data suggest that every year of delay in entry to EIBI may result in ~ 6 fewer expressive language standard score points gained over 24 months relative to normative expectation and decreases the probability of a low intensity educational support placement by as much as 11%.

### Limitations and Future Directions

The primary limitation of this study was the use of a clinical convenience sample, obtained from one EIBI program using available clinical measures. Additionally, the program evolved over time to use more naturalistic teaching methods over time, data were collected over a long enrollment period, with variable length of treatment, and no comparison group was available. As a result, it is not possible to infer that EIBI alone generated language improvements or if these would have occurred naturally as part of development. It is also not possible to discern whether the language increases observed would have occurred if the children had only received small doses of outpatient speech/language therapy, as speech/language therapy was embedded into the program with close consultation between the speech/language therapist and the rest of the EIBI team. Furthermore, the limited representativeness of this clinical sample limits the likelihood that the findings generalize to other academic or community settings.

Another limitation is the decreased availability of language data over time, language patterns at later time points may be less accurately and precisely estimated. At minimum, future research should include planned collection of language data at each time point during the intervention period. Intriguingly, however, calendar year of entry into the EIBI program was associated with more positive expressive language and vocabulary trajectories, which is consistent with improvements in EIBI treatment over the last two decades that corresponds with a shift toward greater use of naturalistic teaching strategies (Schreibman et al. 2015).

A substantial proportion of individuals with significant developmental and cognitive delays were included in the current study to address the criticism regarding the limited variance of ASD functional levels in other studies. Thus, the current findings may not be reflective of more recent ASD cohorts where higher functioning presentations are better identified and included. However, this potential difference would suggest that the language improvements identified in the present study may be *under-estimated* relative to cohorts with a lower proportion of cognitively-delayed children. Future multi-site, clinical trials, including both academic and community sites and recruiting children with ASD across the full spectrum of severity and impairment, are needed to make real-world inferences about the efficacy of EIBI as a treatment approach and to identify predictors specific to treatment response. In addition, multi-site, longitudinal studies enrolling very large cohorts ($n$'$s > 200$) of children with ASD who are followed through the natural history of EIBI treatment to exit into educational placements are needed to replicate the present results and generalize them to other clinical and community settings. Specifically, the present study is limited by the observed values for predictors. Larger studies are likely to include a fuller range of predictor values and therefore, less biased and more efficient estimates for possible predictors of outcome. Inclusion of longitudinal information beyond language, such as ASD symptom severity, executive functioning, adaptive behavior, and potentially other cognitive or physiological measurements would also provide a richer view of EIBI efficacy.

Finally, the current study was limited by the use of multiple measures of ASD symptom severity that necessitated a dichotomous clinical rating rather than a poly-ordinal or continuous measurement. This may have attenuated the relationship between ASD symptom severity and outcome measures, and as a result, the observed significant relationships between lower ASD symptom severity and better language trajectories and placement outcomes are likely under-estimated. As previously mentioned, language measures were consistently collected during the first 18 months of treatment for most individuals who remained in the program, but dropped after 18 months. This could have resulted in a mis-estimation of the later segments of language trajectories and may account for the downturn in language growth from 24 to 36 months. However, given the use of growth models that are resistant to attrition, as they assume data are missing at random, and the fact that the present analyses indicated that attrition was largely due to exit from the program, estimates of later time points may be reasonable for initial descriptive purposes.

It is important that future investigations diligently collect a full range of longitudinal data to characterize and predict language development during EIBI treatment as well as educational placement outcomes. The consistent collection of ASD symptoms, cognitive ability, adaptive skills, and quality of life measures longitudinally could inform mechanisms of EIBI-based language improvement and determine whether EIBI treatment influences broader family functioning. This type of longitudinal collection would allow for simultaneous estimation of growth processes and their inter-relationships over time. In addition, more consistent and comprehensive data could be used to develop powerful predictive models of (a) improved ASD symptomology, (b) language, (c) cognitive ability, (d) social skills, and (e) educational placement outcomes, along with providing an early indication of target EIBI trajectories with the potential for individualized alterations to achieve the most optimal outcomes for each child with ASD.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s10803-021-04900-5.

**Acknowledgments** We are grateful to the participants in this study and thank our clinical research staff for their data collection and data reduction.

**Author Contributions** TWF had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. TWF was responsible for study concept and design, as well as the acquisition of data. TWF and EAY were responsible for statistical analysis and all authors were responsible for interpretation of the data. TWF drafted the initial manuscript and all authors were responsible for critical revision of the manuscript. All authors provided final approval of the manuscript.

**Funding** This study was made possible by a generous donation from the Stephan and Allison Cole Family Research Fund (to TWF). Dr. Frazier's participation was also supported by Autism Speaks and the Hartwell Foundation. The funding source had no role in the design, conduct, analysis, interpretation, or writing for the present manuscript.

## Compliance with Ethical Standards

**Conflict of interest** TWF has received funding or research support from, acted as a consultant to, received travel support from, and/or received a speaker's honorarium from Quadrant Biosciences, Impel NeuroPharma, F. Hoffmann-La Roche AG Pharmaceuticals, the Cole Family Research Fund, Simons Foundation, Ingalls Foundation, Forest Laboratories, Ecoeos, IntegraGen, Kugona LLC, Shire Development, Bristol-Myers Squibb, Roche Pharma, National Institutes of Health, and the Brain and Behavior Research Foundation and has an investor stake in Autism EYES LLC. EWK has received support from Kugona LLC. EAY has consulted with Pearson, Lundbeck, Janssen, Joe Startup Technologies, and Western Psychological Services about psychological assessment, and received royalties from the American Psychological Association and Guilford Press. AYH has received research funding from Forest Pharmaceuticals and Bristol Myers Squibb and is a consultant to IntegraGen. GWG has consulted to Autism Speaks.

**Ethical Approval** This project was a secondary analysis of clinical data collected as part of routine practice. The local institutional review board reviewed and approved the creation of a database containing the clinical data for future research use including the analyses conducted for this study.

**Informed Consent** As the data were collected as part of routine clinical practice, consent was not obtained but all patients at the site provide consent for clinical data collection as part of their clinic visit. The authors certify that the study was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments.

## References

Almirall, D., DiStefano, C., Chang, Y. C., Shire, S., Kaiser, A., Lu, X., et al. (2016). Longitudinal effects of adaptive interventions with a speech-generating device in minimally verbal children with ASD. *Journal of Clinical Child & Adolescent Psychology, 45*(4), 442–456. https://doi.org/10.1080/15374416.2016.1138407.

Bedford, R., Pickles, A., & Lord, C. (2016). Early gross motor skills predict the subsequent development of language in children with autism spectrum disorder. *Autism Research, 9*(9), 993–1001. https://doi.org/10.1002/aur.1587.

Ben-Itzchak, E., Watson, L. R., & Zachor, D. A. (2014). Cognitive ability is associated with different outcome trajectories in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44*(9), 2221–2229. https://doi.org/10.1007/s10803-014-2091-0.

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*(3), 287–303. https://doi.org/10.1016/j.ridd.2006.03.002.

Bentler, P. M. (1988). Comparative fit indexes in structural equation models. *Psychological Bulletin, 107,* 238–246.

Berkowitz, S. A., Traore, C. Y., Singer, D. E., & Atlas, S. J. (2015). Evaluating area-based socioeconomic status indicators for monitoring disparities within health care systems: Results from a primary care network. *Health Services Research, 50*(2), 398–417. https://doi.org/10.1111/1475-6773.12229.

Bolte, S., Poustka, F., & Constantino, J. N. (2008). Assessing autistic traits: Cross-cultural validation of the social responsiveness scale (SRS). *Autism Research, 1*(6), 354–363. https://doi.org/10.1002/aur.49.

Bopp, K. D., Mirenda, P., & Zumbo, B. D. (2009). Behavior predictors of language development over 2 years in children with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research, 52*(5), 1106–1120. https://doi.org/10.1044/1092-4388(2009/07-0262).

Bradshaw, J., Koegel, L. K., & Koegel, R. L. (2017). Improving functional language and social motivation with a parent-mediated intervention for toddlers with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 47*(8), 2443–2458. https://doi.org/10.1007/s10803-017-3155-8.

Chlebowski, C., Green, J. A., Barton, M. L., & Fein, D. (2010). Using the childhood autism rating scale to diagnose autism spectrum disorders. *Journal of Autism and Developmental Disorders, 40*(7), 787–799. https://doi.org/10.1007/s10803-009-0926-x.

Cohen, H., Amerine-Dickens, M., & Smith, T. (2006). Early intensive behavioral treatment: replication of the UCLA model in a community setting. *Journal of Developmental and Behavioral Pediatrics, 27*(2 Suppl), S145–S155.

Constantino, J. N., & Gruber, C. P. (2012). *The social responsiveness scale manual, second edition (SRS-2)*. Torrance: Western Psychological Services.

Cooper, J., Heron, T., & Heward, W. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River: Pearson/Merrill-Prentice Hall.

Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., et al. (2010). Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. *Pediatrics, 125*(1), e17-23. https://doi.org/10.1542/peds.2009-0958.

Dunlap, G. (1984). The influence of task variation and maintenance tasks on the learning and affect of autistic children. *Journal of Experimental Child Psychology, 37*(1), 41–64.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test* (4th ed.). Bloomington: NCS Pearson Inc.

Dunn, L. M., & Dunn, L. M. (1997). *Examiner's manual for the peabody picture vocabulary test* (3rd ed.). Circle Pines: American Guidance Service.

Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2002). Intensive behavioral treatment at school for 4- to 7-year-old children with autism. A 1-year comparison controlled study. *Behavior Modification*, *26*(1), 49–68. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11799654.

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child & Adolescent Psychology, 38*(3), 439–450. https://doi.org/10.1080/15374410902851739.

Esler, A. N., Bal, V. H., Guthrie, W., Wetherby, A., Ellis Weismer, S., & Lord, C. (2015). The autism diagnostic observation schedule, toddler module: Standardized severity scores. *Journal of Autism and Developmental Disorders, 45*(9), 2704–2720. https://doi.org/10.1007/s10803-015-2432-7.

Etscheidt, S. (2003). An analysis of legal hearings and cases related to individualized education programs for children with autism. *Research and Practice for Persons with Severe Disabilities, 28,* 51–69.

Fountain, C., Winter, A. S., & Bearman, P. S. (2012). Six developmental trajectories characterize children with autism. *Pediatrics, 129*(5), e1112-1120. https://doi.org/10.1542/peds.2011-1601.

Garfin, D. G., McCallon, D., & Cox, R. (1988). Validity and reliability of the Childhood Autism Rating Scale with autistic adolescents. *Journal of Autism and Developmental Disorders*, *18*(3), 367–378. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3170455.

Gartlehner, G., Hansen, R. A., Nissman, D., Lohr, K. N., & Carey, T. S. (2006). A simple and valid tool distinguished efficacy from effectiveness studies. *Journal of Clinical Epidemiology, 59*(10), 1040–1048. https://doi.org/10.1016/j.jclinepi.2006.01.011.

Georgiades, S., & Kasari, C. (2018). Reframing optimal outcomes in autism. *JAMA Pediatrics, 172*(8), 716–717. https://doi.org/10.1001/jamapediatrics.2018.1016.

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. https://doi.org/10.1007/s10803-008-0674-3.

Granpeesheh, D., Tarbox, J., & Dixon, D. R. (2009). Applied behavior analytic interventions for children with autism: a description and review of treatment research. *Annals of Clinical Psychiatry, 21*(3), 162–173.

Green, J., Pickles, A., Pasco, G., Bedford, R., Wan, M. W., Elsabbagh, M., et al. (2017). Randomised trial of a parent-mediated intervention for infants at high risk for autism: Longitudinal outcomes to age 3 years. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 58*(12), 1330–1340. https://doi.org/10.1111/jcpp.12728.

Hardan, A. Y., Gengoux, G. W., Berquist, K. L., Libove, R. A., Ardel, C. M., Phillips, J., et al. (2015). A randomized controlled trial of Pivotal Response Treatment Group for parents of children with autism. *Journal of Child Psychology and Psychiatry, 56*(8), 884–892. https://doi.org/10.1111/jcpp.12354.

Howard, J. S., Sparkman, C. R., Cohen, H. G., Green, G., & Stanislaw, H. (2005). A comparison of intensive behavior analytic and eclectic treatments for young children with autism. *Research in Developmental Disabilities, 26*(4), 359–383. https://doi.org/10.1016/j.ridd.2004.09.005.

Howlin, P., Magiati, I., & Charman, T. (2009). Systematic review of early intensive behavior interventions for children with autism. *American Journal of Intellectual and Developmental Disabilities, 114*(1), 23–41.

IBM Corp. (2018). IBM SPSS statistics for windows. *Version 26.0*. New York: IBM Corp.

Ingersoll, B. (2010). Pilot randomized controlled trial of Reciprocal Imitation Training for teaching elicited and spontaneous imitation to children with autism. *Journal of Autism and Developmental Disorders, 40*(9), 1154–1160. https://doi.org/10.1007/s10803-010-0966-2.

Ingersoll, B., Schreibman, L., & Stahmer, A. (2001). Brief report: Differential treatment outcomes for children with autistic spectrum disorder based on level of peer social avoidance. *Journal of Autism and Developmental Disorders, 31*(3), 343–349.

Kasari, C., Paparella, T., Freeman, S., & Jahromi, L. B. (2008). Language outcome in autism: Randomized comparison of joint attention and play interventions. *Journal of*

*Consulting and Clinical Psychology, 76*(1), 125–137. https://doi.org/10.1037/0022-006X.76.1.125.

Koegel, L., Koegel, J., Harrower, J., & Carter, C. (1999). Pivotal response intervention I: Overview of approach. *The Journal of the Association for Persons with Severe Handicaps, 24*(3), 174–185.

Koegel, R. L., O'Dell, M. C., & Koegel, L. K. (1987). A natural language teaching paradigm for nonverbal autistic children. *Journal of Autism and Developmental Disorders, 17*(2), 187–200.

Koegel, R. L., & Williams, J. A. (1980). Direct versus indirect response-reinforcer relationships in teaching autistic children. *Journal of Abnormal Child Psychology, 8*(4), 537–547.

Landa, R. (2007). Early communication development and intervention for children with autism. *Mental Retardation and Developmental Disabilities Research, 13*(1), 16–25. https://doi.org/10.1002/mrdd.20134.

Landa, R. J., & Kalb, L. G. (2012). Long-term outcomes of toddlers with autism spectrum disorders exposed to short-term intervention. *Pediatrics, 130*(Suppl 2), S186–S190. https://doi.org/10.1542/peds.2012-0900Q.

Lord, C., Luyster, R., Guthrie, W., & Pickles, A. (2012a). Patterns of developmental trajectories in toddlers with autism spectrum disorder. *Journal of Consulting and Clinical Psychology, 80*(3), 477–489. https://doi.org/10.1037/a0027214.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2002). *Autism diagnostic observation schedule: ADOS manual*. Torrance: Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012b). *Autism diagnostic observation schedule, second edition (ADOS-2) manual (part 1): Modules 1–4*. Torrance: Western Psychological Services.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55,* 3–9.

Luyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., et al. (2009). The Autism Diagnostic Observation Schedule-toddler module: A new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(9), 1305–1320. https://doi.org/10.1007/s10803-009-0746-z.

Mahoney, G., & Perales, F. (2003). Relationship-focused intervention to enhance the social-emotional functioning of young children with autism spectrum disorders. *Topics in Early Childhood Special Education, 23*(2), 74–86.

McDaniel, J., Yoder, P., Woynaroski, T., & Watson, L. R. (2018). Predicting receptive-expressive vocabulary discrepancies in preschool children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research, 61*(6), 1426–1439. https://doi.org/10.1044/2018_JSLHR-L-17-0101.

McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal of Mental Retardation*, *97*(4), 359–372; discussion 373–391. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8427693.

Mohammadzaheri, F., Koegel, L. K., Rezaee, M., & Rafiee, S. M. (2014). A randomized clinical trial comparison between pivotal response treatment (PRT) and structured applied behavior analysis (ABA) intervention for children with autism. *Journal of Autism and Developmental Disorders, 44*(11), 2769–2777. https://doi.org/10.1007/s10803-014-2137-3.

Molloy, C. A., Murray, D. S., Akers, R., Mitchell, T., & Manning-Courtney, P. (2011). Use of the autism diagnostic observation schedule (ADOS) in a clinical setting. *Autism, 15*(2), 143–162. https://doi.org/10.1177/1362361310379241.

Mullen, E. M. (1995). *Mullen scales of early learning*. Circle Pines: American Guidance Service Inc.

Mundy, P., & Crowson, M. (1997). Joint attention and early social communication: implications for research on intervention with autism. *Journal of Autism and Developmental Disorders*, *27*(6), 653–676. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9455727.

Mundy, P., Sigman, M., & Kasari, C. (1990). A longitudinal study of joint attention and language development in autistic children. *Journal of Autism and Developmental Disorders, 20*(1), 115–128.

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide. Seventh Edition*. (Seventh ed.). Muthén & Muthén.

National Research Council, & Division of Behavioral and Social Sciences Education. (2001). *Educating children with autism*. Washington, D.C.: National Academy Press.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14*(4), 535–569.

Ospina, M. B., Krebs Seida, J., Clark, B., Karkhaneh, M., Hartling, L., Tjosvold, L., et al. (2008). Behavioural and developmental interventions for autism spectrum disorder: A clinical systematic review. *PLoS ONE, 3*(11), e3755. https://doi.org/10.1371/journal.pone.0003755.

Peters-Scheffer, N., Didden, R., Korzilius, H., & Matson, J. (2012). Cost comparison of early intensive behavioral intervention and treatment as usual for children with autism spectrum disorder in The Netherlands. *Research in Developmental Disabilities, 33*(6), 1763–1772. https://doi.org/10.1016/j.ridd.2012.04.006.

Peters-Scheffer, N., Didden, R., Korzilius, H., & Sturmey, P. (2011). A meta-analytic study on the effectiveness of comprehensive ABA-based early intervention programs for children with Autism Spectrum Disorders. *Research in Autism Spectrum Disorders, 5*(1), 60–69.

Raftery, A. (1995). Bayesian model selection in social research. *Sociological Methodology, 25,* 111–163.

Risi, S., Lord, C., Gotham, K., Corsello, C., Chrysler, C., Szatmari, P., et al. (2006). Combining information from multiple sources in the diagnosis of autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 45*(9), 1094–1103.

Robins, D. L., Adamson, L. B., Barton, M., Connell, J. E., Jr., Dumont-Mathieu, T., Dworkin, P. H., et al. (2016). Universal autism screening for toddlers: Recommendations at odds. *Journal of Autism and Developmental Disorders, 46*(5), 1880–1882. https://doi.org/10.1007/s10803-016-2697-5.

Rogers, S. J., Estes, A., Lord, C., Munson, J., Rocha, M., Winter, J., et al. (2019). A multisite randomized controlled two-phase trial of the early start Denver model compared to treatment as usual. *Journal of the American Academy of Child and Adolescent Psychiatry*. https://doi.org/10.1016/j.jaac.2019.01.004.

Rogers, S. J., Vismara, L., Wagner, A. L., McCormick, C., Young, G., & Ozonoff, S. (2014). Autism treatment in the first year of life: A pilot study of infant start, a parent-implemented intervention for symptomatic infants. *Journal of Autism and Developmental Disorders, 44*(12), 2981–2995. https://doi.org/10.1007/s10803-014-2202-y.

Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal of Mental Retardation, 110*(6), 417–438. https://doi.org/10.1352/0895-8017(2005)110[417:IBTFCW]2.0.CO;2.

Schopler, E., Van Bourgondien, M., Wellman, G., & Love, S. (2010). *Childhood autism rating scale* (2nd ed.). Torrance: Western Psychological Services.

Schreibman, L. (2005). *The science and fiction of autism*. Cambridge: Harvard University Press.

Schreibman, L., Dawson, G., Stahmer, A. C., Landa, R., Rogers, S. J., McGee, G. G., et al. (2015). Naturalistic developmental behavioral interventions: Empirically validated treatments for autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(8), 2411–2428. https://doi.org/10.1007/s10803-015-2407-8.

Sherer, M. R., & Schreibman, L. (2005). Individual behavioral profiles and predictors of treatment effectiveness for children with autism. *Journal of Consulting and Clinical Psychology, 73*(3), 525–538. https://doi.org/10.1037/0022-006X.73.3.525.

Smith, T., Groen, A. D., & Wynn, J. W. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal of Mental Retardation, 105,* 269–285.

Smith, T., Klorman, R., & Mruzek, D. W. (2015). Predicting outcome of community-based early intensive behavioral intervention for children with autism. *Journal of Abnormal Child Psychology, 43*(7), 1271–1282. https://doi.org/10.1007/s10802-015-0002-2.

Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism. In F. Volkmar (Ed.), *Handbook of autism and pervasive developmental disorders*. Hoboken: Wiley.

Ventola, P. E., Kleinman, J., Pandey, J., Barton, M., Allen, S., Green, J., et al. (2006a). Agreement among four diagnostic instruments for autism spectrum disorders in toddlers. *Journal of Autism and Developmental Disorders, 36,* 839–847.

Ventola, P. E., Kleinman, J., Pandey, J., Barton, M., Allen, S., Green, J., et al. (2006b). Agreement among four diagnostic instruments for autism spectrum disorders in toddlers. *Journal of Autism and Developmental Disorders, 36*(7), 839–847. https://doi.org/10.1007/s10803-006-0128-8.

Vivanti, G., Prior, M., Williams, K., & Dissanayake, C. (2014). Predictors of outcomes in autism early intervention: Why don't we know more? *Frontiers in Pediatrics, 2,* 58. https://doi.org/10.3389/fped.2014.00058.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gotzsche, P. C., Vandenbroucke, J. P., & Initiative, S. (2007). The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Epidemiology, 18*(6), 800–804. https://doi.org/10.1097/EDE.0b013e3181577654.

Warren, Z., McPheeters, M. L., Sathe, N., Foss-Feig, J. H., Glasser, A., & Veenstra-Vanderweele, J. (2011). A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics, 127*(5), e1303-1311. https://doi.org/10.1542/peds.2011-0426.

Wechsler, D. (2002). *The Wechsler preschool and primary scale of intelligence, third edition (WPPSI-III)*. San Antonio: The Psychological Corporation.

Wechsler, D. (2012). *Wechsler preschool and primary scale of intelligence* (4th ed.). San Antonio: The Psychological Corporation.

Williams, K. (1997). *Expressive vocabulary test*. London: Pearson Education Inc.

Williams, K. T. (2007). *Expressive vocabulary test, second edition (EVT-2)*. Bloomington: NCS Pearson Inc.

Yoder, P., Watson, L. R., & Lambert, W. (2015). Value-added predictors of expressive and receptive language growth in initially nonverbal preschoolers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 45*(5), 1254–1270. https://doi.org/10.1007/s10803-014-2286-4.

Zimmerman, I. L., Steiner, V. G., & Pond, E. (2011). *Preschool language scales-fifth edition (PLS-5)*. London: Pearson.