Title

Investigation of Early Symptom Presentation in Children Under Age Three with Risk for Autism

Author
Bacon, Elizabeth Catherine

Publication Date
2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Investigation of Early Symptom Presentation in Children Under Age Three with Risk for Autism

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Psychology

by

Elizabeth Catherine Bacon

Committee in charge:

Professor Laura Schreibman, Chair
Professor Leslie Carver
Professor Eric Courchesne
Professor Karen Dobkins
Professor Karen Pierce
Professor Aubyn Stahmer

2014

The Dissertation of Elizabeth Catherine Bacon is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2014

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank Laura Schreibman for her mentorship, guidance, and support. Her commitment to conducting innovative research, upholding ethical responsibilities, and improving the lives of families of children with autism has provided me an incredible education. I am grateful for her confidence in me and the motivation she has provided me to pursue my passions. She has truly been an inspirational role model and I am thankful to have had the opportunity to learn from her. I would also like to thank Aubyn Stahmer for her advice, encouragement, and feedback throughout my graduate school experience. Her enthusiasm for research and teaching has been greatly influential on my growth as a researcher. I am grateful to Karen Pierce and Eric Courchesne for welcoming me into their lab to conduct this project and graciously offering mentorship and feedback along the way. I would also like to thank Mark Appelbaum, Julian Parris, and Kevin Smith for their patience and assistance navigating the statistics conducted in this project. Many thanks to the staff of the Autism Intervention Research Program and the Autism Center of Excellence for their daily support, feedback, and much needed humor throughout my graduate schooling. I would also like to acknowledge the many undergraduate research assistants that shared a genuine passion for this project and dedicated many hours to coding countless videos. A special thank you to the families and children that participated in the research and taught me so much about autism. I dedicate this work to my family, and especially to Dan Bacon, for encouraging me daily, and for an incredible amount of patience and understanding throughout the completion of this project.

VITA

EDUCATION

Ph.D. Psychology, University of California, San Diego, 2014

M.A. Psychology, University of California, San Diego, 2011

B.A. Psychology (Speech and Hearing Sciences minor), Arizona State University, 2009

CERTIFICATION

Board Certified Behavior Analyst (certificate number: 1-12-1027)

RESEARCH EXPERIENCE

2009-present **Graduate Student Researcher**
Conduct independent and collaborative research.
Advisor: Dr. Laura Schreibman, University of California, San Diego.

2013-present **Research Consultant**
Conduct assessments of teachers and students participating in a Classroom Pivotal Response Teaching study.
Advisor: Dr. Aubyn Stahmer, Rady Children's Hospital, San Diego, CA.

2012 **Abstract Reviewer**
Reviewed scientific abstracts for acceptance as poster or oral presentations at a conference.
Organization: 2013 International Meeting for Autism Research.

2012 **Article Reviewer**
Reviewed rigor of journal articles concerning interventions for autism spectrum disorders to identify evidence-based practices within the field.
Organization: Evidence-Base Practice Update Workshop, National Professional Development Center on ASD.

2012 **Article Reviewer**
Reviewed scientific merit of empirical studies regarding interventions for individuals with autism spectrum disorders.
Organization: National Standard Project II, National Autism Center.

2011 **Graduate Student Researcher**
Conducted behavioral assessments and recorded electromyography measurements of facial expressions in children with autism spectrum disorder.
Advisor: Dr. Marian Bartlett, University of California, San Diego.

2007-2009 **Undergraduate Research Assistant**
Conducted and coded behavioral assessments of children.
Advisors: Drs. Nancy Eisenberg and Tracy Spinrad, Arizona State University.

2008-2009 **Undergraduate Research Seminar**
Created and implemented a research project, and presented a poster to faculty members.
Advisors: Drs. Tamiko Azuma and Sid Bacon, Arizona State University.

CLINICAL EXPERIENCE

2008-2009 **Senior Habilitator**
Implemented interventions based on applied behavior analysis (ABA) with children with autism spectrum disorder, monitored and reported on children's progress, and assisted with new employee's training.
Advisor: Dr. Daniel Openden, Southwest Autism Research and Resource Center, Phoenix, AZ.

2007-2008 **Junior Habilitator**
Implemented interventions based on applied behavior analysis (ABA) with children with autism spectrum disorder.
Advisor: Dr. Daniel Openden, Southwest Autism Research and Resource Center, Phoenix, AZ.

2006-2007 **In-Home Therapist**
Implemented interventions based on applied behavior analysis (ABA) with children with autism spectrum disorder.
Agency: Honoring and Optimizing Potential in Everyone Group, Phoenix, AZ.

SPECIALIZED TRAINING

2012 **Classroom Pivotal Response Teaching**

TEACHING EXPERIENCE

2009-present **Research Supervisor**
Train and supervise undergraduate research assistants.
UCSD Department of Psychology, La Jolla, CA.

2009-present **Teaching Assistant**
UCSD Department of Psychology, La Jolla, CA.
Abnormal Psychology: 2011
Adolescence: 2013, 2014
Behavioral Foundations: 2013
Behavior Modification: 2009, 2010, 2011, 2013
Eating Disorders: 2014
Introduction to Clinical Psychology: 2013
Introduction to Principles of Behavior: 2012
Learning and Motivation: 2011, 2014
Operant Psychology Lab: 2011, 2014
Psychological Disorders of Childhood: 2013

2009-present **Guest Lecturer**
UCSD Department of Psychology, La Jolla, CA.
Topics:
Autism in Adolescence
Autism and Behavioral Therapy
Discrete Trial Training
Feeding Issues in Autism
Functional Analysis
Pivotal Response Training
Research Methods
Speech Disorders
Test Anxiety

PEER-REVIEWED PUBLICATIONS

Bacon, E. C., Dufek, S. A., Schreibman, L., Stahmer, A. C., Pierce, K., & Courchesne, E. (2014). Measuring outcomes in an early intervention program for toddlers with autism spectrum disorder: Use of a curriculum-based assessment. *Autism Research and Treatment, 2014,* 1-9. doi:10.1155/2014/964704

INVITED TALKS AND WORKSHOPS

Bacon, E. C. (2014, April). *Investigation of early symptom presentation in children with risk for autism under age three.* Presented at the Autism Seminar Series, Rady Children's Hospital, San Diego, CA.

Bacon, E. C. & Dufek, S. A. (2012, November). *Pivotal response training.* Workshop for practitioners working with ASD populations, University of California, San Diego, La Jolla, CA.

PRESENTATIONS AT PROFESSIONAL MEETINGS

Bacon, E. C., Chen, M., Schreibman, L., Stahmer, A. C., Carter, C., Courchesne, E., & Pierce, K. (2014, May). *Reduced curiosity and exploration as an early warning sign of ASD.* Paper presented at the 13th Annual International Meeting for Autism Research, Atlanta, GA.

Worcester, E. C., Schreibman, L., & Stahmer, A. C. (2013, May*). Evaluating social communicative behaviors across treatment settings for children with autism.* Poster presented at the 12th Annual International Meeting for Autism Research, San Sebastian, Spain.

Worcester, E. C., Schreibman, L., & Stahmer, A. C. (2013, May). *Evaluating social communicative behaviors across treatment settings for children with autism*. Poster presented at the Autism Speaks Toddler Treatment Network Conference, San Sebastian, Spain.

Worcester, E. C., Dufek, S. A., Schreibman, L., Stahmer, A. C., Courchesne, E., & Pierce, K. (2012, May). *Evaluation of early intervention outcome in children with risk for ASD.* Poster presented at the 11th Annual International Meeting for Autism Research, Toronto, Ontario, Canada.

Worcester, E. C., Dufek, S. A., Schreibman, L., Stahmer, A. C., Courchesne, E., & Pierce, K. (2012, May). *Evaluation of early intervention outcome in toddlers with risk for ASD.* Poster presented at the Autism Speaks Toddler Treatment Network Conference, Toronto, Ontario, Canada.

Dufek, S. A., **Worcester, E. C.,** Schreibman, L., Stahmer, A. C., Courchesne, E., & Pierce, K. (2012, May). *Stress in parents of toddlers with risk for ASD in an early intervention program.* Paper presented at the Autism Speaks Toddler Treatment Network Conference, Toronto, Ontario, Canada.

Schreibman, L., Stahmer, A. C., **Worcester, E. C.,** Gutierrez, R., Pierce, K., & Courchesne, E. (2012, February). *Examining curriculum-based assessment as a measure of early intervention outcome.* Poster presented at the 30th Annual Western Regional Conference for the California Association of Behavior Analysis, Garden Grove, CA.

Schreibman, L., Stahmer, A. C., **Worcester, E. C.,** Gutierrez, R., Pierce, K., & Courchesne, E. (2011, May). *Examining curriculum-based assessment as a* *measure of early intervention outcome.* Poster presented at the 10th Annual International Meeting for Autism Research, San Diego, CA.

HONORS AND AWARDS

2014 **University of California Graduate Research Advocacy Representative**
Selected as a graduate student representative for the University of California, San Diego and advocated for graduate education and research with legislators at the California State Capitol.

2013, 2014 **IMFAR Student Travel Award**

2012-2014 **Norman Anderson Student Travel and Research Award**

MEMBERSHIPS AND AFFILIATIONS

Association for Applied Behavior Analysis International
California Association of Behavior Analysis
International Society for Autism Research

ABSTRACT OF THE DISSERTATION

Investigation of Early Symptom Presentation in Children Under Age Three with Risk for Autism

by

Elizabeth Catherine Bacon

Doctor of Philosophy in Psychology

University of California, San Diego, 2014

Professor Laura Schreibman, Chair

Given the rise in frequency of autism spectrum disorder (ASD) diagnoses and the importance of early diagnosis for access to intervention services, there has been a push for early identification. Several early markers of ASD have been identified, however, these markers have largely been established in baby siblings of children with ASD, and the extent of generalization to a non-sibling population is unknown. Additionally,

diagnostic stability at young ages is somewhat variable, pointing to a need for further research to improve early identification processes.

The current project studied 299 toddlers, including early-identified cases of ASD (identified at-risk for ASD at initial and subsequent evaluations), late-identified cases of ASD (initially considered nonspectrum, then identified at-risk for ASD at a subsequent evaluation), children with language delay, and typically developing children. Every six to twelve months children participated in a battery of assessments including developmental and diagnostic tests, eye-tracking, an exploration task, and a free play observation. Aims were to: 1) assess whether early markers of ASD identified in the baby sibling literature were replicable within the current sample, 2) identify early behavioral markers within late-identified ASD cases, and 3) analyze trajectories of development until age three.

A reduction in social-communication skills was seen in both ASD groups at initial assessments, including increased preference for non-social stimuli, increased stereotypic play, and reduced exploration, use of gestures, social vocalizations, and social referencing. However, the late-identified cases of ASD were difficult to differentiate from children with language delay. ASD groups showed different developmental trajectories; the early-identified cases showed more impairment initially, but showed greater improvement over time than the late-identified ASD group.

Many of the early behavioral markers identified in the baby sibling literature were replicated. Increased preference for geometric stimuli, increased stereotypic play, and reduced exploration and social referencing indicated the highest risk for ASD and may be useful for identifying ASD in toddlers. Additionally, different developmental trajectories between early and late-identified children with ASD point to possible subgroups of ASD.

These data provide important information regarding early development of ASD and provide direction for future refinement of the early detection process.

INTRODUCTION

The full clinical phenotype of any developmental disorder is rarely present at birth, but rather, emerges across a span of time. This is especially true in the case of autism spectrum disorders (ASD). Researchers are continually investigating the emergence of identifiable differences between typically developing children and children who develop ASD, and are striving to identify children with ASD at the earliest ages possible. For example, Ozonoff and colleagues (2010) followed infants from 6 to 36 months, and identified differences between children that did and did not develop ASD. Over that period of time a slow decline in social skills was observed in children who eventually developed ASD, pointing to possible behavioral markers before age three. Ozonoff and colleagues found a decline in social engagement, social smiling, and looking toward faces in the group of children that developed ASD, as well as fewer vocalizations as compared to the typically developing children. These differences were not significantly different until 12-18 months of age, and the two groups were indistinguishable at younger ages. Several other studies have found a similar pattern with an emergence of symptoms starting at 12 months of age. Case studies examining the development of ASD in infants have noted motor atypicalities, stereotyped behaviors, atypical responses to sensory input and trouble regulating behavior, as well as several social communication deficits including; poor eye-contact, a lack of facial expressions, and a lack of verbal communication beginning around one year of age (Bryson, et al., 2007; Dawson, Osterling, Meltzoff, & Khul, 2000). Several prospective studies have demonstrated that a decline of skills from 6 to 36 months on the Mullen Scales of Early Learning (Mullen, 1995), which measures receptive language, expressive language, visual reception, fine motor and gross motor skills, is associated with the development of an ASD (Bryson, et al., 2007; Landa & Garret-Mayer, 2006; Landa, Gross, Stuart, & Bauman, 2012; Zwaigenbaum et al., 2005). Other early behavioral indicators consistently found in infants who later develop an ASD include several communication deficits such as a lack of response to one's name being called and a lack of gestures (Nadig at el., 2007; Talbott, Nelson, & Tager-Flusberg, 2013), and social deficits including a lack of social smiling, interest, and affect (Macari et al., 2012; Zwaigenbaum et al., 2005). Differences in visual attention have also been documented including atypical eye-contact, difficulties shifting visual attention, and atypicalities in social referencing and joint attention (Chawarska, et al. 2014; Cornew, Dobkins, Akshoomoff, McCleery, & Carver, 2012; Ibanez, Grantz, & Messinger, 2013; Macari et al., 2012; Rozga, et al. 2011; Zwaigenbaum et al., 2005). Lastly, differences in play have also been identified, such as lower activity level or passive behavior, poor imitation, lack of imaginative play, and repetitive behaviors (Chawarska, et al. 2014; Christensen et al., 2010; Macari, et al., 2012; Young et al., 2011; Zwaigenbaum, et al., 2005). In sum, many early behaviors have been associated with the development of ASD. However, there is not a clear behavioral marker that is seen in all cases of ASD, rather toddlers show various combinations of early behaviors associated with ASD (Tager-Flusberg, 2010).

Population Limitations

One limitation of this body of work is the heavy use of baby sibling populations and the lack of replication in a non-sibling cohort of children from the general population. Siblings of children with ASD offer an ease of study, as families are already educated about ASD and are arguably more motivated to track their child for possible delays as up to 18% of siblings go on to develop an ASD themselves (Ozonoff et al., 2011). However, it has been demonstrated that typically developing siblings of children with ASD present differently than typically developing children without a sibling with ASD (e.g. Georgiades et al., 2013). It is also possible that multiplex cases, or families with multiple children with ASD, may demonstrate a slightly different symptom presentation or trajectory of development than simplex cases of ASD, or families with only a single child with ASD. Genetic differences have been found between simplex and multiplex cases of ASD, with simplex cases showing an increased frequency of do novo copy number variations, or a higher frequency of spontaneous mutation of the genome, than multiplex cases (Sebat et al., 2007). Due to these differences behavioral markers of ASD need to be replicated using a more diverse sample. Additionally, since only a small percentage of baby siblings go on to develop an ASD themselves, much of the previous research has included only small sample sizes of children with an ASD, with most of the sample being used as a control group of siblings that do not develop an ASD. Many of these studies had fewer than 25 participants who developed ASD (Bryson, et al., 2007; Christensen et al., 2010; Landa & Garret-Mayer, 2006; Ibanez, Grantz, & Messinger, 2014; Macari, et al., 2012; Nadig et al., 2007; Talbott, Nelson, & Tager-Flusberg, 2013; Young et al., 2011; Zwaigenbaum et al., 2005). Therefore, a larger investigation of symptom presentation over time with the general population at this young age is imperative to establish a more comprehensive understanding of symptom development, increase the specificity of the behavioral markers of ASD that have been established, and increase generalization of these findings to a wider population.

Methods for Identifying Behaviors Associated with ASD

Successful methods for identifying observable behavioral differences between young children with ASD and typically developing peers have included eye-tracking, observation of exploration, and observation of play. For instance, Pierce, Conant, Hazin, Stoner, and Desmond (2011) studied visual preferences for social or non-social stimuli measured through a geometric preference eye-tracking test in children with an ASD, children with a developmental delay, and typically developing children between 12 and 42 months of age. Non-social stimuli (a video of moving geometric shapes) and social stimuli (a video of children engaging in aerobics) were presented side-by-side and preference for one type of stimuli was measured through the amount of time spent looking at each stimulus. They found a preference for geometric stimuli within the ASD group only. Any child that spent 69% or more of the time looking at geometric stimuli were exclusively children with ASD. Forty percent of the children with ASD showed this preference for geometric stimuli, indicating that not all children with ASD showed a preference for non-social stimuli; however, when children did show this preference, it was highly indicative of ASD rather than typical development or even other developmental delays. This type of eye-tracking provides a novel method for examining potential markers of ASD through visual preference.

A reduction in the exploration of the environment has also been associated with ASD. Pierce and Courchesne (2001) examined how children with ASD and typically developing children between ages three and eight years explored their environment. Several items and toys (Slinky®, dolls, string, etc.) were placed about an observation room with some in containers, making them more difficult to access, and some simply on the floor. Compared to the typically developing children, the children with ASD spent less time exploring items, explored fewer containers, and exhibited more stereotyped behaviors during the observation. Ozonoff et al. (2008) and Christensen et al. (2010) also examined explorative behaviors in baby siblings and typical controls between 12 and 18 months of age. The children that went on to develop an ASD engaged in more stereotyped behaviors and repetitive play, such as spinning objects, rotating objects, and unusual visual exploration of objects compared to the children that did not develop an ASD. However, these studies focused on younger siblings of children with ASD, and very few of the participants went on to receive a diagnosis of ASD, so the results are again difficult to translate to a more comprehensive sample. However, explorative behaviors also offer a potential avenue for exploring early markers for ASD in infants and toddlers.

Lastly, differences in play behavior have been documented in toddlers with ASD. For example, Wan et al. (2013) examined differences in play between infant baby siblings and typical controls during play sessions between the child and parent. After watching a video of parent and child playing, an examiner completed a questionnaire rating the levels of social-communication behaviors. By age three, approximately one-third of the baby siblings had developed an ASD, and these infants showed lower interaction and attentiveness to their parent, and less positive affect at 12 months as compared to the infants who did not receive an ASD diagnosis. All behaviors were rated using a Likert scale, rather than quantifying the actual rates of behavior for a more detailed analysis of differences in behavior. Wan and colleagues also focused on siblings of children with autism, rather than a general population sample. However, behavioral observation of play with a parent offers a way to capture differences in social behavior within a more natural environment, rather than through a more traditional structured assessment. Eye-tracking and behavioral observations provide unique avenues for measuring diverse behavioral differences associated with ASD, and may be particularly useful for assessing very young populations.

Developmental Trajectories

Researchers have identified differing patterns of learning, and have begun to attempt to classify different developmental trajectories. Researchers have found variable rates of skill acquisition among children with ASD (Weiss, 1999) and have been able to identify subgroups of children with ASD based on functioning level that are predictive of long-term outcomes (Stevens et al., 2000). Fountain, Winter, and Bearman (2012) established multiple trajectories of social communication development including categories of low to high development, wherein children beginning at higher developmental levels showed more improvement and maintained a higher level of functioning than those more severely affected. Fountain and colleagues also identified one group of children with a very different developmental trajectory where children started off more severely affected and experienced extremely rapid gains in skills, resulting in an accelerated trajectory of learning. Regressive trajectories have also been documented, where children initially displayed average developmental functioning, and then began to show a decline in development as they age (Landa, Gross, Stuart, & Bauman, 2012; Shumway et al., 2011). This phenomenon is referred to as developmental worsening, late-onset ASD, or regression; all referring to a process of children becoming more developmentally delayed over time. These differences in the display of symptoms over time may be contributing to the difficulty of maintaining stable diagnoses over time, and the wide range of ages at which children receive their diagnoses. It is vital to better understand developmental trajectories to inform future research, diagnostics, and treatment methodologies.

Diagnostic Stability

Related to early symptom presentation, researchers have begun to explore the stability of ASD diagnoses in young children (see Tables 1 and 2). Woolfenden, Sarkozy, Ridley, and Williams (2012) recently conducted a review of the literature regarding the stability of ASD diagnoses. The review revealed that diagnoses of autistic disorder under the age of three were relatively unstable, with studies reporting movement to ASD or to nonspectrum diagnoses for 0 to 30% of children. Diagnoses of other ASDs including PDDNOS, atypical autism, and Asperger's syndrome made under the age of three were even more unstable, with studies indicating movement to nonspectrum diagnoses or to an autistic disorder diagnosis in 0 to 53% of participants. Researchers also have compared stability of ASD relative to nonspectrum diagnoses in children. For example, Lord and colleagues (2006) evaluated the stability of diagnoses from two to nine years of age. In a sample of 172 children including both children diagnosed with ASD and those identified as nonspectrum, 5% of children with a diagnosis of ASD at age two received a non-spectrum diagnosis at age nine, whereas 26% of children identified as nonspectrum at age two were later identified with an ASD, indicating that many children were missed. Similarly, Chawarska, Klin, Paul, Macari, and Volkmar (2009) found that in a sample of 89 children diagnosed with ASD at age two, none of the children moved off of the spectrum when re-evaluated at age four, and 10% of children initially identified as nonspectrum, were later identified as being on the spectrum. However, several studies have reported larger and smaller rates of movement onto the spectrum within their samples, resulting in an unclear picture of what percentage of children have a change in diagnosis (see Table 2). Guthrie, Swineford, Nokke, and Wetherby (2013) also looked at diagnosis stability in children evaluated before age two until three to four years of age. Children were either initially identified as ASD, developmentally delayed, or typically developing, however, some participants were difficult to classify and were not given an initial diagnosis. Any child that received a specific diagnosis initially, retained that diagnosis, but the children that were not given a diagnosis, were scattered across diagnostic groups at follow up. Therefore, clear cases showed high diagnostic stability, but unclear cases had variable outcomes. What these studies highlight is how nonspectrum diagnoses may also be unstable in children under age three. Diagnosticians tend to be conservative at young ages and symptoms may be very subtle, resulting in a group of children that are either misdiagnosed as non-spectrum, or are not diagnosed with an ASD until a later date.

This lack of stability of diagnoses in children under age three may be due to unstable behavioral presentation or subtle symptom presentation at younger ages that current diagnostic measures do not accurately capture, as they were originally designed for children at older ages. For example, Charman and colleagues (2005) found that scores on standard psychometric measures and symptom severity as indexed by the Autism Diagnostic Interview-Revised (Rutter, LeCouteur, & Lord, 2003) conducted at age two were not predictive of diagnostic and cognitive outcome at age seven. However, scores on the same assessments at age three were highly correlated with outcome at age seven. This lack of consistency points to several diagnostic problems, including the possibility that some diagnostic symptoms may simply not be present at younger ages and symptom presentation may be less stable at younger ages.

Changes in Diagnoses

There has been an increased focus on children who no longer meet diagnostic criteria for an ASD over time. This work has examined how children's symptoms change, and how children who lose the ASD diagnosis differ from children who retain the diagnosis (Kleinman et al., 2008; Turner & Stone, 2007; Woolfenden, Sarkozy, Ridley, & Williams, 2012). There has been less focus on subtle symptoms exhibited by children that are not identified as having symptoms indicating risk for an ASD at early ages but later are identified as being on the spectrum. Most studies are not designed to capture this type of movement as the focus of the research is usually to specifically study children with ASD, and not typically developing children (Charman et al., 2005; Daniels et al., 2011; Jonsdottir et al., 2007; Paul, Chawarska, Cicchetti, & Volkmar, 2008; Turner & Stone, 2007; Turner, Stone, Pozdol, & Coonrod, 2006). Therefore, those children who do not initially meet criteria for ASD are not included in the study, and those children who move onto the spectrum later are not captured within the sample. There have been several studies designed to create a sample of children that included nonspectrum cases and children with ASD (Chawarska, Klin, Paul, Macari, & Volmar, 2009; Guthrie, Swineford, Nottke, & Wetherby, 2013; Kleinman et al., 2008; Lord et al., 2006; Scambler, Hepburn, & Rogers, 2006; van Daalen et al., 2009). These studies report the changes in diagnoses within both groups over time, and many report children that move onto the spectrum later, however, any information beyond the change in diagnosis is rarely reported (see Table 2). Considering the impact of early intervention, (Boyd, Odom, Humphreys, & Sam, 2010; Corsello, 2005; Dawson, 2008; Eldevik et al., 2009) earlier identification of these children is imperative, and information about specific behavioral characteristics that may predict later diagnosis would be essential in helping to improve the early identification process. Research has not yet carefully examined the behavioral characteristics in children initially identified as nonspectrum and later diagnosed with ASD.

Benefits of Early Identification

The main benefit of identifying patterns of early symptoms is the opportunity to provide earlier diagnoses and intervention for these children. Researchers and clinicians have been working toward diagnosing children with ASD at the youngest ages possible, but the emergence of symptoms during the first two years of life results in difficulty in diagnosing until enough clear symptoms become apparent. Therefore clinicians may miss a significant group of children as evidenced by the fact that a portion of participants in diagnostic stability studies are sometimes misdiagnosed as nonspectrum initially, and are later diagnosed with an ASD. (Chawarska et al., 2009; Lord, et al., 2006; Scambler, Hepburn, & Rogers, 2006; van Daalen, et al., 2009). The current research sought to inform not only theoretical concepts by illuminating early displays of symptoms and developmental progression, but also applied concepts by furthering our understanding of how ASD may differentially present itself in this young population, which can inform future diagnostic practices and access to early intervention services.

Current Investigation

The current investigation sought to identify behavioral markers of ASD that are observable under the age of two. Unlike much of the previous research on early markers of ASD, the current investigation focuses on markers present within a general population rather than a baby sibling population only. Additionally, the current study sought to characterize the behavioral presentation of late-identified cases of ASD, or cases of ASD that presented initially as nonspectrum, but later presented as having an ASD, in addition to studying early-identified cases of ASD that were initially and continually diagnosed with an ASD. Early markers of ASD and patterns of development were investigated for both groups of children and compared to children identified as typically developing or diagnosed with a language delay, unrelated to a diagnosis of ASD. Specific aims were:

1. *To identify early behavioral markers specific to ASD present between 12-24 months of age through eye-tracking behavior and behavior during observations of exploration of toys and play with a parent.* Performance during these assessments was compared across diagnostic groups at their initial evaluations only, to assess for early differences. Early and late-identified cases were also specifically evaluated to examine whether the late-identified cases of ASD were showing any behaviors associated with ASD at their initial evaluation, before receiving a diagnosis of ASD. It was hypothesized that the late-identified cases of ASD would show performance more similar to the early-identified cases of ASD than to the typically developing or language delay cases, even prior to their diagnosis of ASD.
2. *To identify how well behavioral markers from initial assessments classify ASD cases versus nonspectrum cases.* In follow-up to the previous analyses, significant variables from initial assessments were then all used in a discriminant analysis to see how well specific behaviors classified between a final ASD and nospectrum diagnosis. It was hypothesized that the variables would distinguish ASD and nonspectrum cases fairly well. The aim was to identify variables that led to the best classification, to see which behaviors showed the largest and most

consistent differences between diagnoses.

3. *To compare trajectories of development across typically developing, language delayed, early-identified ASD and late-identified ASD cases.* Longitudinal data on developmental assessments was modeled across diagnostic groups to compare trajectories of development over time. It was hypothesized the ASD groups would show marked delays in comparison to the typically developing children and even the language delayed children. Differences in development between the early and late-identified children were expected. Early-identified children were expected to initially present as more severely delayed than late-identified children, although differences were not expected to be as evident over time as late-identified children begin to show more symptoms of ASD over time.

METHODS

Participants

Participants were recruited for a large scale study examining the development of infants and toddlers with risk for ASD. Child participants were identified through community referral and by using the One-Year Well-Baby Check-Up Approach (Pierce et al., 2011). Using this approach, pediatricians in the San Diego community were recruited to implement the Communication and Symbolic Behavior Scales Developmental Profile Infant-Toddler Checklist (CSBS; Wetherby & Prizant, 2002) to screen all toddlers at their 12-month check-up for developmental delays. The CSBS is a short questionnaire completed by the parent that assesses the development of communication and social skills of children ages 6-24 months (see Appendix A). The CSBS provides scoring information and cut-off scores indicating a concern of delay based on the child's age (see Appendix B). Toddlers who met the cut-off for concern of possible developmental delays as indicated by the score received on the questionnaire were referred by pediatricians to the research team for further evaluation. A portion of children who did *not* meet the cut-off for concern were also randomly selected by pediatricians and were also referred to participate in the study for further evaluation. This method created a diverse sample of children, including typically developing children, children with language delays and developmental delays such as ASD.

Measures

Once recruited, children received a comprehensive diagnostic and developmental evaluation by an experienced psychologist. Parents were then asked to return with their child every six to twelve months until age three to complete the same evaluation to track development over time. At these evaluations, child functioning was assessed using several standardized and novel assessments (see Table 3 for a summary).

Background Information

Identification of Simplex and Multiplex Cases. Information about the siblings of the children in the study was gathered through caregiver interview and questionnaires. This information was used to determine whether children with an ASD diagnosis would be considered a simplex case, multiplex case, or stoppage/only child case. Simplex cases were defined as a child with ASD that had one or more siblings, but no siblings with and ASD. Multiplex cases were defined as a child with ASD that had one or more siblings with ASD. Stoppage cases were instances when the child with ASD was the only child the family, and therefore simplex or multiplex status could not be determined.

Treatment Information. Parents completed a questionnaire regarding any services their child had ever received. The questionnaire asked about services for speech therapy, physical therapy, and occupational therapy as well as any autism specific treatment (e.g. behavioral or developmental-based therapy). Parents were asked to report the start and end dates of treatment and the average hours of treatment received each week for each treatment service. Parent completed this survey at each longitudinal evaluation, but only the questionnaire completed at the latest age available was analyzed, as it provided the most information about treatment received across time.

Standardized Assessments

Autism Diagnostic Observation Schedule (ADOS). The ADOS is a semi-structured assessment used to measure behavioral features of ASD (Lord, Rutter, DiLavore, & Risi, 2002; Lord et al., 2012; Lord, Luyster, Gotham, & Guthrie, 2012). An examiner conducts a series of activities with the individual that are designed to allow the examiner to observe various aspects of social and communication behaviors associated with ASD. During the observation, behaviors are noted by the examiner and later coded. Overall scores are determined and cut-off scores for autism or autism spectrum disorder are provided. The ADOS is designed to be used as a tool to help inform the clinician's overall diagnostic judgment. The ADOS has several different modules for use across the lifespan. The Toddler Module is specifically designed for 12 to 30 month olds who do not consistently use phrase speech, Module 1 is designed for children 31 months and older who do not consistently use phrase speech, Module 2 is used for children of any age who use phrase speech but are not verbally fluent, Module 3 is used for verbally fluent children and young adolescents, and finally Module 4 is for verbally fluent older adolescents and adults. Children in the study were administered the Toddler Module, Module 1, and 2 throughout the course of the study depending on age and ability. Originally, the first edition of the ADOS (Lord, Rutter, DiLavore, & Risi, 2002) was administered, but the second edition (Lord et al., 2012; Lord, Luyster, Gotham, & Guthrie, 2012) was used once it was released. The second edition of the ADOS also provides standardized severity scores which allow for a comparison of symptoms of ASD between individuals, while taking into account the individual's age and developmental and communication level. Severity scores for the first edition of the ADOS were also calculated for comparison. Severity scores range from 1 to 10 with 10 indicating very high severity of symptoms. Severity scores are comparable across Modules 1-3 of the ADOS. Severity score calculations have not yet been released for the ADOS Toddler Module or Module 4. Severity scores on the ADOS Modules 1 and 2 were used within analyses to compare severity of impairment at age three in children with ASD.

Mullen Scales of Early Learning (MSEL). The MSEL is used to assess overall developmental functioning of children between birth and 68 months (Mullen, 1995). An examiner measures child functioning level through a series of play-like tasks over five domains; visual reception, receptive language, expressive language, fine motor, and gross motor skills. For each scale, the assessment derives a standardized score with a mean of 50 and standard deviation of 10, a percentile score, and an age equivalent score indicating at what developmental age the person is performing. An Early Learning Composite Score is calculated from the total of scores on all scales (excepting the gross motor scale) with a mean of 100 and standard deviation of 15. The MSEL is a common developmental test used when assessing children with ASD, and has shown good convergent validity with multiple developmental assessments (Akshoomoff, 2006; Bishop, Guthrie, Coffing, & Lord, 2011). Standardized scores for subdomains across longitudinal evaluations were used to analyze trajectories of development over time.

Vineland Adaptive Behavior Scales (VABS). The VABS is used to provide a measure of adaptive behavior through caregiver report (Sparrow, Balla, & Cicchetti, 1984; Sparrow, Cicchetti, & Balla, 2005). The caregiver completes a questionnaire regarding the individual’s current level of functioning across five domains: communication, daily living skills, socialization, motor skills, and maladaptive behavior. All scales provide standard scores with a mean of 100 and a standard deviation of 15, a percentile score, and an age equivalent score indicating at what developmental age the individual is performing. Scores on all scales are combined to obtain an overall Adaptive Behavior Composite Score with a mean of 100 and a standard deviation of 15. Two versions of the VABS exist, the first (Sparrow, Balla, & Cicchetti, 1984) and second editions (Sparrow, Cicchetti, & Balla, 2005). At the start of the study, the first edition of the VABS was used at 73 assessment timepoints, while the second edition was used for the remaining 850 assessment timepoints, as the new edition was released shortly after the start of the study. The VABS provides evidence of validity for individuals with delays, and specifically for individuals with ASD, and has been widely used with individuals with ASD within the research literature (Ventola, Saulnier, Steinberg, Chawarska, & Klin, 2014). Standardized scores for subdomains across longitudinal evaluations were used to analyze trajectories of development over time.

Additional Assessments

Geometric Preference Test. An eye-tracking task, the geometric preference test, developed by Pierce, Conant, Hazin, Stoner, & Desmond (2011) was performed to measure visual stimuli preferences. A Tobii T120 Eye Tracker (2013) was used which measures visual fixations through infrared light sources and cameras that are integrated into a 17-inch thin film transistor monitor. The Tobii eye-tracker recorded the coordinates of the child's eye position through corneal reflection at a frequency of 120 Hz or 7200 data collections per min. During the assessment, children were seated in their parent's lap, 60 centimeters from a computer screen in a darkened room. Before engaging in the eye-tracking task itself, the children participated in a calibration task. During calibration, children were shown a picture of a cat in nine different locations on the screen, and the eye tracker measured characteristics of the toddler's eyes in order to calculate the gaze data. The Tobii system provided feedback on the quality of calibration, and if necessary, calibration was repeated until a good quality was reached. During the eye-tracking task, two dynamic images were presented side-by-side for a total of 60 seconds. One side featured a social stimulus, with scenes of children engaging in aerobics and dancing (Wenig & Landon, 2004) whereas, the other side featured a non-social stimulus of moving geometric shapes recorded from screen saver programs. Both stimuli provided active, colorful images for the child to observe. No audio information was provided. The visual angle of the rectangle that enclosed both images measured 26.30 degrees horizontally and 9.05 degrees vertically, at the 60 centimeter distance. In order to control for any differences due to spatial location, half of children were presented with the non-social stimuli on the left side, with the social stimuli on the right, and the other half were shown the same stimuli but with the locations interchanged. Each child was videotaped and eye-tracking behavior was monitored on a second screen by an experimenter using the live tracker software from the Tobii system (Tobii Studio 1.3, 2013) which allowed the experimenter to observe the child's gaze position in real time and monitor for any technical difficulties or behavioral problems from the child. Visual fixations on each stimulus was then calculated using a 35-pixel radius filter using Tobii software. Any time looking outside the area of each stimuli was excluded. The percentage of time spent looking at social or nonsocial stimuli at initial evaluations was compared across diagnostic groups. The geometric preference test has been used in previous research and has demonstrated that a subset of children with ASD spend more time attending to nonsocial stimuli than social stimuli, as compared to TD and developmentally delayed children (Pierce, Conant, Hazin, Stoner, & Desmond, 2011).

Exploration. In the exploration task children were instructed to play in a 12 by 12 foot room with toys placed in standardized locations throughout the room (Pierce & Courchesne, 2001; see Figure 1 for a depiction of the observation room). Some of the toys were functional items (e.g. a ball) while others were nonfunctional (e.g. a piece of string). Additionally, some of the toys were placed in containers that were difficult to open, while the remaining toys were simply placed on the floor. Exploration tasks completed at initial evaluations were coded for the number of items and containers explored, the amount of movement about the observation room, the percentage of appropriate exploration and play, percentage of stereotypic exploration and play, and percentage of off-task behavior, and were compared across diagnostic groups. See coding procedures below for further information. Previous research has shown that children with ASD show reduced rates of exploration, increased passivity, and often play nonfunctionally or stereotypically, in comparison to typically developing peers (Pierce & Courchesne, 2001).

Parent-Child Interaction (PCI). The PCI consisted of a 10-minute free-play interaction between the child and one parent. The parent-child dyad was given access to a standardized set of age-appropriate toys placed in standardized locations about the observation room (see Figure 2 for a depiction of the observation room set-up). The parent was instructed to play with their child as they normally would at home. PCI observations completed at initial evaluations were coded for gestures, approach to parent, social vocalizations, orientation toward parent, social referencing, and child affect and were compared across diagnostic groups. See coding procedures below for further information. Differences in social responsiveness reported by caregivers or observed during interactions with caregivers have been identified as early markers of ASD (Wan et al. 2012; Zwaigenbaum, et al., 2005).

Diagnostic Groups

Diagnoses were determined using best practice guidelines for diagnosing young children with ASD and other developmental disorders, including the use of a standardized observational measures of child behavior and parent report (Perry, Condillac, & Freeman 2002). Psychologists with specialized experience in child development and ASD interviewed the parents about the child's development and observed the child's performance in a battery of assessments (described above), and used clinical judgment to make a final diagnosis. Procedures for diagnoses followed these general guidelines: *Typically developing* children scored within average to above average across domains of the MSEL and VABS and did not meet criteria for ASD on the ADOS. Children identified with a *language disorder* scored at least one standard deviation below the mean on language and/or communication domains of the MSEL or VABS, and usually did not meet criteria for ASD on the ADOS. Children identified with *ASD* could show variable scores on the MSEL and VABS, but most notably met criteria for ASD on the ADOS. The psychologists used these general guidelines for assessing performance on the assessments, but also relied on their own clinical judgment to provide the most accurate diagnosis as they saw fit based on the information available. Clinical expertise, in addition to the use of diagnostic assessment such as the ADOS, is considered an important part of an ASD diagnosis, especially for less clear cut cases, or when evaluating for a myriad of developmental delays, as opposed to testing for ASD only (Mazefsky & Oswald, 2006). If there was concern for an ASD, children were given provisional or "at-risk" diagnoses, prior to turning age three. At age three final diagnoses were given, since research supports that diagnostic status at age three is predictive of diagnostic status at a later age, whereas diagnoses at younger ages continue to be somewhat unstable (Charman et al., 2005). Children identified as at-risk for delays or ASD were referred to appropriate community intervention services.

Inclusion Criteria

To be included in the current study, children were required to have at least two assessments prior to the age of 36 months with their first assessment occurring before 24 months (mean age at initial assessment = 16.84, standard deviation = 3.87), and last assessment after 24 months (mean age at last assessment = 34.44, standard deviation = 3.90). These criteria resulted in a population of 130 typically developing children (TD), 59 children with language delay (LD), 69 children with an early-identified ASD (EI-ASD) diagnosis meaning they were diagnosed with ASD at their initial evaluation and continued to have an ASD diagnosis at all subsequent evaluations, and 41 children with a late-identified ASD (LI-ASD) diagnosis who were initially misdiagnosed, meaning they moved from a nonspectrum diagnosis to an ASD diagnosis across assessment periods (see Tables 4 and 5 for further participant characteristics). At the initial assessment the 41 children in the LI-ASD group included 11 children identified as typically developing, 18 children diagnosed with a language delay, nine children with a developmental delay, one child with a motor delay, and two children were labeled as having features of ASD but were not given a diagnosis of ASD as they did not show enough symptoms of ASD and scored within normal ranges on assessments. Additionally, 45 of the 59 children diagnosed with a language delay at their initial evaluation did not retain that diagnosis at a subsequent evaluation, and 14 children had a consistent diagnosis of language delay on all assessments through age three. However, all of these children had a diagnosis of a language delay at their initial evaluation, which is the time point of focus for the analyses.

Children in the TD group and EI-ASD group were required to have consistent identification of TD or ASD diagnosis, respectively, and any children that showed variation in diagnoses were not included. For example, if a child was initially identified as TD, then received a diagnosis of LD at a later evaluation, but later tested as TD again, they would not be included in the analysis, since their diagnosis was too variable and they did not clearly belong to one diagnostic group. An additional comparison group of children with developmental delay was not included due to the small number of children with this diagnosis within the sample (n=20). T-tests revealed there were significant differences between groups at the age of first evaluation (all $p < .01$), with the EI-ASD on average being 2.98 months older than the average age of all participants when first evaluated. Therefore, age at first assessment was considered within the analyses. There were no significant differences on age at last evaluation between the TD, LD, and ASD groups. There were also differences in the proportions of male and female children across diagnostic groups, with higher proportions of male participants in the EI-ASD, LI-ASD, and LD groups, compared to the TD group (see Table 4). Thus, gender was also considered within analyses as well. All 299 children were assessed using the ADOS, MSEL, and VABS at all evaluations. Of these 299 children, 245 had data available for the exploration task at initial assessments, 202 for the PCI at initial assessments, and 128 for the geometric preference eye-tracking test at initial assessments. Ninety-five children completed all three of the additional assessments at intake. See Table 6 for a breakdown of participants by assessment.

Coding Procedures

Additional measures that required behavioral coding were scored from videotape for various behaviors. Coding procedures for each paradigm are described below.

Quantity of Exploration Coding

Exploration observations at initial assessments were video recorded and later coded using a continuous five-second partial-interval scoring procedure. Using this procedure, the 10-minute observation was broken down into five-second intervals during which the observer recorded whether the child was exploring any of the assessment items. If the child explored an item during the interval, the interval was marked for exploration. The interval was not marked if the behavior did not occur during the interval. Exploration of each item in the assessment was recorded individually, allowing for a count of the number of toy items and containers explored to be calculated. Videos were also coded to identify the child's movement about the observation room. The assessment room was divided into four equal quadrants marked by masking tape. An additional fifth quadrant was added to account for additional alcove space surrounding the door to the assessment space (see Figure 1 for diagram of the layout). Continuous five-second quasi whole-interval scoring was used to identify which quadrant the child was in for the majority of the interval. The number of quadrant changes was then calculated for each child to quantify the amount of movement during the observation. If the child was not in view of the camera, the interval was marked as "unscorable." The number of items explored, the number of containers explored, and the amount of movement about the observation room were used in the analyses. See Appendices C and D for a sample scoring sheet and scoring definitions.

Type of Exploration Coding

Exploration observations at initial assessments were coded for the type of exploration or play the child was engaging in. Videos were coded using a continuous five-second quasi whole-interval scoring procedure. Using this procedure, the 10-minute observation was broken down into five-second intervals and the scorer marked a single play category to best represent the entire interval. Scorers selected one of three main coding categories; appropriate exploration and play, stereotypic exploration and play, and off-task behavior. Subcategories of more specific behavior were selected within each main category to create more specific definitions of behavior to ease the coding process. Only main categories of behavior were used within the analyses. If the child engaged in behaviors representative of more than one play category the category that best fit the behavior taking place for the majority of the interval was chosen. Only one play behavior category was selected for each interval. One exception to this rule was regarding stereotypic behaviors. Stereotypic behaviors could be marked along with another category if the stereotypic behavior occurred for the minority of the interval. Usually, any behavior occurring for the minority of the interval would not be marked, but an exception was made for stereotypic behaviors because the behaviors were often brief, leading to an underestimation of the occurrence of the behavior. For example, many children engaged in hand flapping behavior with the toy nets included the assessment, and this behavior often occurred for one or two seconds only, rather than the majority of the interval, and then the child might play appropriately with the toy for the majority of the interval. In this instance, the stereotypic behavior category would be selected, along with the appropriate code to identify the child's behavior for the majority of the interval. If the child was out of the view of the camera, the interval was marked as "unscorable." The percentage of the total intervals in which the target behavior occurred was recorded for each behavior category, resulting in the percentage of time spent engaging in appropriate exploration or play, stereotypic behavior and play, or off-task and these scores were then used in the data analyses. See Appendices E and F for a sample scoring sheet and scoring definitions.

PCI Coding

PCI observations at initial assessments were coded for several child social behaviors. Videos were coded using a continuous 5-second partial-interval scoring procedure. Using this procedure, the 10-minute observation was broken down into 5-seconds intervals and the scorer recorded whether any of the target behaviors; reaching, pointing, showing, giving, approach, social vocalizations, or social referencing, occurred at any point during the interval. The interval was not marked if the behavior did not occur during the interval, and the interval was only marked once if the behavior occurred multiple times during the interval. Orientation toward parent was coded if the child was oriented toward their parent for the majority of the interval. If the child was out of the view of the camera, the interval was marked as “unscorable.” The total number of intervals in which the target behavior occurred was recorded for each target behavior. Additionally, every 30 seconds ratings of child affect were made using a 1-7 Likert scale, with 1 indicating high negative affect, 4 indicating a neutral affect, and 7 indicating a high positive affect. Ratings were then averaged to create an overall affect rating for each child. As individual gestures occurred somewhat infrequently, scores for gestures were compiled in two ways. First, all gestures (reaching, pointing, showing, and giving) were totaled and measured as one category, then only “social” gestures were totaled, which included only pointing, showing, and giving. In comparison to all other gestures, all groups engaged in reaching most frequently. Reaching was often used to gain access to a toy that the parent controlled, and it was hypothesized that reaching was not always social in nature, whereas the other types of gestures needed another partner to be involved. The number of all gestures combined, number of social gestures, number of approaches to parent, number of social vocalizations, number of intervals oriented toward parent, number of social references, and average rating of child affect were included in the analyses. See Appendices G and H for a sample scoring sheet and scoring definitions.

Exclusionary Criteria for Coding Paradigms

For the exploration and freeplay tasks, at least 50% of the observation was required to be "scoreable" to be included in the analysis. If an observation was cut short (i.e. child was tantrumming and observation was discontinued) or the child was out view of the camera for more than 50% of the observation, these videos were excluded from the analyses. Three exploration observations were excluded (all from the EI-ASD group), resulting in a total of 245 videos included in the analyses. The majority of videos provided a full 10 minutes of scoreable data (mean length = 9.67 minutes, SD = 0.67). No PCI observations were excluded for this reason, with the majority of videos also being the full ten minutes in length (mean length = 9.83 minutes, SD = 0.29). For the eye-tracking paradigm, 154 of the 299 children participated in the assessment. Of these 154 children, 128 had usable data and were included in the analysis. Initially, 10 children (5 EI-ASD, 2 LD, 3 TD) were excluded because there were difficulties in calibration. Calibration is mostly an automated process in the Tobii system that teaches the system the characteristics of the user's eye, in order to track where the user is looking.

Difficulties can arise during calibration, especially if the child is moving. Beyond technical difficulties, the main inclusion criteria, was the availability of 50% or more of the data. The eye-tracking video is 60 seconds long, and children were required to attend to at least 30 seconds of the video to be included within the analyses. Twelve children (2 EI-ASD, 4 LI-ASD, 2 LD, 4 TD) did not attend to the video for a long enough period of time. However, most children attended to the video for majority of the time (mean = 49.75 seconds, SD = 9.97). Additionally, four children were excluded for other behavioral interferences (i.e. crying, tantrumming) during the task, as their attention may be directed elsewhere and vision impaired.

Inter-rater Reliability

Inter-rater reliability was calculated for each coding paradigm. Undergraduate research assistants were trained to score practice videos of children not included in the current study because they did not meet the inclusion criteria (e.g. did not complete more than one evaluation, did not have a consistent diagnosis, etc.) for each coding paradigm. These student observers were instructed on the scoring procedures and definitions and participated in trial scoring. Feedback on scoring was provided and practice continued until the research assistant reached inter-rater reliability of 85% or above, across at least three consecutive sample videos. Two coders independently coded 30% of videos for each paradigm to check reliability of scoring. All coders were kept blind to child diagnoses throughout the coding process.

Reliability was assessed using single measures intraclass correlation (two-way random effect model using absolute agreement). For the quantity of exploration coding paradigm, high agreement was found across all variables; number of items explored (r =

.989, $p < .001$), number of containers explored ($r = .940$, $p < .001$), and number of quadrant changes ($r = .983$, $p < .001$). There was also high agreement across type of exploration coding variables; appropriate exploration and play ($r = .957$, $p < .001$), stereotypic exploration and play ($r = .906$, $p < .001$), and off-task behavior ($r = .951$, $p < .001$). For the PCI coding paradigm, there was also fairly high agreement across variables; all gestures ($r = .830$, $p < .001$), social gestures ($r = .884$, $p < .001$), approach ($r = .791$, $p < .001$), social vocalizations ($r = .879$, $p < .001$), orientation toward parent ($r = .904$, $p < .001$), social referencing ($r = .863$, $p < .001$), and child affect ($r = .896$, $p < .001$).

Data Analysis

Identifying Early Behavioral Characteristics of ASD

To address aim one, and identify early behavioral markers specific to ASD present between 12-24 months of age, performance on eye-tracking, exploration tasks, and PCI observations at initial evaluations were compared across diagnostic groups to identify behavioral characteristics specifically associated with the development of an ASD. Additionally, the LI-ASD group was compared to the EI-ASD group, TD group and the LD group to examine whether the LI-ASD groups displayed any behavioral characteristics associated with ASD at their initial assessment, before receiving a diagnosis of ASD. Initial performance on the quantity of exploration, type of exploration, and PCI paradigms were compared across the EI-ASD, LI-ASD, LD, and TD groups using a multivariate ANOVA, to analyze differences in behavior between groups on these measures. Differences in the geometric preference eye-tracking test between groups was assessed using an ANOVA. Analyses were carried out using IBM

SPSS Version 22 (IBM Corporation).

Predictors of Diagnosis at Age Three

To address aim two, and identify how well behavioral markers from initial assessments predicted final diagnoses, exploratory discriminant analyses were conducted. After identifying significant behavioral variables from the multivariate ANOVA and ANOVA analyses of the geometric preference test, exploration, and PCI assessments, exploratory discriminant analyses were performed to see how well the significant variables predicted final diagnosis. First, these variables were used to predict diagnoses of EI-ASD or a nonsprectrum diagnosis, including LD and TD cases. Then the discriminant analyses were performed again, and the same variables were used to predict cases of LI-ASD or nonspectrum diagnoses. ASD groups were analyzed separately to look more closely at differences between EI-ASD and LI-ASD groups. Discriminant analyses were carried out using JMP Pro 11 (SAS Institute Inc.).

Developmental Trajectories

To address aim three, and compare trajectories of development across diagnostic groups, growth curve analysis (also referred to as linear mixed modeling within the literature) was used to examine longitudinal trajectories of development on the MSEL and VABS from approximately 12 months up to 42 months of age. Growth curve analysis is a type of multilevel modeling used for longitudinal data that estimates between-subject differences while taking into account within-subject change through the use of both fixed and random components. In this type of analysis, separate intercepts and slopes are calculated for each child in order to control for the high correlations among repeated measures on the same individuals over time. Thus, a growth trajectory is estimated for each subject, and then is combined with estimates from the other individuals to estimate an overall mean growth rate for the entire group. Due to this design, growth curve analysis offers flexibility in dealing with repeated observations and variable amounts of time between repeated measurements across subjects (Curran, Obeidat, & Losardo, 2010). These benefits made growth curve analysis useful within the constraints of the current data.

The MSEL and VABS were selected as they are both standardized assessments of development that are normed on typical development. Standardized scores on the MSEL Expressive Language domain and the VABS Socialization domain from intake to exit were modeled and were the primary focus of the analysis as behaviors measured on these domains are associated with diagnostic criteria of ASD. Scores on the MSEL Visual Reception and Receptive Language domains and the VABS Communication and Daily Living Skills domains were also modeled as exploratory analyses as these domains were hypothesized to show interesting differences between diagnostic groups as well. Growth curve modeling was used to compare the four diagnostic groups with respect to the initial scores at 12 months of age (i.e., the intercept), and the rate of change from 12 to 42 months of age (i.e., the slope). Within each model of MSEL and VABS scores, diagnostic group, age at the time of assessment, and the interaction between diagnostic group and age at assessment were considered fixed effects. Age at which each subject's assessment was conducted was also specified as a random effect in the model. This specification allowed each individual to have his or her own slope (growth rate of each MSEL or VABS score) in the model apart from the population slope (growth rate). Because subjects had variable amounts of assessments, ranging from 2 assessments to 5 assessments over the course of the study, linear models were used, as higher-order models could not be modeled for all subjects due to a lack of sufficient number of timepoints. Analyses of trajectories were carried out using JMP Pro 11 (SAS Institute Inc.

Additional Characteristics of EI and LI-ASD Groups

Lastly, additional information regarding the background of the EI-ASD and LI-ASD groups was collected to further characterize the groups. Information concerning family history of ASD, severity of symptoms at age three, and information regarding the amount of treatment received from intake to exit was collected and compared between groups.

Genetic Differences. The proportions of simplex, multiplex, and stoppage cases were identified within the EI-ASD group and the LI-ASD group, and were compared using a chi-square test to assess for differences in proportions of simplex, multiplex, and stoppage cases across groups. This analysis was carried out using IBM SPSS Version 22 (IBM Corporation).

ADOS Severity Scores. ADOS severity ratings were calculated for the EI-ASD and LI-ASD groups at age three to determine if there were any differences in the severity of ASD symptoms between the EI-ASD and LI-ASD groups. Fourteen of the children in the EI-ASD (n=6) and LI-ASD (n=8) groups received a Toddler ADOS at their final appointment. Severity scores have not yet been released for the Toddler ADOS and severity scores were not calculated for these children. The Toddler ADOS may be used for children up to 30 months of age, and these children were last seen at ages younger than 31 months. However, most children were followed until 36 months or later, and were given a different module of the ADOS. Severity scores were calculated for 96 children (63 EI-ASD and 33 LI-ASD). A two-tailed, independent samples t-test was conducted to test for differences in ratings between groups.

Differences in Treatment. Lastly, latency to treatment start and the amount of treatment received was compared for the EI-ASD, LI-ASD, and LD groups, as they were likely to receive services. The amount of time from the child's first assessment with the research group to beginning autism-related treatment or developmental services such as speech therapy, physical therapy, or occupational therapy, was calculated to determine how quickly children received services. An estimate of the average number of hours of autism related treatment and developmental services per week was also calculated. The latency to starting treatment and the average treatment hours per week were averaged for each group and compared for any large discrepancies across groups.

RESULTS

The specific aims of this investigation were to identify early behavioral markers specific to ASD present between 12-24 months of age, identify how well these behavioral markers predicted final diagnoses, and examine trajectories of development across diagnostic groups. Investigation of early behavioral markers for ASD were analyzed through eye-tracking behavior, and behavior during observations of exploration of toys and play with a parent at initial assessments. A series of ANOVA analyses were performed to examine differences between groups on these measures. It was hypothesized the paradigms analyzed would indicate several behavioral delays associated with not only the EI-ASD group, but the LI-ASD group as well. The LI-ASD group was predicted to show subtle behavioral deficits associated with ASD even at their initial assessment, prior to a diagnosis of ASD. To follow up these analyses, discriminant analyses were performed to see how well behavioral markers identified in the previous analyses, then went on to correctly predict ASD or nonspectrum diagnoses. It was hypothesized that the variables would distinguish ASD and nonspectrum cases fairly well. Lastly, growth curve modeling was performed to examine trajectories of development across diagnostic groups on standardized developmental assessments. It was hypothesized the ASD groups would show marked delays in comparison to LD and TD cases. Additionally, the EI-ASD group was expected to initially present as more severely delayed than LI-ASD group, although differences are not expected to be as evident over time as LI-ASD children begin to show more symptoms of ASD.

Early Behavioral Markers

Analyses were conducted to examine differences in behaviors across diagnostic groups at initial evaluations. Analyses centered around examining for differences in rates or patterns of behavior that were associated with the ASD groups, as opposed to the TD or LD groups to identify behavioral markers of ASD. A series of behavioral assessments were examined, including the eye-tracking paradigm, exploration task, and PCI.

Geometric Preference Test

For the geometric preference eye-tracking test, an ANOVA was performed to analyze the differences between diagnostic groups on the percentage of time looking at the geometric stimuli, rather than the social stimuli at initial evaluations. Bonferroni corrections were used to account for multiple comparisons. The results showed an overall significant difference in the amount of time looking at geometric stimuli across diagnostic groups ($F_{(3, 124)} = 16.084$, $p < .001$). Diagnostic group accounted for approximately 28 percent of the variance in the percentage of time looking at geometric stimuli ($\eta^2 = .280$). Differences between diagnostic groups were also examined; the TD group spent the least amount of time looking at geometric stimuli (M = 16.81%, SD = 11.82), followed by the LD group (M = 24.45%, SD = 15.22), then the LI-ASD group (M = 27.19%, SD = 22.13), with the EI-ASD group looking at the geometric stimuli for the largest amount of time (M = 44.84%, SD = 25.83; see Figure 3). Post-hoc pairwise comparisons between diagnostic groups revealed a significant difference between the EI-ASD group and the TD group ($p < .001$), the EI-ASD group and the LD group ($p < .001$), and the EI-ASD group and LI-ASD ($p = .012$). There were no significant differences between the TD, LD, and LI-ASD groups.

Quantity of Exploration

For the exploration paradigm, a MANOVA was used to examine differences in the quantity of exploration; the number of items explored, number of containers explored, and the amount of movement about the observation room, across diagnostic groups at initial evaluations. Bonferroni corrections were used to account for multiple comparisons. There was a significant overall effect for diagnoses on overall quantity of exploration (Wilk's Lambda = .838, $F_{(9, 581.8)} = 4.871$, $p < .001$). Diagnostic group accounted for approximately five percent of the variance in the overall quantity of exploration ($\eta^2 = .057$).

Differences across diagnostic groups were also examined on the specific variables of the number of items explored, the number of containers explored, and the amount of movement about the observation room. Means and standard deviations for each diagnostic group on each variable are listed in Table 7. The results of the between subjects tests indicated that there was a significant difference between diagnostic groups on the number of items explored ($F_{(3, 241)} = 11.593$, $p < .001$, $\eta^2 = .126$, see Figure 4), and on the number of containers explored ($F_{(3, 241)} = 4.162$, $p = .007$, $\eta^2 = .049$, see Figure 5), but not on the amount of movement about the observation room ($F_{(3, 241)} = .352$, $p = .787$, $\eta^2 = .004$, see Figure 6). Post-hoc pairwise comparisons between diagnostic groups were conducted to examine for differences on the number of items explored and the number of containers explored. On the number of items explored, there was a significant difference between the TD group and the EI-ASD group ($p < .001$), the TD group and the LI-ASD group ($p < .001$), and the LD group and the EI-ASD group ($p = .009$). For the number of containers explored, there was a significant difference between the TD and EI-ASD group ($p = .026$) and the TD and LI-ASD group ($p = .043$).

Type of Exploration and Play

A MANOVA was utilized to examine the differences in the type of exploration and play; the percentage of the observation engaging in appropriate exploration and play, percentage of stereotypic exploration and play, and the percentage of time off-task, across diagnostic groups at initial evaluations. Bonferroni corrections were used to account for multiple comparisons. There was a significant overall effect for diagnoses on overall type of exploration and play (Wilk's Lambda = .850, $F_{(9, 581.8)} = 4.469$, $p < .001$). Diagnostic group accounted for approximately five percent of the variance in the overall type of exploration and play ($\eta^2 = .053$).

Differences across diagnostic groups were also examined on the specific variables of the percentage of the observation engaging in appropriate exploration and play, the percentage of stereotypic exploration and play, and the percentage of off-task behavior. Means and standard deviations for each diagnostic group on each variable are listed in Table 8. The results of the between subjects tests indicated that there was a significant difference between diagnostic groups on the amount of appropriate exploration and play ($F_{(3, 241)} = 4.432$, $p = .005$, $\eta^2 = .052$, see Figure 7), and the amount of stereotypic exploration and play ($F_{(3, 241)} = 8.434$, $p < .001$, $\eta^2 = .095$, see Figure 8), but not on the amount of off-task behavior ($F_{(3, 241)} = 1.669$, $p = .174$, $\eta^2 = .020$, see Figure 9). Post-hoc pairwise comparisons between diagnostic groups were conducted to examine differences of the amount of appropriate exploration and play and the amount of stereotypic exploration and play. For the amount of appropriate exploration and play, there was a significant difference between the TD group and the EI-ASD group ($p = .031$), and the TD group and the LI-ASD group ($p = .016$). There were no significant differences between the remaining group comparisons. For the amount of stereotypic exploration and play, there was a significant difference between the TD and EI-ASD group ($p < .001$), the LD and EI-ASD group ($p = .001$). There were no significant differences between the remaining groups.

PCI

Social behaviors measured during the PCI, including all gestures (reaching, pointing, showing, and giving), social gestures (pointing, showing, and giving), approach, social vocalizations, orientation, social referencing, and child affect, were examined using a MANOVA to assess for differences between diagnostic groups at initial evaluations. Bonferroni corrections were used to account for multiple comparisons. There was a significant overall effect for diagnoses on all social behaviors measured during PCI (Wilk's Lambda = .727, $F_{(21, 551.8)} = 3.091$, $p < .001$). Diagnostic group accounted for approximately 10 percent of the variance in social behaviors ($\eta^2 = .101$).

Differences across diagnostic groups were also examined on the specific variables of all gestures, social gestures, approach, social vocalizations, orientation, social referencing, and child affect. Means and standard deviations for each diagnostic group on each variable are listed in Table 9. The results of the between subjects tests indicated that there was a significant difference between diagnostic groups on the amount of all gestures ($F_{(3, 198)} = 4.097$, $p = .008$, $\eta^2 = .058$, see Figure 10), the amount of social gestures ($F_{(3, 198)} = 6.888$, $p < .001$, $\eta^2 = .094$, see Figure 11), the amount of social vocalizations ($F_{(3, 198)} = 5.246$, $p = .002$, $\eta^2 = .074$, see Figure 12), and the amount of social referencing ($F_{(3, 198)} = 6.044$, $p = .001$, $\eta^2 = .084$, see Figure 13), but not on the amount of approach to parent ($F_{(3, 198)} = 1.232$, $p = .299$, $\eta^2 = .018$, see Figure 14), orientation towards parent ($F_{(3, 198)} = .811$, $p = .489$, $\eta^2 = .012$, see Figure 15), or child affect ($F_{(3, 198)} = 2.163$, $p = .094$, $\eta^2 = .032$, see Figure 16). Post-hoc pairwise comparisons between diagnostic groups were conducted to examine for differences on all gestures, social gestures, social vocalizations, and social referencing. For the amount of all gestures, there was a significant difference between the TD group and the EI-ASD group ($p = .006$). For the amount of social gestures, there was a significant difference between the TD and EI-ASD group ($p = .001$), the TD and LI-ASD group ($p = .015$), and the LD and EI-ASD group ($p = .030$). For the amount of social vocalizations, there was a significant difference between the TD group and the EI-ASD group ($p = .008$), and the TD group and the LD group ($p = .034$). For the amount of social referencing, there was a significant difference between the TD group and the EI-ASD group ($p =.001$) and the TD and LD group ($p = .024$).

Age as a Covariate. On average, the EI-ASD group was slightly older than the other groups (see Table 4). Thus it is possible age at intake could be affecting the results of the analyses. The ANOVA and MANOVA analyses were performed again with age at first assessment included in the analysis as a covariate, to test whether including age as a covariate affected the interpretation of the results.

Geometric Preference Test. For the geometric preference test analysis the results of the ANCOVA showed an overall significant difference in the amount of time looking at geometric stimuli across diagnostic groups ($F_{(3, 123)} = 11.384$, $p < .001$, $\eta^2 = .217$). There was not a significant covariate effect for age ($F_{(1, 123)} = 0.837$, $p =.362$, $\eta^2 = .007$). Since there was not a significant effect of age as a covariate, and there is still a significant effect for diagnostic group when age is included as a covariate, it is unlikely the results of the model are driven by any age differences between groups, and there is a true difference in scores between diagnostic groups.

Quantity of Exploration. The results of the MANCOVA analysis for the quantity of exploration showed there was a significant overall effect for diagnoses on overall quantity of exploration (Wilk's Lambda = .830, $F_{(9, 579.4)} = 5.107$, $p < .001$, $\eta^2 = .060$). There was also a significant covariate effect for age (Wilk's Lambda = .903, $F_{(3, 238)} =$

8.550, $p < .001$, $\eta^2 = .097$). The results of the between subjects tests indicated that there was a significant difference between diagnostic groups on the number of items explored ($F_{(3, 240)} = 13.439$, $p < .001$, $\eta^2 = .144$), and a borderline significant difference in the number of containers explored ($F_{(3, 240)} = 2.533$, $p = .058$, $\eta^2 = .031$), but no significant difference in the amount of movement about the observation room ($F_{(3, 240)} = .070$, $p = .976$, $\eta^2 = .001$). In this analysis, there was only a borderline significant effect for the number of containers explored, but there was a significant effect in the previous analysis when age was not included as covariate. Therefore, the interpretation of differences between groups on the number of containers explored should be interpreted with caution.

Type of Exploration. For the type of exploration analysis, the results of the MANCOVA showed a significant overall effect for diagnoses on overall type of exploration (Wilk's Lambda = .856, $F_{(9, 579.4)} = 4.240$, $p < .001$, $\eta^2 = .050$). There was not a significant covariate effect for age (Wilk's Lambda = .988, $F_{(3, 238)} = .958$, $p = .413$, $\eta^2 = .012$). The results of the between subjects tests indicated that there was again a significant difference between diagnostic groups on the amount of appropriate exploration and play ($F_{(3, 240)} = 5.196$, $p < .002$, $\eta^2 = .061$), the amount of stereotypic exploration and play ($F_{(3, 240)} = 6.746$, $p < .001$, $\eta^2 = .078$), but not for the amount of time off-task ($F_{(3, 240)} = 1.993$, $p = .116$, $\eta^2 = .024$). In this analysis there was not a significant covariate effect of age, and the same variables were once again significant as in the previous analysis.

PCI. For the PCI analysis, the results of the MANCOVA showed a significant overall effect for diagnoses on social behaviors measured during PCI (Wilk's Lambda = .608, $F_{(21, 549)} = 4.946$, $p < .001$, $\eta^2 = .153$), and a significant covariate effect for age (Wilk's Lambda = .670, $F_{(7, 191)} = 13.410$, $p < .001$, $\eta^2 = .330$). The results of the between subjects tests indicated that there was a significant difference between diagnostic groups on the amount of all gestures ($F_{(3, 197)} = 7.949$, $p < .001$, $\eta^2 = .108$), and the amount of social gestures ($F_{(3, 197)} = 12.319$, $p < .001$, $\eta^2 = .158$), the amount of social vocalizations ($F_{(3, 197)} = 20.478$, $p < .001$, $\eta^2 = .238$), the amount of social referencing ($F_{(3, 197)} = 7.697$, $p < .001$, $\eta^2 = .105$), and again there were no significant differences between groups on the amount of approach to parent ($F_{(3, 197)} = 2.385$, $p = .070$, $\eta^2 = .035$) or orientation towards parent ($F_{(3, 197)} = .706$, $p = .550$, $\eta^2 = .011$). However, there was a significant effect for child affect ($F_{(3, 197)} = 4.013$, $p = .008$, $\eta^2 = .058$), that was not significant in the previous model. Therefore, differences between groups on child affect should be interpreted with caution.

Gender as a Covariate. Additionally, it was hypothesized gender may have an impact of the results as there are different proportions of gender between groups (see Table 4). The analyses were run once again with gender included as a covariate to examine whether gender had any effect of the results.

Geometric Preference Test. For the geometric preference test analysis, the results of the ANCOVA showed an overall significant difference in the amount of time looking at geometric stimuli across diagnostic groups ($F_{(3, 123)} = 16.088$, $p < .001$, $\eta^2 = .282$). There was not a significant covariate effect for gender ($F_{(1, 123)} = 1.155$, $p = .285$, $\eta^2 = .009$). Given there was not a significant effect of gender as a covariate, and there is still a significant effect for diagnostic group as in the main analysis, it is unlikely the results of the model are driven by any gender differences between groups.

Quantity of Exploration. The results of the MANCOVA analysis for the quantity of exploration showed there was a significant overall effect for diagnoses on overall quantity of exploration (Wilk's Lambda = .859, $F_{(9, 579.4)} = 4.166$, $p < .001$, $\eta^2 = .050$).

There was not a significant covariate effect for gender (Wilk's Lambda = 1.874, $F_{(3, 238)}$ = 1.874, p = .135, η^2 = .023). The results of the between subjects tests indicated that there was a significant difference between diagnostic groups on the number of items explored ($F_{(3, 240)}$ = 10.022, p < .001, η^2 = .111), and for the number of containers explored ($F_{(3, 240)}$ = 2.877, p = .037, η^2 = .035), and again there was not a significant effect on the amount of movement about the observation room ($F_{(3, 240)}$ = .244, p = .866, η^2 = .003). The results with gender included as a covariate are very similar to the original analysis, suggesting gender is not driving the results of the model.

Type of Exploration. For the type of exploration analysis, the results of the MANCOVA showed a significant overall effect for diagnoses on overall type of exploration (Wilk's Lambda = .863, $F_{(9, 579.4)}$ = 4.023, p < .001, η^2 = .048). There was not a significant covariate effect for gender (Wilk's Lambda = .992, $F_{(3, 238)}$ = .677, p = .567, η^2 = .008). The results of the between subjects tests indicated that there was again a significant difference between diagnostic groups on the amount of appropriate exploration and play ($F_{(3, 240)}$ = 4.092, p = .007, η^2 = .049), and the amount of stereotypic exploration and play ($F_{(3, 240)}$ = 7.389, p < .001, η^2 = .085), but not for the amount of time off-task ($F_{(3, 240)}$ = 1.735, p = .160, η^2 = .021). In this analysis there was not a significant covariate effect of gender, and the same variables were once again significant as in the previous analysis.

PCI. For PCI, the results of the MANCOVA showed a significant overall effect for diagnoses on social behaviors measured during PCI (Wilk's Lambda = .749, $F_{(21, 549)}$ = 2.769, p < .001, η^2 = .092), but there was not a significant covariate effect for gender (Wilk's Lambda = .946, $F_{(7, 191)}$ = 1.556, p = .151, η^2 = .054). The results of the between subjects tests indicated that there was a borderline significant difference between diagnostic groups on the amount of all gestures ($F_{(3, 197)} = 2.550$, $p = .057$, $\eta^2 = .037$), and significant effects for the amount of social gestures ($F_{(3, 197)} = 4.997$, $p = .002$, $\eta^2 = .071$), the amount of social vocalizations ($F_{(3, 197)} = 4.276$, $p = .006$, $\eta^2 = .061$), the amount of social referencing ($F_{(3, 197)} = 5.483$, $p < .001$, $\eta^2 = .077$), and again there were no significant differences between groups the amount of approach to parent ($F_{(3, 197)} = 1.242$, $p = .296$, $\eta^2 = .019$), orientation towards parent ($F_{(3, 197)} = 1.188$, $p = .315$, $\eta^2 = .018$) or for child affect ($F_{(3, 197)} = 2.484$, $p = .062$, $\eta^2 = .036$). In this analysis as well, there was not a significant covariate effect of gender, and the same variables were once again significant as in the previous analysis.

Overall, the additional analyses including gender and age as covariates, indicate age at intake and gender had little effect on the results and difference in age or gender were not solely driving the results of the model. Thus the differences found between groups is likely due to actual differences between diagnostic groups, rather than due to any differences in age or gender between groups.

Predictors of Diagnostic Classification

Exploratory discriminant analyses were performed in follow up to the previous ANOVA and MANOVA analyses to examine how well the behaviors from initial evaluations previously measured predicted final diagnostic outcome. All of the significant variables from the previous analyses (1. percentage of time looking at geometric stimuli during geometric preference test, 2. number of items explored, 3. amount of appropriate play during the exploration task, 4. amount of stereotypic play during the exploration task, 5. frequency of social gestures during freeplay, 6. frequency of social vocalizations during freeplay, 7. frequency of social referencing during freeplay) were included in the discriminant analyses. First, these variables were used to predict diagnoses of EI-ASD or nonsprectrum diagnosis, including LD and TD cases. Then the discriminant analyses were performed again, and the same variables were used to predict cases of LI-ASD or nonspectrum diagnoses. Only children with all assessments (geometric preference test, exploration, and PCI) were included in the analysis. A total of 95 children (25 EI-ASD, 12 LI-ASD, 17 LD, 41 TD) had all assessments.

EI-ASD vs Nonspectrum

The results of the discriminant analysis focusing on EI-ASD and nonspectrum cases showed a significant overall effect for all seven behavioral variables predicting diagnostic group (Wilk's Lambda = .485, $F_{(7, 75)} = 11.381$, $p < .001$). The canonical structure of the discriminant function indicates factor loadings of each variable and represents the correlations between the observed variables and the dimensions created with the discriminant functions. The predictor variables with the highest factor loadings were percentage of time looking at geometric stimuli (r = 0.874), amount of stereotypic play (r = 0.564), number of items explored (r = -0.549), amount of social referencing (r = -0.333), with the amount of social gestures (r = -0.192), amount of social vocalizations (r = -0.159), and the amount of appropriate play (r = -0.096) representing much smaller factor loadings. Lastly, the discriminant analysis correctly classified 90.36% (75 of 83) into the correct diagnostic group based on the seven behavioral measurements. For EI-ASD cases, 84% (21 of 25) were correctly classified as being in the EI-ASD group and 93.10% (54 of 58) nonsprectrum cases were correctly classified as nonspectrum.

LI-ASD vs Nonspectrum

The results of the discriminant analysis focusing on LI-ASD and nonspectrum cases did not show a significant overall effect for all seven behavioral variables predicting diagnostic group (Wilk's Lambda = .871, $F_{(7, 62)} = 1.313$, $p = 0.259$). According to the canonical structure report, the factor loadings of each variable were as follows; percentage of time looking at geometric stimuli ($r = -0.643$), number of items explored ($r = 0.599$), amount of social gestures ($r = 0.423$), amount of social vocalizations ($r = 0.403$), amount of social referencing ($r = 0.402$), percentage of appropriate play ($r = 0.399$), and percentage of stereotypic play ($r = -0.276$). The discriminant analysis classified 68.57% (48 of 70) into the correct diagnostic group based on the seven behavioral measurements. Of the LI-ASD cases, 75.00% (9 of 12) were correctly classified as being in the LI-ASD group and 67.24% (39 of 58) nonsprectrum cases were correctly classified as nonspectrum.

Developmental Trajectories

Analyses were conducted to determine differences in trajectories of development across diagnostic groups, and specifically identify any differences in development between the EI-ASD and LI-ASD groups. Scores on the MSEL and VABS assessments from intake to exit (12-42 months) were modeled across diagnostic groups.

MSEL

For the MSEL, standardized scores on the expressive language, receptive language, and visual reception domains were modeled separately. Scores from the expressive language domain were the main analysis of interest, as expressive language is often impaired in children with ASD. Receptive language and visual reception were also modeled as supplementary analyses, as differences were also hypothesized between groups within these domains.

MSEL Expressive Language. Fixed effect tests for expressive language standard scores revealed a significant effect of age ($F_{(1, 236.4)} = 47.785$, $p < .001$), diagnostic group ($F_{(3, 260.7)} = 39.009$, $p < .001$) and a significant interaction between age and diagnostic group ($F_{(3, 239.2)} = 7.583$, $p < .001$). In other words, age at assessment had an impact on scores, the diagnostic group of the child had an effect on scores, and diagnostic groups performed differently at different ages and demonstrated different slopes (see Figure 17). Table 10 depicts the parameter estimates for the fixed effects and random effects of the growth curve model for MSEL expressive language scores. The intercept of the model indicates the average MSEL expressive language score was 33.22 at intake. The positive and significant effect of age indicates increases in MSEL expressive language scores over time. The fixed effects parameter estimates are difficult to interpret with regard to multiple diagnostic groups. It is important to note the TD group is treated as the reference group and parameter estimates of each diagnostic group are compared to the reference group, but not to each other. Results of the parameter estimates should not be generalized to interpret differences to mean differentiating from all diagnostic groups. With that said, the LD and EI-ASD group showed lower scores with respect to the TD group. The coefficients for interactions between diagnostic group and age depicted different slopes, with the LD demonstrating more rapid increases over time than the TD group, and the LI-ASD showing negative change in comparison to the TD group. Random effects parameter estimates are also depicted in Table 10. There was high variance due to individual subject differences.

MSEL Receptive Language. Fixed effect tests for receptive language standard scores revealed a significant effect of age ($F_{(1, 237.8)} = 20.891$, $p < .001$), and diagnostic group ($F_{(3, 236.9)} = 25.417$, $p < .001$), but there was not a significant interaction between age and diagnostic group ($F_{(3, 241.8)} = 1.166$, $p = .323$; see Figure 18). The parameter estimates for the fixed effects and random effects of the model are depicted in Table 11. The intercept of the model indicates the average MSEL receptive language score was 38.81 at intake. The positive and significant effect of age indicates increases in MSEL receptive language scores over time. The LD and EI-ASD group showed lower scores with respect to the TD group. None of the coefficients for interactions between diagnostic group and age depicted significantly different slopes. Random effects parameter estimates again showed there was high variance due to individual subject differences.

MSEL Visual Reception. Fixed effect tests for visual reception standard scores did not show a significant effect for age ($F_{(1, 220)} = 0.590$, $p = .443$) or diagnostic group ($F_{(3, 207.3)} = 1.234$, $p = .298$), but there was a significant interaction between age and diagnostic group ($F_{(3, 244.1)} = 9.342$, $p < .001$; see Figure 19). The parameter estimates for the fixed effects and random effects of the model are outlined in Table 12. The intercept of the model indicates the average MSEL visual reception score was 52.09 at intake. There was not a significant effect of age, but the parameter estimate indicates a negative trend, with an overall decrease in scores over time. None of the parameter estimates for each diagnostic group were significant. The coefficients for interactions between diagnostic group and age showed the LD demonstrating more rapid increases over time than the TD group. Random effects parameter estimates once again showed there was high variance due to individual subject differences.

Limitations. Approximately nine percent of the sample floored (i.e., scored at the bottom of the scorable range) on the MSEL in at least one domain, with the majority of instances occurring in the receptive or expressive language domains. Individuals in the EI-ASD group were most likely to floor, with 65% of those that floored in at least one domain belonging to the EI-ASD group, 29% belonging to the LI-ASD group and the remaining six percent in the LD group. In order to test whether the floor effects were affecting the interpretation of the trajectories of development, trajectories on the expressive language, receptive language, and visual reception domains were also modeled using quintile scoring. Standard scores on each domain were divided into five equal groups and assigned a scores of 1, 2, 3, 4, or 5. Floor effects create a lack of variability in scores at the lower level, and the quintile scoring method removes variability in scores across multiple levels of scores. This allows for comparison of trajectories when modeled using standard scores and when modeled using quintile scores to examine for any differences or lack thereof between each variation of the model that may be driven by floor effects. Quintile scoring was used to model scores of the expressive language, receptive language, and visual reception domains. Fixed effect tests for expressive language quintile scores revealed a significant effect of age ($F_{(1, 213.1)} = 47.917$, $p < .001$), diagnostic group ($F_{(3, 249.5)} = 32.050$, $p < .001$) and a significant interaction between age and diagnostic group ($F_{(3, 216.1)} = 7.174$, $p < .001$; see Figure 20). Fixed effect tests for receptive language quintile scores revealed a significant effect of age ($F_{(1,253.5)} = 14.179$, $p < .001$), and diagnostic group ($F_{(3, 248.6)} = 22.298$, $p < .001$), with no significant interaction between age and diagnostic group ($F_{(3, 257.8)} = 1.239$, $p = .296$; see Figure 21). Fixed effect tests for visual reception quintile scores once again did not show a significant effect of age ($F_{(1, 218.6)} = 1.584$, $p = .210$) or diagnostic group ($F_{(3, 209.4}$ $= 1.117$, $p = .343$), but a significant interaction between age and diagnostic group ($F_{(3, 222.8)}$ $= 6.846$, $p < .001$; see Figure 22) as seen with the standard scores as well. Quintile scoring for expressive language, receptive language, and visual reception showed a very similar pattern of results to those found using standard scores, suggesting floor effects on the MSEL were not driving the effects of the model. Additionally, at 92 of the longitudinal evaluations, the participants were given the Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI; Wechsler, 2012), when considered too advanced to be given the MSEL. Seven of these cases were in the EI-ASD group, five in the LI-ASD group, 16 in the LD group, and 64 in TD group. These timepoints were not included in the trajectory models, as there are not equivalent scores on the WPPSI. As most of the participants were from the TD group, any differences due to the exclusion of these children would predominantly effect the scores of the TD group. However, the TD group showed average to above average scores across domains without these children, so adding in the highest performing children would likely have only inflated their scores. Thus, the interpretation of the performance of the TD group would likely still be much higher on average in comparison to the other groups.

EI-ASD and LI-ASD Comparison. As the ASD groups were the groups of most interest, the model was re-run with only the ASD groups included to specifically examine for differences in developmental trajectories of the EI-ASD and LI-ASD groups. This allowed for the direct comparison of significant differences between the two groups that was not possible using the full model, with four diagnostic comparison groups. It is important to note that this model is not directly translatable to the full model including all diagnostic groups, as the parameter estimates are affected with fewer subjects included in the model. However, this exploratory analysis comparing the EI-ASD and LI-ASD group allows for a preliminary examination of potential differences between the ASD groups.

Fixed effect tests on the expressive language domain revealed a significant effect of age ($F_{(1, 93.6)}$ = 8.194, p = .005), and diagnostic group ($F_{(1, 89.4)}$ = 6.005, p = .016), but there was not a significant interaction between age and diagnostic group ($F_{(1, 93.6)}$ = 1.732, p = .191. The parameter estimates for the fixed effects and random effects of the model are outlined in Table 13. Fixed effect tests for receptive language scores revealed a significant effect of age ($F_{(1, 93.5)}$ = 8.695, p = .004), and diagnostic group ($F_{(1, 88.5)}$ = 9.154, p = .003), with no significant interaction between age and diagnostic group ($F_{(1, 93.5)}$ = 2.424, p = .123). See Table 14 for full parameter estimates. Fixed effect tests for visual reception scores showed a significant effect of age ($F_{(1, 79.5)}$ = 5.335, p = .024), but not for diagnostic group ($F_{(1, 71.5}$ = 0.947, p = .334), or for the interaction between age and diagnostic group ($F_{(1, 79.5)}$ = 0.787, p = .378). See Table 15 for full parameter estimates.

VABS

Standardized scores on the socialization, daily living skills, and communication domains were modeled separately. Scores from the socialization domain comprised the main analysis of interest, as social deficits are a main diagnostic criteria for ASD. Scores on the communication and daily living skills domains were also modeled as supplementary analyses, as differences were also hypothesized between groups within those domains.

VABS Socialization. Fixed effect tests on the socialization domain revealed a significant effect of age ($F_{(1, 269.6)}$ = 21.130, p < .001), diagnostic group ($F_{(3, 270.7)}$ = 3.633, p = .013) and a significant interaction between age and diagnostic group ($F_{(3, 265.9)}$ = 10.946, p < .001; see Figure 23). The parameter estimates for the fixed effects and random effects of the model are depicted in Table 16. The intercept of the model indicates the average VABS socialization domain score was 99.66 at intake. The negative and significant effect of age indicates decreases in VABS socialization scores over time. The EI-ASD group showed lower scores with respect to the TD group. The coefficients for interactions between diagnostic group and age showed the LI-ASD group showed negative change in comparison to the TD group. Random effects parameter estimates showed high variance due to individual subject differences.

VABS Communication. Fixed effect tests for the communication domain revealed a significant effect of age ($F_{(1, 286.6)} = 62.919$, $p < .001$) and diagnostic group ($F_{(3, 286.6)} = 28.272$, $p < .001$), and there was a significant interaction between age and diagnostic group ($F_{(3, 283.3)} = 6.633$, $p < .001$; see Figure 24). The parameter estimates for the fixed effects and random effects of the model are depicted in Table 17. The intercept of the model indicates the average VABS communication score was 79.83 at intake. The effect of age was significant indicating increases in VABS communication scores over time. The EI-ASD and LD groups showed lower scores with respect to the TD group. The coefficients for interactions between diagnostic group and age depicted different slopes, with the LD demonstrating a steeper positive slope than the TD group, while the LI-ASD showed negative change in comparison to the TD group. Random effects parameter estimates again showed high variance due to individual subject differences.

VABS Daily Living Skills. Lastly, the fixed effect tests for the daily living skills domain did not show a significant effect for age ($F_{(1, 269.7)} = 1.782$, $p = .183$), and the effects for diagnostic group were borderline significant ($F_{(3, 272.3)} = 2.572$, $p = .055$), but there was a significant interaction between age and diagnostic group ($F_{(3, 265.3)} = 11.950$, $p < .001$; see Figure 25). The parameter estimates for the fixed effects and random effects of the model are depicted in Table 18. The intercept of the model indicates the average VABS daily living skills score was 92.84 at intake. The estimate for age was not significant. The LD group showed lower scores with respect to the TD group, but the parameter estimate for the EI-ASD and LI-ASD groups were not significant. The coefficients for interactions between diagnostic group and age depicted different slopes across groups, with the LD demonstrating more rapid increases over time than the TD group, while the EI-ASD and LI-ASD showed negative change in comparison to the TD group. Random effects parameter estimates again showed there was high variance due to individual subject differences.

Limitations. A second edition of the VABS was released after the first evaluations of this study were conducted. The first edition of the VABS was used in 73 of the longitudinal evaluations (7 EI-ASD cases, 6 LI-ASD cases, 20 LD cases, and 40 TD cases), while the second edition was used for the remaining 850 longitudinal evaluations (196 EI-ASD cases, 120 LI-ASD cases, 159 LD cases, and 375 TD cases). The VABS manual provides a description of the differences between editions, with the overall change being additional questions in the second edition (see Table 19 for differences described in the manual). The developers of the VABS manual conducted a small scale study where they compared performance on the VABS 1 and 2 within 24 subjects between 0-24 months, and 29 subjects between ages 3-6 (see Table 20). Correlations are fairly robust ranging from 0.65-0.94, with the lowest correlation at 0.65 between VABS 1 and 2 on the communication domain for children between 0-24 months. Demographic differences within our own sample were examined comparing the timepoints when participants received the VABS 1 or VABS 2 (see Table 21). Participants who received the VABS 1 were younger on average (19.11 months) than those participants who received the VABS2 (25.85 months). This difference is expected as many children initially received the VABS 1, but received the VABS 2 later on once it was released. Proportions of gender were similar across versions, with the majority of participants being male across versions. Differences in ethnicity and race were minimal, but more categories of race were indicated within those that received the VABS 2, which may be attributed simply to the larger number of individuals that received the VABS 2. Overall, the groups of children that received the VABS 1 or 2 were very similar.

EI-ASD and LI-ASD Comparison. Once again, the model was re-run with only the ASD groups included to conduct an exploratory analysis comparing the developmental trajectories of the EI-ASD and LI-ASD groups. Fixed effect tests on the socialization domain revealed a significant effect of age ($F_{(1, 64.8)} = 33.407$, $p < .001$), and diagnostic group ($F_{(1, 66.5)} = 5.035$, $p = .028$), but there was not a significant interaction between age and diagnostic group ($F_{(1, 64.8)} = 2.085$, $p = .154$. The parameter estimates for the fixed effects and random effects of the model are outlined in Table 22. Fixed effect tests for the communication domain revealed a significant effect of age ($F_{(1, 92.9)} = 7.978$, $p = .006$), and diagnostic group ($F_{(1, 94.2)} = 6.172$, $p = .015$), but no significant interaction between age and diagnostic group ($F_{(1, 92.9)} = 2.316$, $p = .131$). See Table 23 for full parameter estimates. Fixed effect tests for daily living skills domain showed a significant effect of age ($F_{(1, 89.8)} = 7.035$, $p = .009$), but not for diagnostic group ($F_{(1, 90.4} = 1.512$, $p = .222$), or for the interaction between age and diagnostic group ($F_{(1, 89.8)} = 0.385$, $p = .537$). See Table 24 for full parameter estimates.

Further Characterization of the LI-ASD and EI-ASD Groups

Additional information regarding the background of the EI-ASD and LI-ASD groups was collected to further characterize the differences in symptoms presentation and development of these two groups of children. Information regarding family history of ASD, severity of symptoms at age three, and the amount of treatment received was collected and compared between groups.

Genetic Differences

The number of simplex, multiplex, and stoppage cases was identified within the EI-ASD and LI-ASD groups. In the EI-ASD group 23.19% (n=16) of cases were identified as multiplex cases, 40.58% (n=28) were identified as simplex cases, 34.78% (n=24) were identified as stoppage cases, and 1.45% (n=1) did not have sibling information available. In the LI-ASD group, 31.71% (n=13) were identified as multiplex cases, 31.71% (n=13) were simplex cases, 31.71% (n=13) were stoppage cases, and 4.88% (n=2) did not have sibling information available. A chi-square test was conducted to assess whether there were different proportions of multiplex, simplex, or stoppage cases across the EI-ASD and LI-ASD groups. The results of the chi-square were not significant, suggesting there were not significant differences of proportions of multiplex, simplex, or stoppage cases across diagnostic groups (χ^2 = 1.304, df = 2, p = .521).

ADOS Severity Scores

ADOS severity ratings were calculated for the EI-ASD and LI-ASD groups at exit to determine if there were any differences in the severity of ASD symptoms between the EI-ASD and LI-ASD groups. Severity scores were calculated for 96 children (63 EI-ASD and 33 LI-ASD) that completed a module 1 or 2 of the ADOS and had severity scores available. Fourteen of the children in the EI-ASD (n=6) and LI-ASD (n=8) groups received a Toddler ADOS at their final appointment and did not have severity scores available. On average, children in the EI-ASD groups had a severity rating of 7.57, with a standard deviation of 1.72. The children in the LI-ASD groups had an average rating of 6.88, with a standard deviation of 2.16. A t-test revealed that there were no significant differences between groups on severity scores at exit ($p = 0.09$).

Differences in Treatment

Since the EI-ASD and LI-ASD groups were identified as at-risk for ASD at different ages (EI-ASD 19.82 months on average, LI-ASD 27.66 months on average), the age at which these children were referred for autism specific services differed as well. Treatment records were available for 95% of the EI and LI-ASD cases. On average, the EI-ASD group started receiving autism related treatments (i.e. ABA based therapies such as Pivotal Response Training or Incidental Teaching) 2.82 months (SD = 0.10) after their first assessment with the research group, and received 11.70 hours (SD = 6.67) of treatment a week. On the other hand, the LI-ASD group began receiving ASD related services 8.60 months (SD = 0.32) after their first evaluation on average, and received 10.31 hours (SD = 8.32) of treatment a week. The ASD groups also often received developmental services such as speech therapy, occupational therapy, and physical therapy services in addition to autism specific therapy. On average, the EI-ASD group began receiving developmental services 4.89 months (SD = 6.17) after their first evaluation and received 1.38 hours (SD = 1.77) of treatment a week on average. The LI-ASD group received additional services 7.19 months (SD = 6.52) after their initial evaluation and received 1.65 hours (SD = 1.47) of therapy. Treatment records were available for 90% of the LD cases. The LD group also received 1.03 hours (SD = 1.02) of speech/occupational/physical therapy a week, and began receiving services 5.67 months (SD = 6.48) after their first evaluation with the research group.

DISCUSSION

This research sought to identify early behavioral markers associated with ASD within a general population sample. Early behaviors associated with ASD were also explored within the LI-ASD group, to examine whether LI-ASD cases were showing signs of ASD prior to receiving a diagnosis. Trajectories of development were also explored across diagnostic groups to identify differences in growth over time. Patterns of growth between EI-ASD and LI-ASD groups were specifically focused on to identify differences in development across groups.

Early Markers

As hypothesized, analyses of the geometric preference test, exploration task, and PCI paradigms led to the identification of several behavioral features associated with the EI-ASD group that distinguished them from the LD and TD groups. These included an elevated preference for geometric stimuli during the eye-tracking paradigm, a reduction of the number of items explored, reduced appropriate play and increased stereotypic play during the exploration task, and lower rates of gestures, social vocalizations and social referencing during PCI. These findings replicated much of the previous literature showing reductions in the aforementioned social behaviors in children with ASD as compared to TD children, and in this case LD children as well (Ozonoff et al., 2010; Pierce & Courchesne, 2001; Pierce, Conant, Hazin, Stoner, & Desmond, 2011; Wan et al. 2012, 2013; Zwaigenbaum et al., 2005). The LI-ASD group was also showing delays in these social behaviors, but the delays were not as severe as seen in the EI-ASD group. The LI-ASD group was often difficult to distinguish from the LD group, therefore making their behavior difficult to distinguish from other non-ASD delays. These analyses were focused on initial evaluations only, at which time the children in the LI-ASD group had not yet received a provisional diagnosis of ASD. It seems this group of children was definitely showing some delays in social behaviors that are commonly seen in children with ASD, but these delays were not very pronounced at this timepoint, which likely led to their initial nonspectrum diagnosis. Overall, the geometric preference test, exploration task, and PCI were all useful methods for capturing behavioral differences in toddlers with ASD, especially in the clear EI-ASD cases. Many of the social deficits associated with the early development of ASD identified with the baby sibling literature were also replicated within this sample suggesting the baby sibling literature may have good generalization to a more general population sample of infants and toddlers with ASD.

Classification

The classification results of the discriminant analyses between the EI-ASD group and nonspectrum cases were highly accurate when all of the significant variables (percentage of time looking at geometric stimuli, number of items explored, appropriate exploration and play, stereotypic exploration and play, social gestures, social vocalizations, and social referencing) were included. These results corroborate the idea that the behaviors analyzed in the current study are important factors for identifying ASD and in combination may lead to the ability to distinguish ASD cases from typical development and non-ASD delays (LD cases). Increased preference for geometric stimuli, increased stereotypic play, and a lack of exploration and social referencing indicated early risk for ASD. Unfortunately, the discriminant analysis for the LI-ASD and nonspectrum groups was far less robust than the EI-ASD and nonspectrum comparison. The differences between the LI-ASD and the LD and TD groups was not nearly as apparent as those differences seen between the EI-ASD and the LD and TD groups. Again, these analyses examined the performance at the children's first evaluation, which occurred prior to any of the LI-ASD children being identified as at-risk for an ASD. Therefore, it would be expected that these children are not showing symptoms as robust or clear as the EI-ASD group at this timepoint. Nonetheless, these analyses provide a method to explore which types of behavior may be the most promising for identifying subtle behavioral differences in the LI-ASD group early on. The percentage of time looking at geometric stimuli, and the number of items explored appeared to be the strongest contributors the model, and were also strong predictors in the EI-ASD vs nonspectrum analysis. Therefore assessing for visual preferences and exploration and play behaviors may be interesting avenues to pursue more in-depth for both populations.

Trajectories

As predicted, diagnostic groups showed variable patterns of development on MSEL and VABS. Also as expected, the TD children initially showed scores right around the mean, however it should be noted that the TD group showed slight increases in scores over time. Increases in scores over time were unexpected, as standardized scores are expected to maintain over time. However, increases were very minimal and could be due to noise in the measurement. The LD group showed similar development on most domains, with the exception of language domains, where they showed substantial increases in language scores over time. Increases in scores for the LD group were expected as several of these children no longer met criteria for a diagnosis of language delay as they aged. Language delay is often considered a transient diagnosis, especially when diagnosed in children under age two, with most children recovering and catching up to the performance level of their peers over time (Paul, 2000). Not surprisingly, both ASD groups showed significant delays in comparison to the TD and LD group, and even showed decreases in non-language domains (MSEL VR, VABS Socialization, and VABS Daily Living Skills). In the additional trajectory analyses focusing on the ASD groups, there was a significant negative impact of age, reflecting these decreases seen across these domain. Only the VABS socialization domain showed a significant difference between the EI-ASD and LI-ASD group, with the VABS daily living skills and MSEL VR domain showing similar results across ASD groups. Differences in performance over time on language domains was particularly interesting for the EI-ASD and LI-ASD groups. The additional trajectory analysis focusing on the ASD groups only showed significant differences between groups overall, but significant differences in slope were not detected across the language domains. This suggests there are overall differences in scores between the groups, but slopes were not distinctive enough to show significant differences between groups, likely due to the high variability within each group. However, upon visual inspection, the same trend in slopes was evident between groups across all language domains. Repeatedly, the EI-ASD group initially showed lower performance than the LI-ASD group, then the EI-ASD group demonstrated rapid progress, indicating an increase in scores close to a full standard deviation over time. The LI-ASD group also showed delays in language initially, but showed higher scores on average than the EI-ASD group. However, the LI-ASD group did not show the same increase in language scores, as the EI-ASD group. The language domains repeatedly depicted the LI-ASD group starting off with higher performance than the EI-ASD group, but then the EI-ASD group made enough progress to reach scores similar to the LI-ASD group at age three. These patterns of development suggest the LI-ASD group as a whole was initially performing at a higher level than the EI-ASD group.

This is not surprising, given the LI-ASD group was not initially identified with risk for ASD. What is concerning is the fact that the LI-ASD group does not maintain this higher level of performance over time. As a group, the LI-ASD cases do not show the same amount of progress in language development as the EI-ASD group.

Overall, these results suggest the EI-ASD and LI-ASD groups show different patterns of development. Both groups start off with different levels of development, but converge to a similar level of performance. These alternative trajectories, particularly the idea of subgroups of children that make either rapid or slower progress within an ASD population mirrors other variable patterns of development identified in previous literature (Fountain, Winter, & Bearman, 2012; Landa, Gross, Stuart, & Bauman, 2012; Shumway et al., 2011). Important information regarding potential ASD subgroups is explored within this project, and patterns of symptom onset may provide insight into differential patterns of development over time. These differences are particularly concerning given the fact that the EI-ASD group and LI-ASD received different amounts of treatment. The LI-ASD group began receiving autism related treatment approximately five months later on average than the EI-ASD group. This delay in treatment is inevitably due to the fact that the LI-ASD group also received a provisional ASD diagnosis later than the EI-ASD group. It is possible that this difference in treatment received impacted the developmental trajectories of each ASD group. This highlights the importance of beginning treatment as soon as possible, as beginning treatment at younger ages has been associated with better outcomes (Itzchak & Zachor, 2011). Additionally, at exit these groups of children showed no differences in severity scores ratings on the ADOS, suggesting both groups of children show similar levels of impairment by age three and one group is not more or less severely impacted by

ASD symptoms. Finally, the EI-ASD and LI-ASD groups showed similar rates of simplex and multiplex cases, suggesting differences between these two groups may not be easily identified though familial inheritance, and distinction between the groups will rely on identifying behavioral differences.

Limitations

One major limitation of the discriminant analyses is the lack of power. Only 95 of 299 cases had all of the assessments of interest, reducing the power of the analyses quite a bit. These analyses were exploratory in nature and therefore should be interpreted cautiously. Due to the exploratory nature and the small number of subjects, the analyses do not have strong explanatory information of the contributions of each factor to predicting diagnosis. Ideally, subsequent discriminant analyses would be performed in the future with a new cohort of children to test how well the variables identified in these analyses generalized to a new cohort of children. However, these analyses may give us information about whether these behaviors are useful as a whole, and some information about which are most useful for classification, but that this needs to be interpreted cautiously.

A limitation of the growth curve modeling was the lack of contrasts of effects between specific diagnostics groups. The fixed effects tests provided information on whether there were differences in scores between diagnostic groups on average, and if there were differences in trajectories of the diagnostic groups, but it does not provide specific information about whether there are significant differences between each of the groups. Currently, there is no ability to look at differences between groups at this level within statistical analysis packages. Thus differences between trajectories of development is limited to visual inspection. Therefore these differences should be should be interpreted with caution, as no information can be provided about whether these differences would be considered significantly different. The growth models were re-run only including the ASD groups to compare differences between these two groups. However, it should be noted that this analysis is not directly comparable to the full model, as fewer subjects are included and the parameter estimates are therefore different. Consequently, these comparisons, in respect to the full model, should also be interpreted with caution.

Additionally, there were relatively few LI-ASD cases in our sample. Future research would benefit from a larger LI-ASD group to further study the characteristics of this group. Also, the LI-ASD group consisted of a rather heterogeneous sample, with members of the groups starting off with a variety of initial diagnoses before receiving a provisional ASD diagnosis. Eleven children were considered typically developing, 18 were considered language delayed, nine were identified as developmentally delayed, one was identified as having a motor delay, and two cases were noted as having features of ASD, but did not show enough symptoms to warrant a diagnosis, and were performing within normal ranges of performance of assessments. Thus, the presentation of this group was variable at first, which inevitably led to variance in the data, which may have impacted the interpretation of the results. Future research to focus of subgroups of children within the LI-ASD cohort would be interesting. It would be valuable to understanding the differences in performance over time for children that are originally considered typically developing, who later received an ASD diagnosis, compared to children who are originally identified with a non-ASD delay and later received an ASD diagnosis. It is possible these groups of children will show different trajectories of development over time.

One overall limitation of this project is the lack of a non-ASD developmental delay comparison group. A language delay group was included as there were a fair number of children seen at the center with a language delay. There were very few cases of children that met criteria for a non-ASD developmental delay that also met the other study inclusion criteria, and were not included because there were not enough cases to allow for a robust comparison group. Future research incorporating a developmentally delayed comparison group in addition to a language delay group would be valuable to further assess behavioral characteristics that are highly specific to ASD. The use of multiple comparison groups, and comparison groups involving children with non-ASD delays are important to see how behavior of children with ASD differ not only from typical development, but also how it differs from other delays unrelated to ASD. It is important to identify behaviors highly specific to ASD, rather than behaviors that simply indicate a delay of some kind, as highly specific behavioral markers will be the most useful for diagnostic purposes (Zwaigenbaum et al., 2007).

Finally, the current study examined early behavioral markers of children between 12 and 24 months. While one year of age is a rather narrow age range, child behavior can look very different at 12 months as compared to 24 months. Explorative and play behaviors may develop extensively over this time period, and different levels of performance may be indicative of delay at different ages. Further, more specific study of the development of play behaviors, and what is considered within normal range is warranted to further the understanding of development and when to be concerned.

Future Research

Future research examining differences over time in the presentation of the behavioral markers studied in this project would be interesting to further our understanding of the development of ASD. For example, it would be particularly interesting to understand how visual stimuli preference for non-social or social stimuli vary across age groups. This information seems particularly important for understanding the LI-ASD group. On the geometric preference test, the LI-ASD group performed similarly to the TD and LD group, and on average preferred to look at the social stimuli. This was significantly different from the preferences of the EI-ASD who preferred to look at the non-social stimuli much more than the other groups. It would be interesting to know if the performance of the LI-ASD group changed over time. Perhaps once the LI-ASD group was given a provisional diagnosis of ASD, and was showing clear symptoms of ASD, their preferences for social or non-social stimuli during eye-tracking would change. A longitudinal analyses would be interesting for other early markers as well, to examine whether there was a reduction in exploration or play skills over time.

The LI-ASD group was sometimes difficult to distinguish from the TD group, and often difficult to distinguish from the LD group. The LI-ASD group did show reduced exploration and play skills, but not to the same extent as the EI-ASD group. Future research should try to identify robust early behavioral markers for the LI-ASD group. It is possible behaviors associated with ASD are even more subtle in the LI-ASD group and behaviors need to be examined at an even finer level. For example, appropriate exploration and play was also reduced in the LI-ASD group, but appropriate play was examined at a gross level. Any type of play (functional, symbolic, etc.) was accepted as appropriate play, but no information was gathered concerning the developmental level of play. Children generally progress through different types of play, with play becoming more sophisticated as they age. Previous research suggests children with ASD may show delays in this progression as compared to typically developing peers (Baranek et al., 2005). It is possible the ASD groups are also delayed in the type of play they are engaging in with regard to typical development. This level of analysis may be promising for determining robust differences identifying the LI-ASD group. Also, exploration of potential biological differences between EI-ASD and LI-ASD cases would help our understanding of these differential patterns of development. For example, overgrowth of the brain during infancy and childhood has been associated with ASD (Redcay & Courchesne, 2005) and differences in brain size or growth would be interesting to compare between groups to see if differences seen in behavioral development mirror any differences in biological development.

Ultimately, it is necessary to develop ways to identify these cases early on, so correct diagnoses can be given immediately. Within the current study families were seen every 6-12 months per the research protocol, but this level of monitoring is often not seen in the general community, and it is possible these children could have gone without receiving a diagnosis of ASD, and getting the services they need for quite some time. Therefore further research improving early identification is imperative to improve diagnosis and access to treatment, in order to have the greatest positive impact on outcomes.

Summary

In general, the existence of EI-ASD and LI-ASD groups highlight the heterogeneity of the ASD population, showing differences across children in behavioral presentation over time. These differences were highlighted in the trajectory analyses and point to possible subgroups of ASD, as the EI-ASD and LI-ASD groups showed different patterns of development. The geometric preference test, exploration task, and PCI provided useful methods to identify characteristics associated with ASD. Most significantly, increased preference for geometric stimuli, increased stereotypic play, and a lack of exploration and social referencing indicated early risk for ASD and may be useful for identifying ASD in toddlers A reduction in these social and communication skills was seen in the EI-ASD group, as well as the LI-ASD group, but the EI-ASD showed more severe impairments than the LI-ASD in these behaviors. The LI-ASD group was often difficult to differentiate from the LD group, but showed more severe impairments in skills than the LD group on average. The results of these analyses indicate that the LI-ASD children are showing observable differences in behavior, even prior to receiving a diagnosis of ASD, suggesting there are signs of ASD present earlier, but methods to detect these symptoms need to be refined. Additionally, many early behavioral features of ASD that have been identified using baby sibling populations, such as reduced social referencing, reduced social vocalizations, and passive behavior, were replicated within the current sample from a more general population. These results suggest that many of the findings from baby sibling research may be generalizable to the wider population.

Ultimately, this study has important implications for improving the early detection process for ASD and provides further direction in this line of research. Assessing visual preferences and play behaviors may be important considerations for the diagnostic process. The geometric preference test used in this study has been established as a method to reliably identify some cases of ASD when they show elevated preference for non-social stimuli (Pierce, Conant, Hazin, Stoner, & Desmond, 2011), and these results have been replicated in the current population focused exclusively on very young toddlers. Observing solitary play may also be very useful in addition to existing diagnostic tools. Observing play behavior is already currently part of the ADOS assessment, however, most of the play is directed by the examiner or involves participation of the examiner. Useful information about the child may become more apparent when the child is left to play on their own, so one can observe how the child independently explores toys or if they play repetitively. The addition of measuring behaviors in these ways offer new methods to examine for behaviors associated with ASD in very young children, and provide additional ways to assess for delays in cases of ASD that are very difficult to identify early on. In conclusion, this study provides important information about the development and different ways ASD may be presented in very young children, and provides insight into ways of refining and adding to current diagnostic practices.

TABLES

Table 1. Summary of Diagnostic Stability Articles Studying ASD Samples Under Age Three Since 2005.

First Author	Year	Participants	Initial Evaluation	Final Evaluation	Remained on Spectrum	Moved off Spectrum	Moved onto Spectrum
Charman	2005	26 children diagnosed with ASD	2 years	3 and 7 years	96.15% (25 of 26)	3.85% (1 of 26)	N/A
Daniels	2011	Survey of 7,106 children with ASD diagnoses from age 6 months to 18 years	6 mos to 18 years	6 mos to 18 years	100%	N/A	N/A
Jonsdottir	2007	41 children with various Pervasive Developmental Disorders	22 to 59 mos	61 to 79 mos	100%	0%	N/A
Paul	2008	37 children diagnosed with autism or PDDNOS	15 to 25 mos	3 to 5 years	100%	0%	N/A
Turner	2007	48 children diagnosed with autism or PDDNOS	2 years	4 years	62.50% (30 of 48)	37.50% (18 of 48)	N/A
Turner	2006	25 children with diagnosis of autism or PDDNOS	2 years	9 years	88% (22 of 25)	12% (3 of 25)	N/A

Table 2. Summary of Diagnostic Stability Articles Studying ASD and Nonspectrum Samples Under Age 3 Since 2005.

First Author	Year	Participants	Initial Evaluation	Final Evaluation	Remained on Spectrum	Moved off Spectrum	Moved onto Spectrum
Chawarska	2009	89 children referred for a differential diagnosis, 43 with autism, 18 with PDDNOS, 28 with developmental delay	13 to 27 mos	30 to 61 mos	100%	0%	10.71% (3 of 28)
Guthrie	2013	82 children that failed the CSBS screener and showed delays during a behavioral observation	15 to 24 mos	30 to 46 mos	100%	0%	0%
Kleinman	2008	77 children recruited for a larger study testing validity of the Modified Checklist for Autism in Toddlers (MCHAT) screener	16 to 35 mos	42 to 82 mos	75.41% (46 of 61)	24.59% (15 of 61)	0% (0 of 16)
Lord	2006	172 children referred for evaluation for possible autism	2 years	9 years	95.38% (124 of 130)	4.61% (6 of 130)	26.19% (11 of 42)
Scambler	2006	19 children with autism, 11 children with other developmental disabilities	2 to 3 years	4 to 6 years	100%	0%	45.45% (5 of 11)
van Daalen	2009	131 children that failed the Early Screening for Autistic Traits (ESAT) screener	26 mos	45 mos	86.79% (46 of 53)	13.21% (7 of 53)	2.56% (2 of 78)

Table 3. Description of Assessments and the Scores Used in the Analyses.

	Assessment	Description	Scores for Analyses
Standardized Assessments	Autism Diagnostic Observation Schedule (ADOS)	Evaluates children for ASD through a series of activities focusing on social and communication behaviors (Lord, Rutter, DiLavore, & Risi, 2002; Lord et al., 2012).	ADOS was used for diagnostic purposes. Severity ratings were used in analyses.
	Mullen Scales of Early Learning (MSEL)	Assesses developmental functioning of children, including fine motor, receptive language, expressive language and visual reception skills, to produce an overall composite score (Mullen, 1995).	Standardized scores on the visual reception, receptive language, & expressive language domains.
	Vineland Adaptive Behavior Scales (VABS)	Parent report measure of adaptive behavior including daily living skills, socialization, communication and motor skills, to produce an overall adaptive behavior score (Sparrow, Balla, Cicchetti, 1984; Sparrow, Cicchetti, & Balla, 2005).	Standardized scores on the communication, daily living skills, & socialization domains.
Additional Assessments	Geometric Preference Test	Assesses for visual preferences for social or non-social (geometric) stimuli, through an eye-tracking procedure (Pierce, Conant, Hazin, Stoner, & Desmond, 2011).	Percentage of time looking at geometric stimuli.
	Exploration Task	Observation of explorative and play behaviors with a standardized set of toys. Includes functional and nonfunctional toys, with some in containers and some easily accessible (Pierce & Courchesne, 2001).	Number of items and containers explored, the amount of movement around the room, and the amount of time engaging in appropriate play, stereotypic play, or off-task behavior.
	Parent Child Interaction (PCI)	A naturalistic observation of parent and child interactions while playing for 10 minutes with a standardized set of toys.	Frequency of gestures, approach, social vocalizations, orientation, social referencing, and ratings of child affect.

Table 4. Participant Demographics.

	All Participants	EI-ASD	LI-ASD	LD	TD
Age in Months	16.84 (3.87)	19.82 (3.63)	16.78 (3.86)	16.61 (3.30)	15.26 (3.23)
Age at Provisional Diagnosis	16.84 (3.87)	19.82 (3.63)	27.66 (5.89)	16.61 (3.30)	15.26 (3.23)
Gender					
Male	217 (72.58%)	59 (85.51%)	34 (82.93%)	47 (79.66%)	77 (59.23%)
Female	82 (27.42%)	10 (14.49%)	7 (17.07%)	12 (20.34%)	53 (40.77%)
Ethnicity					
Hispanic or Latino	75 (25.08%)	27 (39.13%)	10 (24.39%)	14 (23.73%)	24 (18.46%)
Not Hispanic or Latino	222 (74.25%)	42 (60.87%)	31 (75.61%)	45 (76.27%)	104 (80.00%)
Not Reported	2 (0.67%)	0 (0%)	0 (0%)	0 (0%)	2 (1.54%)
Race					
African American	8 (2.68%)	1 (1.45%)	1 (2.44%)	1 (1.69%)	5 (3.85%)
American Indian/Alaska Native	2 (0.67%)	0 (0%)	0 (0%)	0 (0%)	2 (1.54%)
Asian	16 (5.35%)	0 (0%)	0 (0%)	1 (1.69%)	15 (11.54%)
Caucasian	226 (75.59%)	45 (65.22%)	31 (75.61%)	53 (89.83%)	97 (74.62%)
Pacific Islander/Native Hawaiian	3 (1.00%)	0 (0%)	2 (4.88%)	0 (0%)	1 (0.77%)
Multiple Races Selected	32 (10.70%)	23 (33.33%)	3 (7.32%)	2 (3.39%)	4 (3.08%)
Not Reported	12 (4.01%)	0 (0%)	4 (9.76%)	2 (3.39%)	6 (4.62%)

Note: Averages with standard deviations in parentheses are listed for ages. Counts with percentages listed in parentheses for gender, ethnicity, and race.

Table 5. Participant Cognitive Scores at Intake.

	All Participants	EI-ASD	LI-ASD	LD	TD
MSEL					
Visual Reception	52.43 (10.87)	44.25 (9.32)	46.88 (11.45)	54.29 (9.69)	57.68 (8.31)
Fine Motor	53.26 (11.03)	43.83 (9.86)	49.76 (11.91)	52.86 (8.05)	59.55 (8.08)
Receptive Language	43.24 (14.03)	28.77 (10.08)	37.85 (11.84)	44.19 (11.54)	52.18 (9.76)
Expressive Language	42.12 (13.30)	30.46 (7.90)	37.20 (10.95)	35.41 (10.65)	52.92 (8.32)
Early Learning Composite	95.96 (19.34)	75.32 (12.17)	86.88 (17.40)	93.24 (11.32)	111.22 (12.01)
VABS					
Communication	89.84 (14.88)	76.77 (10.64)	83.61 (13.57)	86.64 (12.46)	100.19 (10.40)
Daily Living Skills	94.22 (11.64)	87.58 (10.78)	91.17 (11.56)	94.24 (10.91)	98.68 (10.50)
Socialization	98.35 (11.10)	86.93 (8.59)	94.27 (9.82)	99.34 (7.35)	105.25 (8.25)
Motor Skills	98.23 (9.67)	96.65 (10.48)	92.10 (11.88)	98.78 (8.43)	100.76 (7.91)
Adaptive Behavior Composite	93.87 (11.93)	84.32 (8.88)	88.22 (11.30)	92.83 (10.28)	101.20 (9.29)
ADOS					
Total Score	8.88 (7.80)	19.65 (5.12)	9.49 (5.90)	6.24 (4.61)	3.52 (2.84)

Note: Averages are listed for each measure with standard deviations in parentheses.

Table 6. Number of Participants that Completed Each Assessment by Diagnostic Group.

Diagnostic Group	All Participants	Exploration	PCI	Geometric Preference Test	Completed All Assessments
TD	130	108	88	56	41
LD	59	44	37	25	17
LI-ASD	41	37	27	16	12
EI-ASD	69	56	50	31	25
TOTAL	299	245	202	128	95

Table 7. Means and Standard Deviations for Diagnostic Groups on Quantity of Exploration Variables.

Variable	TD	LD	LI-ASD	EI-ASD
Number of Items Explored	11.74 (4.44)	10.93 (3.67)	8.57 (3.58)	8.29 (3.98)
Number of Containers Explored	4.30 (3.65)	3.21 (3.09)	2.60 (3.00)	2.73 (2.87)
Amount of Movement	15.01 (11.17)	15.32 (8.60)	15.57 (14.23)	16.89 (11.20)

Table 8. Means and Standard Deviations for Diagnostic Groups on Type of Exploration Variables.

Variable	TD	LD	LI-ASD	EI-ASD
Appropriate Exploration and Play	58.24 (18.02)	53.33 (14.08)	47.63 (22.68)	49.71 (18.84)
Stereotypic Exploration and Play	2.57 (4.78)	2.29 (5.96)	6.78 (12.65)	9.72 (14.95)
Off-Task Behavior	40.60 (18.37)	45.42 (16.25)	48.40 (24.57)	42.8 (21.25)

Table 9. Means and Standard Deviations for Diagnostic Groups on PCI Variables.

Variable	TD	LD	LI-ASD	EI-ASD
All Gestures	12.32 (7.77)	11.41 (8.22)	9.37 (7.27)	7.98 (5.86)
Social Gestures	6.22 (6.05)	5.89 (6.91)	2.67 (2.48)	2.64 (3.08)
Approach	2.39 (2.33)	3.03 (3.29)	1.70 (2.48)	2.52 (3.12)
Social Vocalizations	13.60 (14.01)	7.43 (8.92)	7.30 (7.52)	7.14 (8.64)
Orientation	34.63 (17.87)	32.43 (15.99)	35.00 (10.82)	38.54 (21.48)
Social Referencing	27.01 (12.18)	20.51 (8.57)	22.00 (10.42)	19.26 (12.29)
Child Affect	4.19 (0.25)	4.12 (0.36)	4.19 (0.43)	4.03 (0.50)

Table 10. Growth Curve Analysis Parameter Estimates for MSEL Expressive Language Standard Scores.

Fixed Effects Parameter Estimates

Parameter		Estimate	SE
Intercept		33.217 ***	1.186
Age		0.352 ***	0.051
Diagnosis			
	EI-ASD	-8.251 ***	2.13
	LI-ASD	1.777	2.354
	LD	-9.678 ***	2.203
	TD	-	-
Slope			
	Age x EI-ASD	0.001	0.088
	Age x LI-ASD	-0.234 *	0.101
	Age x LD	0.392 ***	0.089
	Age x TD	-	-

Random Effects Covariance Parameter Estimates

Parameter	Estimate	SE
Intercept Variance	123.623	30.804
Slope Variance	0.289	0.060
Intercept/Slope Covariance	-5.149	1.298

Note. Dashes indicate reference group.
* $p < .05$ **$p < .01$ *** $p < .001$.

Table 11. Growth Curve Analysis Parameter Estimates for MSEL Receptive Language Standard Scores.

Fixed Effects Parameter Estimates

Parameter		Estimate	SE
Intercept		38.805 ***	1.214
Age		0.212 ***	0.046
Diagnosis			
	EI-ASD	-14.979 ***	2.228
	LI-ASD	-1.375	2.386
	LD	4.913 *	2.059
	TD	-	-
Slope			
	Age x EI-ASD	0.134	0.082
	Age x LI-ASD	-0.138	0.091
	Age x LD	-0.007	0.080
	Age x TD	-	-

Random Effects Covariance Parameter Estimates

Parameter	Estimate	SE
Intercept Variance	57.72	32.852
Slope Variance	0.061	0.046
Intercept/Slope Covariance	-1.055	1.157

Note. Dashes indicate reference group.
* $p < .05$ **$p < .01$ *** $p < .001$.

Table 12. Growth Curve Analysis Parameter Estimates for MSEL Visual Reception Standard Scores.

Fixed Effects Parameter Estimates

Parameter		Estimate	SE
Intercept		52.090 ***	1.189
Age		-0.039	0.051
Diagnosis			
	EI-ASD	-3.985 *	2.177
	LI-ASD	1.461	2.335
	LD	2.377	2.026
	TD	-	-
Slope			
	Age x EI-ASD	-0.112	0.088
	Age x LI-ASD	-0.312 **	0.100
	Age x LD	0.068	0.088
	Age x TD	-	-

Random Effects Covariance Parameter Estimates

Parameter	Estimate	SE
Intercept Variance	51.647	34.913
Slope Variance	0.171	0.061
Intercept/Slope Covariance	-2.378	1.397

Note. Dashes indicate reference group.
* $p < .05$ **$p < .01$ *** $p < .001$.

Table 13. Growth Curve Analysis Parameter Estimates for MSEL Expressive Language Standard Scores for ASD Groups Only.

Fixed Effects Parameter Estimates

Parameter		Estimate	SE
Intercept		29.612 ***	2.029
Age		0.250 **	0.087
Diagnosis			
	EI-ASD	-4.971 *	2.029
	LI-ASD	-	-
Slope			
	Age x EI-ASD	0.115	0.087
	Age x LI-ASD	-	-

Random Effects Covariance Parameter Estimates

Parameter	Estimate	SE
Intercept Variance	151.417	58.582
Slope Variance	0.385	0.109
Intercept/Slope Covariance	-6.484	2.415

Note. Dashes indicate reference group.
* p < .05 **p < .01 *** p < .001.

Table 14. Growth Curve Analysis Parameter Estimates for MSEL Receptive Language Standard Scores for ASD Groups Only.

Fixed Effects Parameter Estimates

Parameter		Estimate	SE
Intercept		29.516 ***	2.280
Age		0.255 **	0.086
Diagnosis			
	EI-ASD	-6.899 **	2.280
	LI-ASD	-	-
Slope			
	Age x EI-ASD	0.135	0.086
	Age x LI-ASD	-	-

Random Effects Covariance Parameter Estimates

Parameter	Estimate	SE
Intercept Variance	162.440	73.128
Slope Variance	0.230	0.103
Intercept/Slope Covariance	-4.743	2.588

Note. Dashes indicate reference group.
* $p < .05$ **$p < .01$ *** $p < .001$.

Table 15. Growth Curve Analysis Parameter Estimates for MSEL Visual Reception Standard Scores for ASD Groups Only.

Fixed Effects Parameter Estimates

Parameter		Estimate	SE
Intercept		50.320 ***	2.395
Age		-0.231 *	0.100
Diagnosis			
	EI-ASD	-2.330	2.395
	LI-ASD	-	-
Slope			
	Age x EI-ASD	0.089	0.100
	Age x LI-ASD	-	-

Random Effects Covariance Parameter Estimates

Parameter	Estimate	SE
Intercept Variance	225.046	94.912
Slope Variance	0.488	0.158
Intercept/Slope Covariance	-9.616	3.787

Note. Dashes indicate reference group.
* $p < .05$ ** $p < .01$ *** $p < .001$.

Table 16. Growth Curve Analysis Parameter Estimates for VABS Socialization Domain Standard Scores.

Fixed Effects Parameter Estimates

Parameter		Estimate	SE
Intercept		99.686 ***	1.096
Age		-0.216 ***	0.047
Diagnosis			
	EI-ASD	-6.002 **	2.031
	LI-ASD	2.476	2.179
	LD	0.530	1.849
	TD	-	-
Slope			
	Age x EI-ASD	-0.125	0.083
	Age x LI-ASD	-0.341 ***	0.095
	Age x LD	0.144	0.080
	Age x TD	-	-

Random Effects Covariance Parameter Estimates

Parameter	Estimate	SE
Intercept Variance	37.221	24.249
Slope Variance	0.171	0.046
Intercept/Slope Covariance	-1.899	1.002

Note. Dashes indicate reference group.
* $p < .05$ ** $p < .01$ *** $p < .001$.

Table 17. Growth Curve Analysis Parameter Estimates for VABS Communication Domain Standard Scores.

Fixed Effects Parameter Estimates

Parameter		Estimate	SE
Intercept		79.827 ***	1.365
Age		0.411 ***	0.052
Diagnosis			
	EI-ASD	-11.314 ***	2.449
	LI-ASD	1.291	2.747
	LD	-5.575 *	2.321
	TD	-	-
Slope			
	Age x EI-ASD	0.021	0.091
	Age x LI-ASD	-0.287 **	0.104
	Age x LD	0.365 ***	0.089
	Age x TD	-	-

Random Effects Covariance Parameter Estimates

Parameter	Estimate	SE
Intercept Variance	223.234	37.794
Slope Variance	0.323	0.054
Intercept/Slope Covariance	-7.133	1.366

Note. Dashes indicate reference group.
* $p < .05$ **$p < .01$ *** $p < .001$.

Table 18. Growth Curve Analysis Parameter Estimates for VABS Daily Living Skills Domain Standard Scores.

Fixed Effects Parameter Estimates

Parameter		Estimate	SE
Intercept		92.839 ***	1.230
Age		0.064	0.048
Diagnosis			
	EI-ASD	-2.025	2.253
	LI-ASD	3.890	2.454
	LD	-4.261 *	2.077
	TD	-	-
Slope			
	Age x EI-ASD	-0.227 **	0.085
	Age x LI-ASD	-0.346 ***	0.097
	Age x LD	0.351 ***	0.082
	Age x TD	-	-

Random Effects Covariance Parameter Estimates

Parameter	Estimate	SE
Intercept Variance	82.94	30.974
Slope Variance	0.153	0.048
Intercept/Slope Covariance	-2.355	1.147

Note. Dashes indicate reference group.
* $p < .05$ **$p < .01$ *** $p < .001$.

Table 19. Number of Items on the VABS 1 and 2 by Domain and Subdomain.

Domain and Subdomain	VABS 1	VABS 2
Communication	67	99
Receptive	20	13
Expressive	54	31
Written	25	23
Daily Living Skills	92	109
Personal	41	39
Domestic	24	21
Community	44	32
Socialization	66	99
Interpersonal Relationships	38	28
Play and Leisure Time	31	20
Coping Skills	30	18
Motor Skills	36	76
Gross	40	20
Fine	36	16
Adaptive Behavior Composite	261	383

Note. Information provided in the VABS 2 manual (Sparrow, Cicchetti, & Balla, 2005).

Table 20. Correlations Between the VABS 1 and 2.

Ages 0-2 (n=24)	VABS 1		VABS 2		Correlations	
	Mean	SD	Mean	SD	*r*	Adj *r*
Communication	94.2	17.0	97.4	13.5	0.65	0.69
Daily Living Skills	94.4	15.1	94.0	14.5	0.75	0.76
Socialization	97.0	13.9	95.8	12.4	0.85	0.89
Motor Skills	94.9	11.7	99.7	13.1	0.91	0.93
Adaptive Behavior Composite	93.5	15.4	95.9	12.3	0.82	0.87

Ages 3-6 (n=29)	VABS 1		VABS 2		Correlations	
	Mean	SD	Mean	SD	*r*	Adj *r*
Communication	86.6	15.6	95.6	12.6	0.86	0.89
Daily Living Skills	78.2	14.6	87.3	9.6	0.90	0.96
Socialization	89.7	18.5	93.3	13.3	0.94	0.95
Motor Skills	88.8	14.3	94.2	12.0	0.86	0.90
Adaptive Behavior Composite	82.7	18.5	91.2	12.6	0.91	0.93

Note. Correlations provided in the VABS 2 manual (Sparrow, Cicchetti, & Balla, 2005).

Table 21. Demographic Information for Children Tested with the VABS 1 and VABS 2.

	VABS 1	VABS 2
Age	19.11 (6.33)	25.85 (8.08)
Gender		
Female	20.55%	30.00%
Male	79.45%	70.00%
Ethnicity		
Hispanic or Latino	21.92%	25.06%
Not Hispanic or Latino	75.34%	74.35%
Not Reported	2.74%	0.59%
Race		
African American	0%	2.94%
American Indian/Alaska Native	0%	0.71%
Asian	6.85%	5.29%
Caucasian	82.19%	75.29%
Pacific Islander/Native Hawaiian	1.37%	1.06%
Multiple Races Selected	6.85%	11.29%
Not Reported	2.74%	3.41%

Table 22. Growth Curve Analysis Parameter Estimates for VABS Socialization Domain Standard Scores for ASD Groups Only.

Fixed Effects Parameter Estimates

Parameter		Estimate	SE
Intercept		97.965 ***	1.928
Age		-0.450 ***	0.078
Diagnosis			
	EI-ASD	-4.327 *	1.928
	LI-ASD	-	-
Slope			
	Age x EI-ASD	0.112	0.078
	Age x LI-ASD	-	-

Random Effects Covariance Parameter Estimates

Parameter	Estimate	SE
Intercept Variance	31.909	57.754
Slope Variance	0.130	0.100
Intercept/Slope Covariance	-1.319	2.326

Note. Dashes indicate reference group.
* $p < .05$ **$p < .01$ *** $p < .001$

Table 23. Growth Curve Analysis Parameter Estimates for VABS Communication Domain Standard Scores for ASD Groups Only.

Fixed Effects Parameter Estimates

Parameter		Estimate	SE
Intercept		74.942 ***	2.475
Age		0.275 **	0.097
Diagnosis			
	EI-ASD	-6.150 *	2.475
	LI-ASD	-	-
Slope			
	Age x EI-ASD	0.148	0.097
	Age x LI-ASD	-	-

Random Effects Covariance Parameter Estimates

Parameter	Estimate	SE
Intercept Variance	255.486	83.432
Slope Variance	0.446	0.132
Intercept/Slope Covariance	-8.95	3.191

Note. Dashes indicate reference group.
* $p < .05$ ** $p < .01$ *** $p < .001$

Table 24. Growth Curve Analysis Parameter Estimates for VABS Daily Living Skills Domain Standard Scores for ASD Groups Only.

Fixed Effects Parameter Estimates

Parameter		Estimate	SE
Intercept		93.791 ***	2.220
Age		-0.225 **	0.085
Diagnosis			
	EI-ASD	-2.729	2.220
	LI-ASD	-	-
Slope			
	Age x EI-ASD	0.053	0.085
	Age x LI-ASD	-	-

Random Effects Covariance Parameter Estimates

Parameter	Estimate	SE
Intercept Variance	186.606	69.719
Slope Variance	0.292	0.104
Intercept/Slope Covariance	-5.795	2.537

Note. Dashes indicate reference group.
* $p < .05$ ** $p < .01$ *** $p < .001$

FIGURES

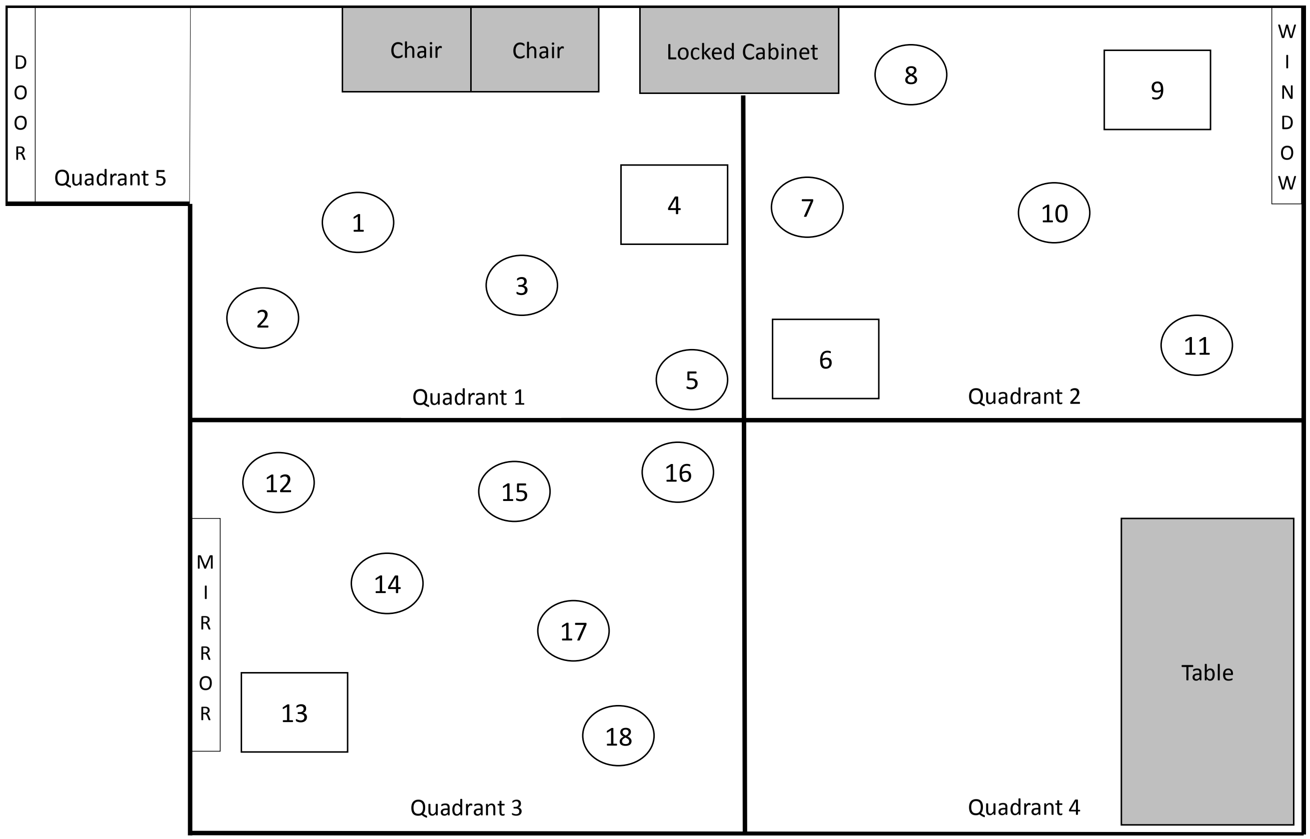

Figure 1. Layout of the exploration paradigm. Standardized locations of items are depicted and description of items are as follows: 1) ball of string, 2) glitter wand, 3) ball with two small nets, 4) shoe box that contains a stretch tube that emits noise, 5) picture book, 6) shoe box that contains a piece of purple cloth, 7) Slinky®, 8) toy tuck, 9) large plastic container with lid that contains an expanding sphere toy, 10) two lizards, 11) toy boat, 12) girl doll and boy doll, 13) clear plastic container with lid that contains a piece of printed cloth, 14) rhino, 15) multi-colored cube with spinning wheels, 16) flat gold circle, 17) Elmo® keys that emit sound effects, 18) Koosh® ball.

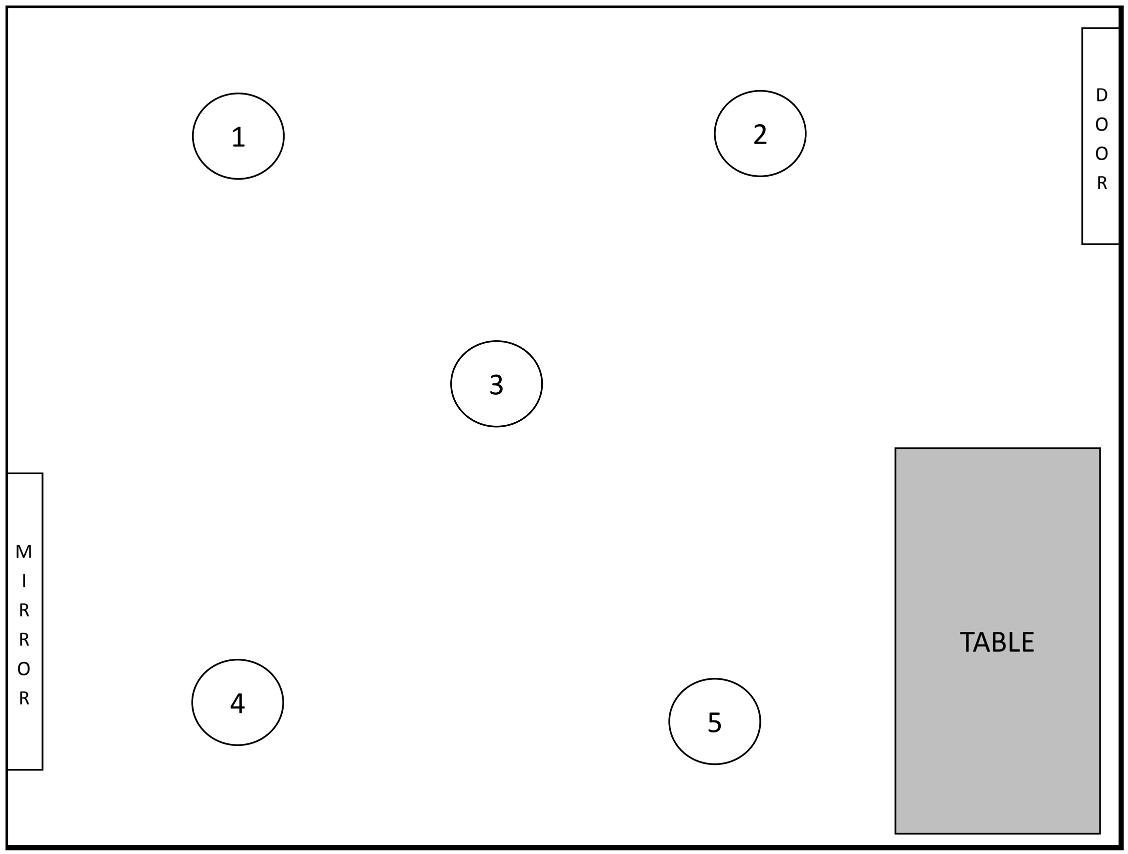

Figure 2. Layout of the PCI paradigm. Standardized locations of items are depicted and description of items are as follows: 1) pirate hat, cowboy hat, tiara, crown, hand puppets, 2) two nets with balls, football, 3) drums, toy radio, maracas, two cans connected by a string, toy microphone, 4) baby doll, stroller, two bottles, 5) small picture book, Mr. Potato Head®, Thomas the Train®, and Raggedy Ann® doll.

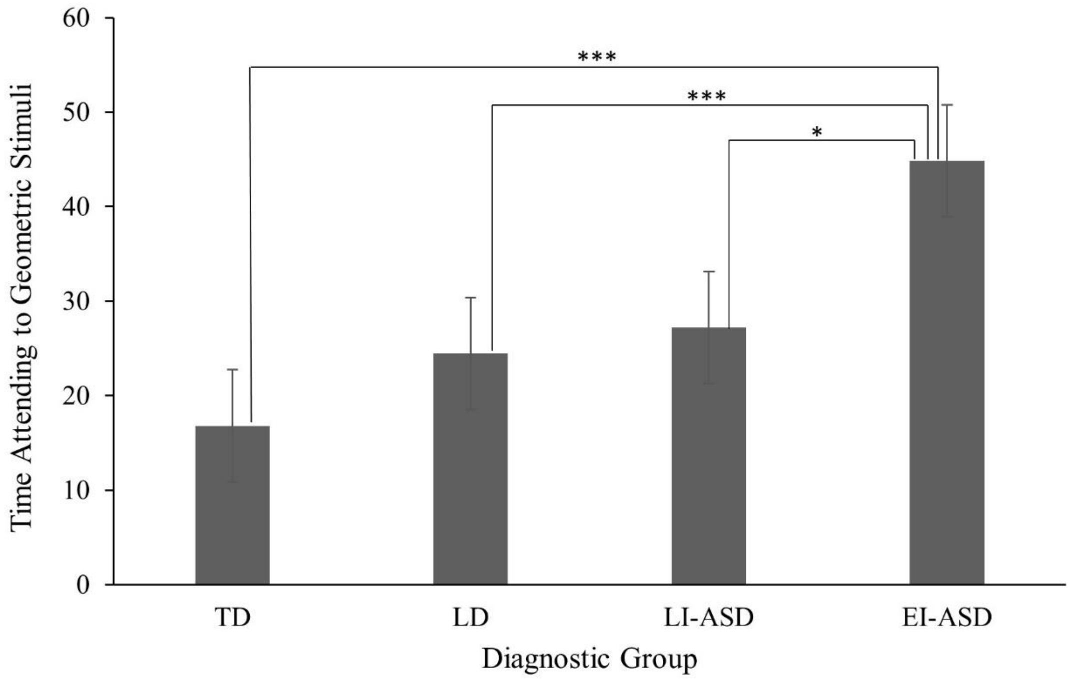

Figure 3. Average percentage of time looking at the geometric stimuli during the geometric preference test across diagnostic groups.
* p < .05, ** p < .01, *** p < .001

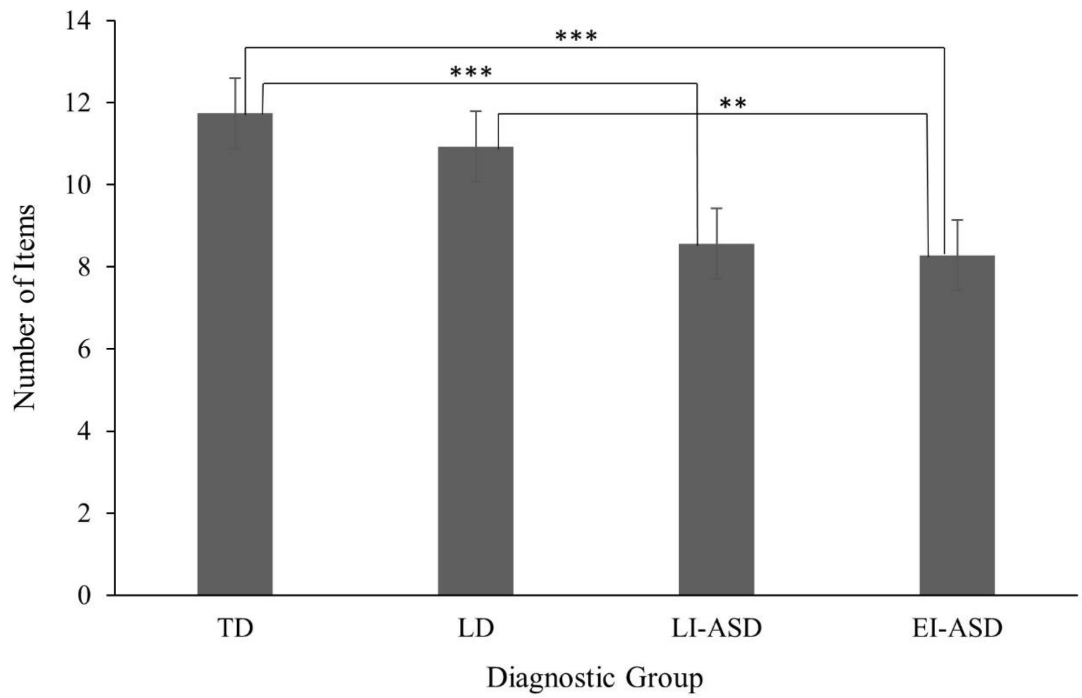

Figure 4. Average number of items explored during the exploration task across diagnostic groups.
* p < .05, ** p < .01, *** p < .001

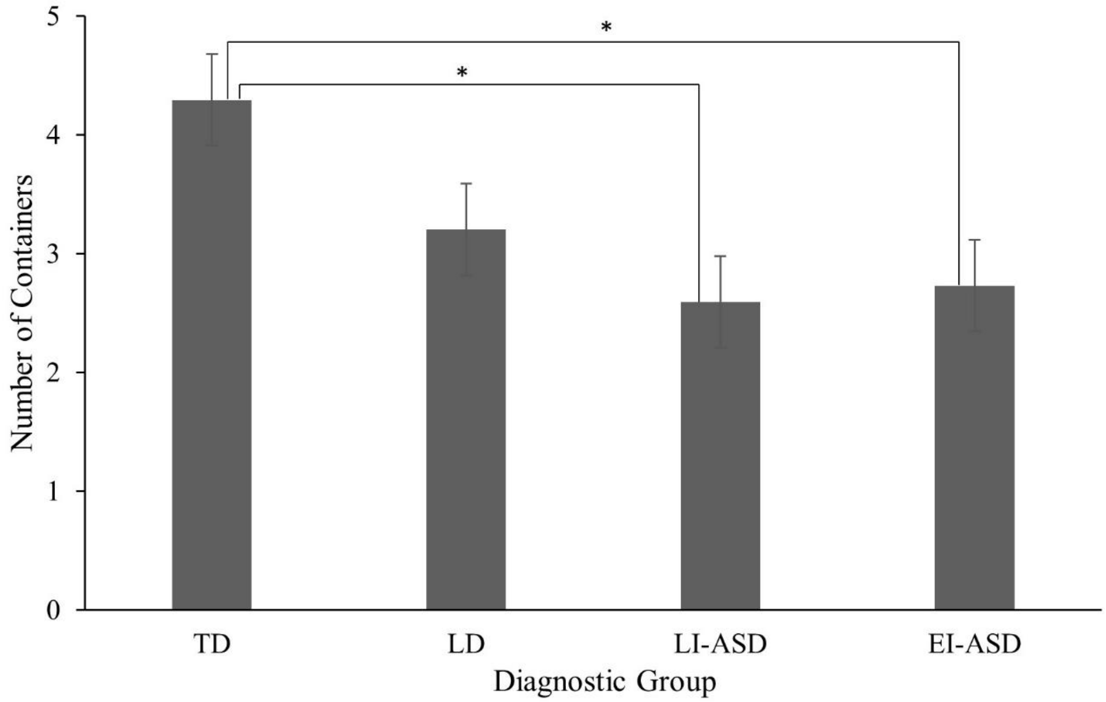

Figure 5. Average number of containers explored during the exploration task across diagnostic groups.
* p < .05, ** p < .01, *** p < .001

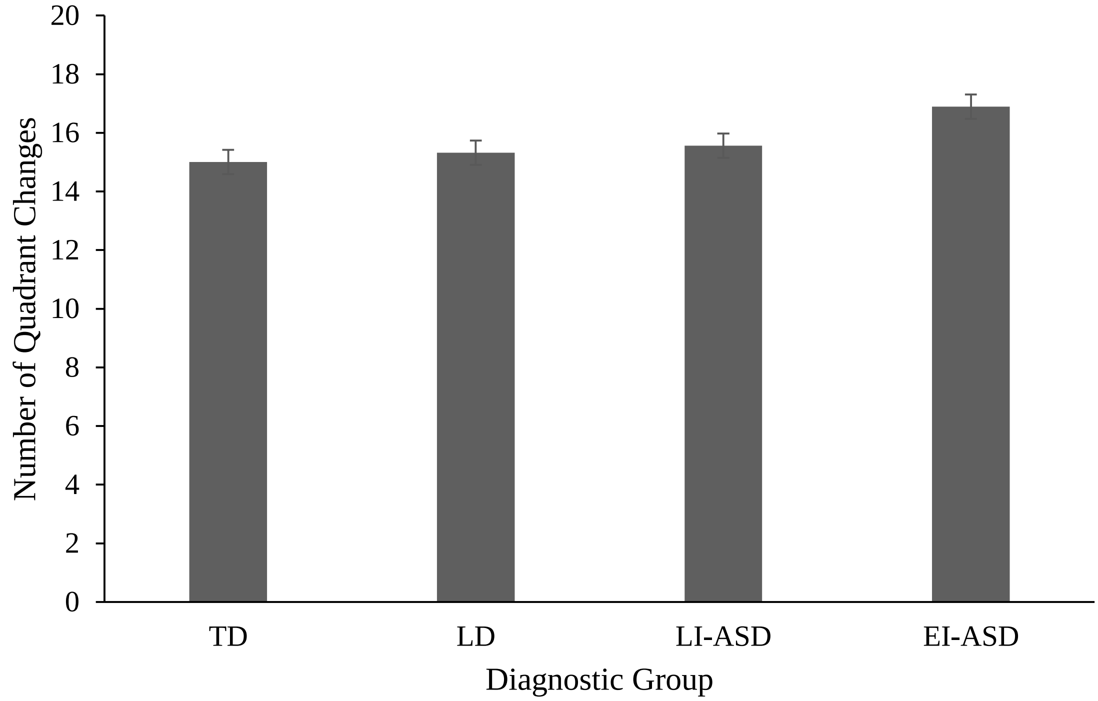

Figure 6. Amount of movement about the observation room during the exploration task across diagnostic groups. Depicted by the average number of quadrant changes
* $p < .05$, ** $p < .01$, *** $p < .001$

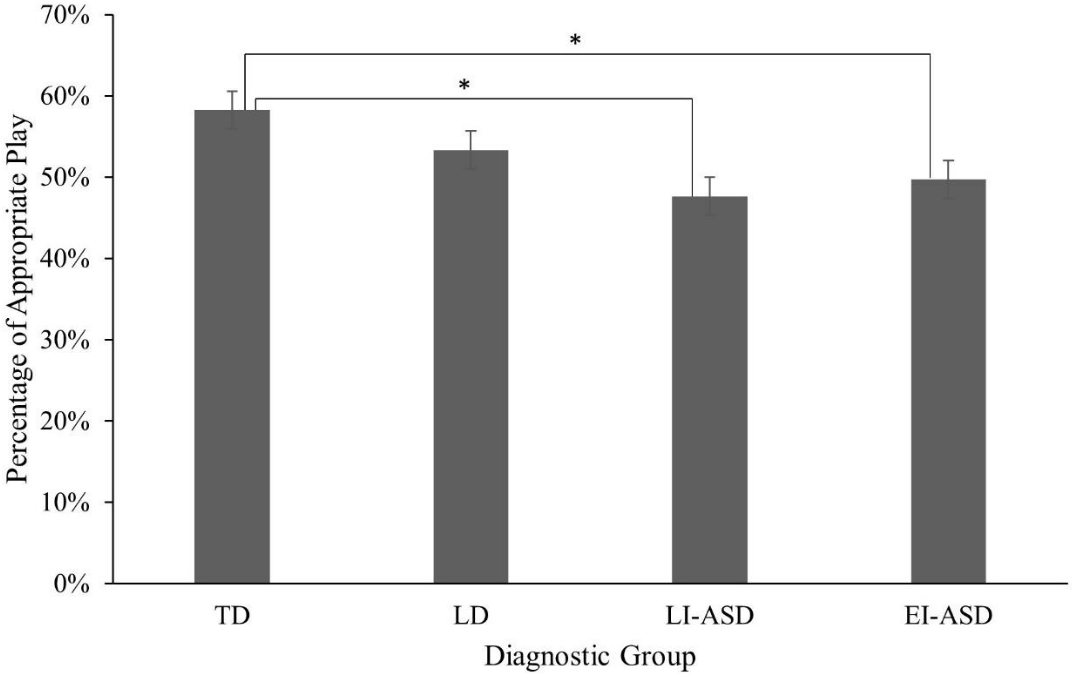

Figure 7. Average percentage of appropriate exploration and play during the exploration task across diagnostic groups.
* p < .05, ** p < .01, *** p < .001

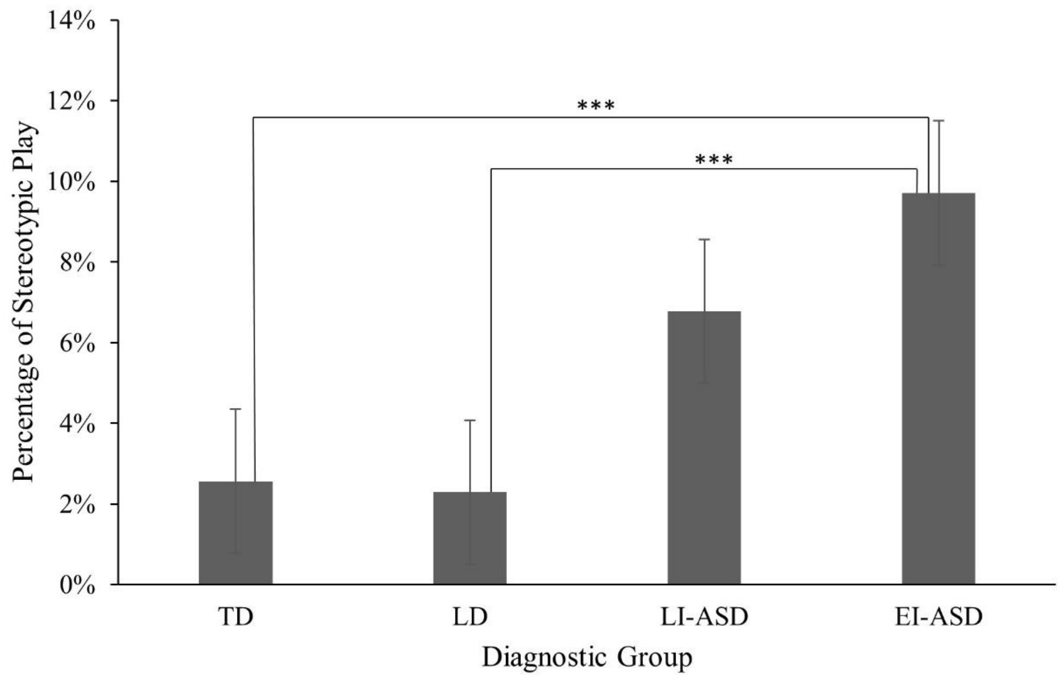

Figure 8. Average percentage of stereotypic exploration and play during the exploration task across diagnostic groups.
* p < .05, ** p < .01, *** p < .001

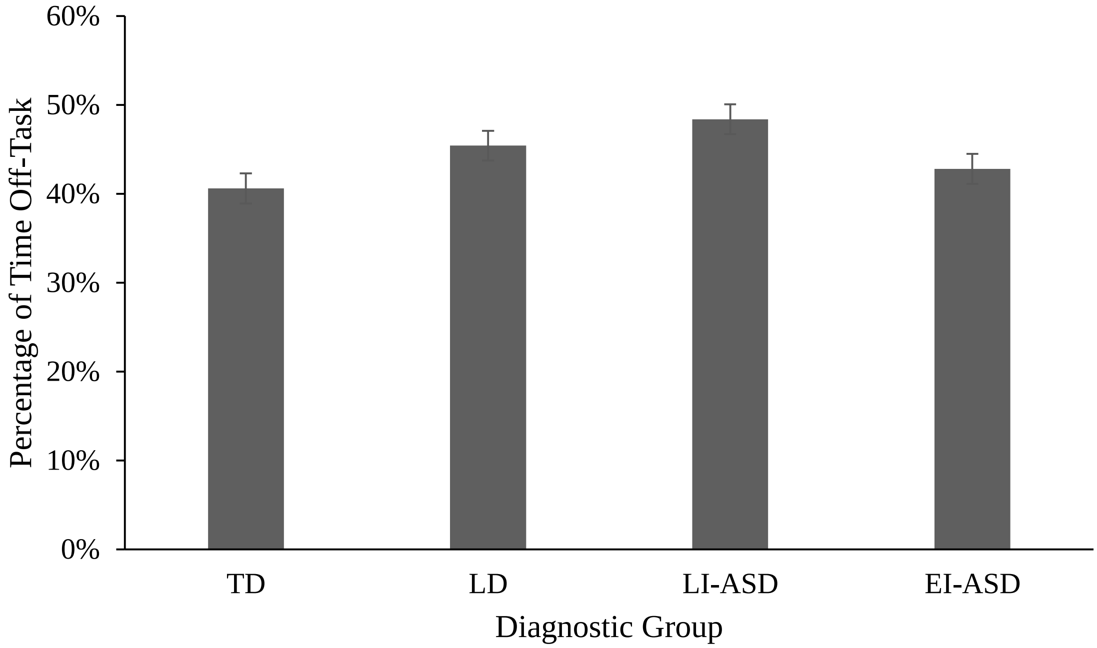

Figure 9. Average percentage of off-task behavior during the exploration task across diagnostic groups.
* p < .05, ** p < .01, *** p < .001

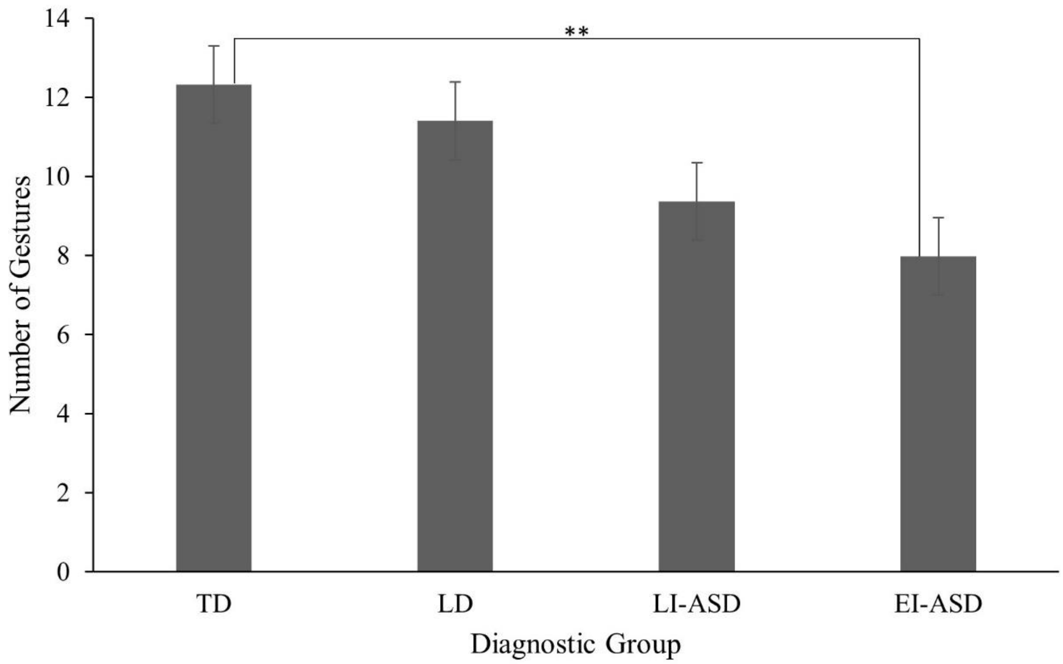

Figure 10. Average number of all gestures during the PCI across diagnostic groups. Includes reaching, pointing, showing, and giving gestures.
* p < .05, ** p < .01, *** p < .001

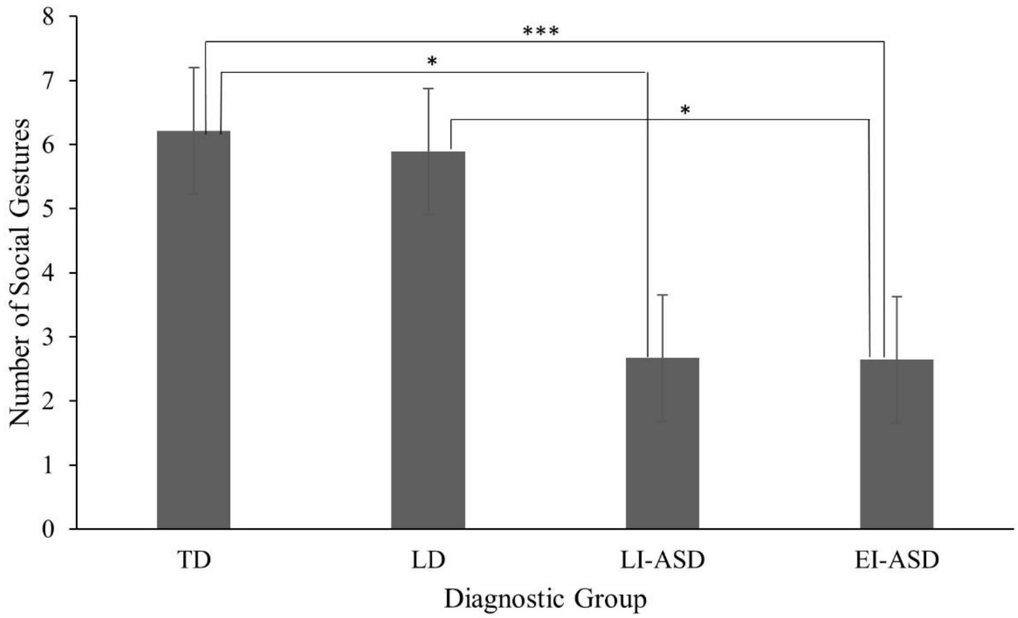

Figure 11. Average number of social gestures during the PCI across diagnostic groups. Includes pointing, showing, and giving gestures.
* p < .05, ** p < .01, *** p < .001

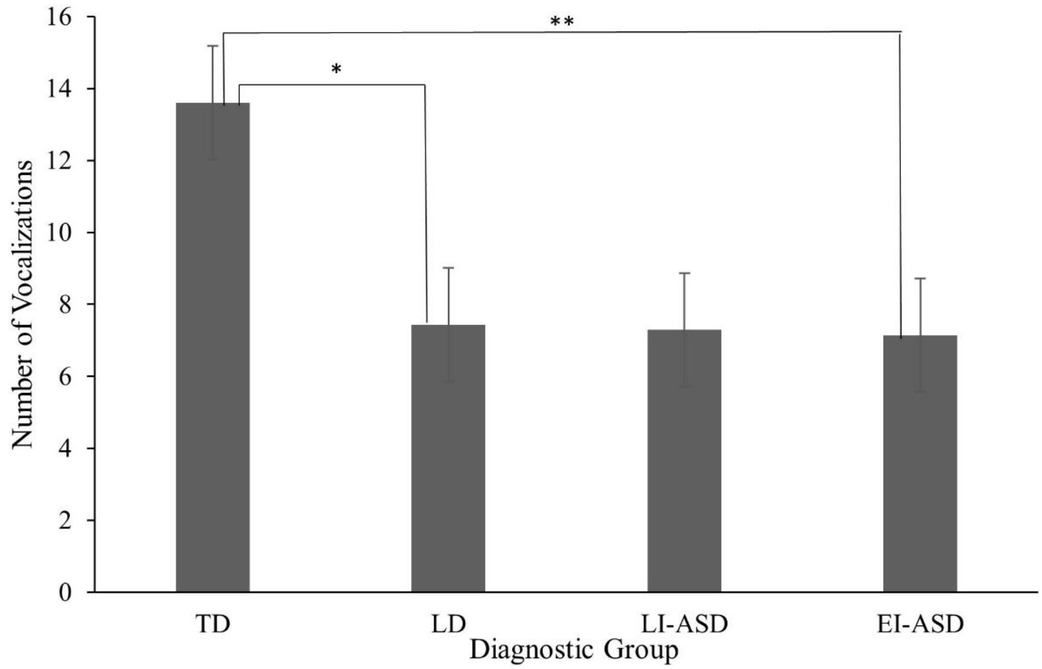

Figure 12. Average number of social vocalizations during the PCI across diagnostic groups.
* p < .05, ** p < .01, *** p < .001

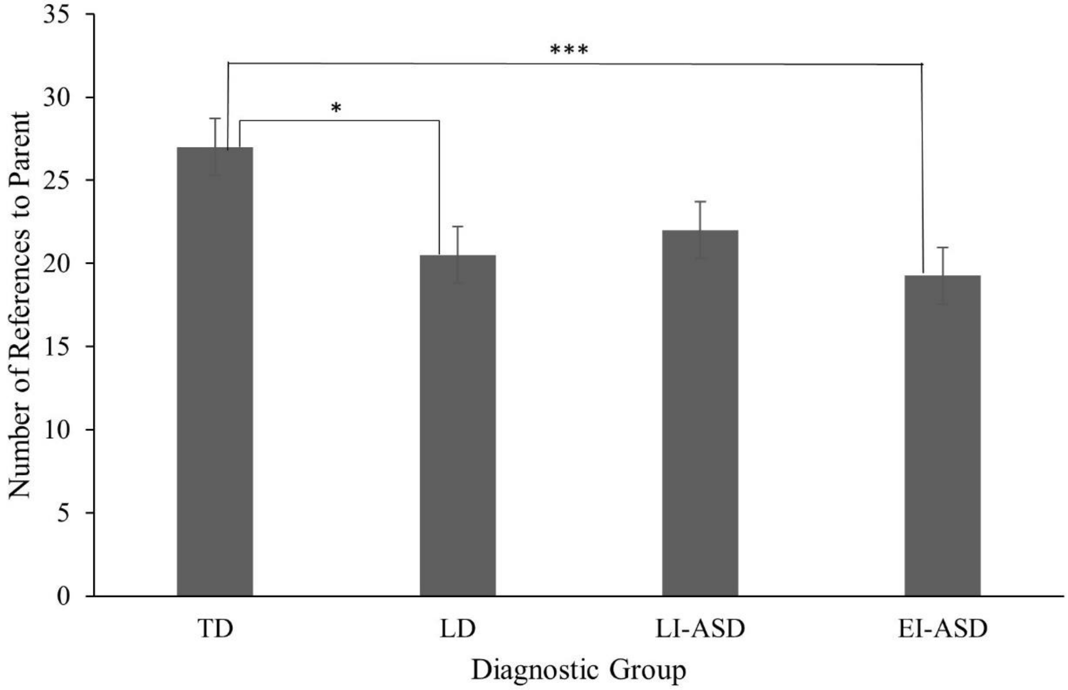

Figure 13. Average number of social references toward parent during the PCI across diagnostic groups.
* $p < .05$, ** $p < .01$, *** $p < .001$

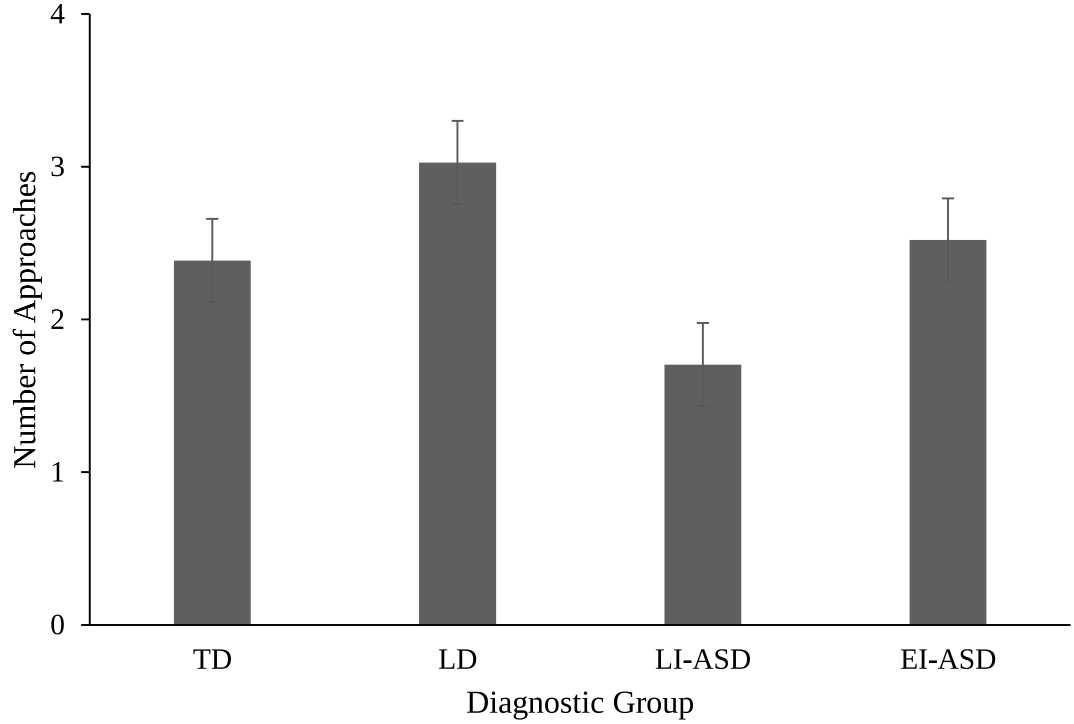

Figure 14. Average number of approaches toward parent during the PCI across diagnostic groups.
* p < .05, ** p < .01, *** p < .001

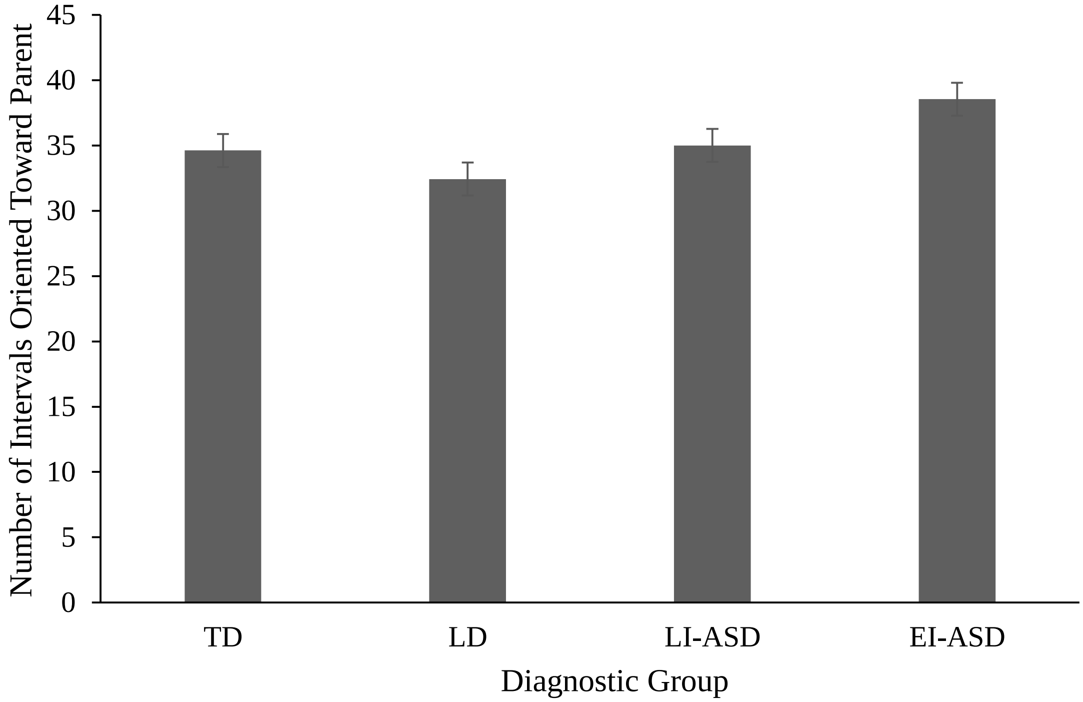

Figure 15. Average number of intervals the child was oriented toward parent during the PCI across diagnostic groups.
* $p < .05$, ** $p < .01$, *** $p < .001$

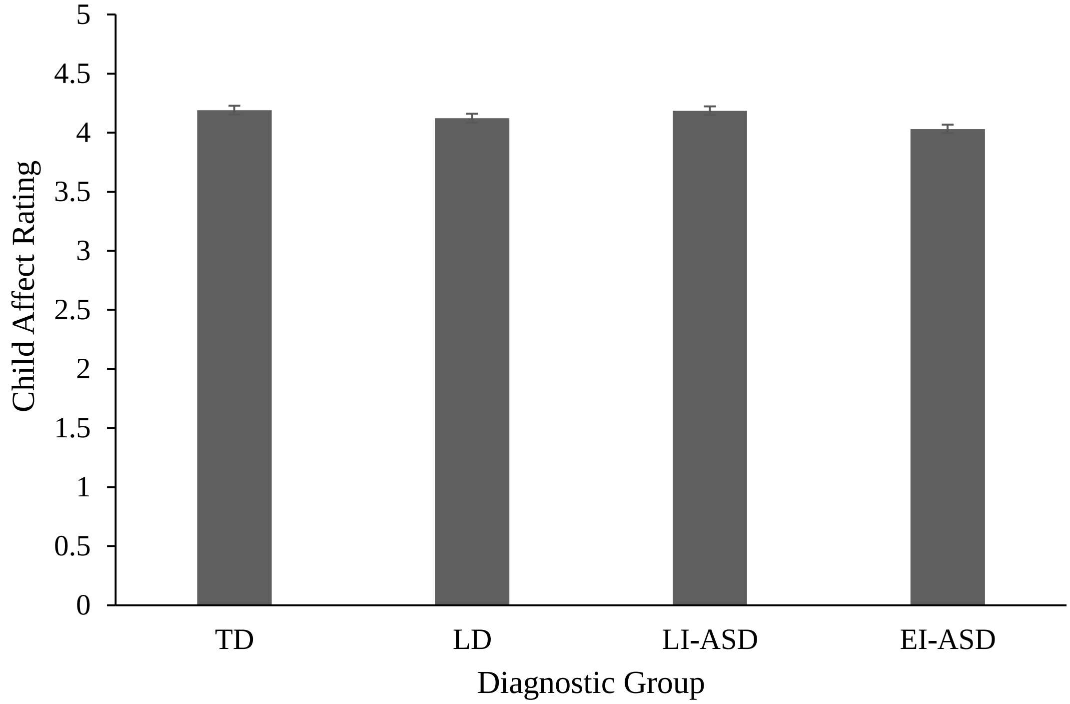

Figure 16. Average rating of child affect during the PCI across diagnostic groups.
* p < .05, ** p < .01, *** p < .001

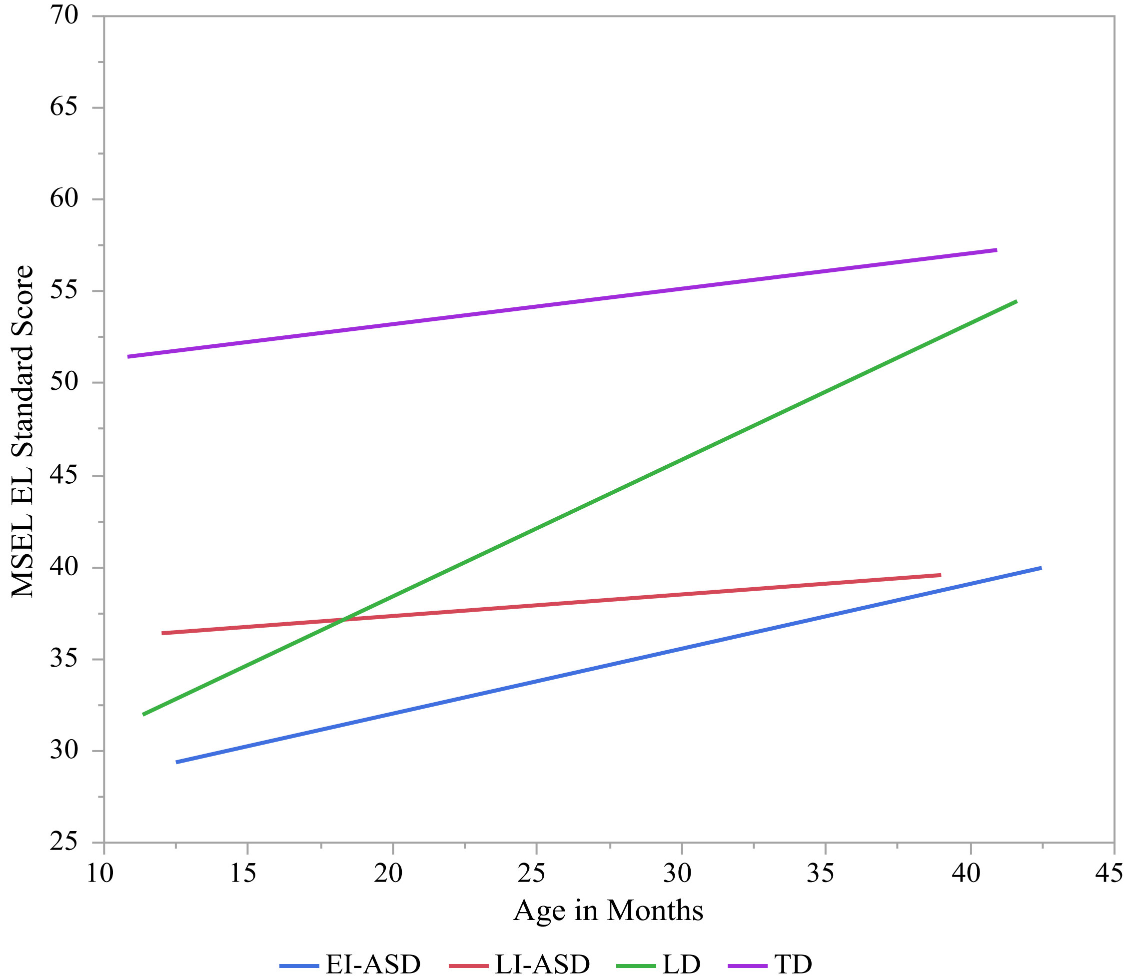

Figure 17. Trajectories of development of the MSEL expressive language domain across diagnostic groups. Standard scores are depicted.

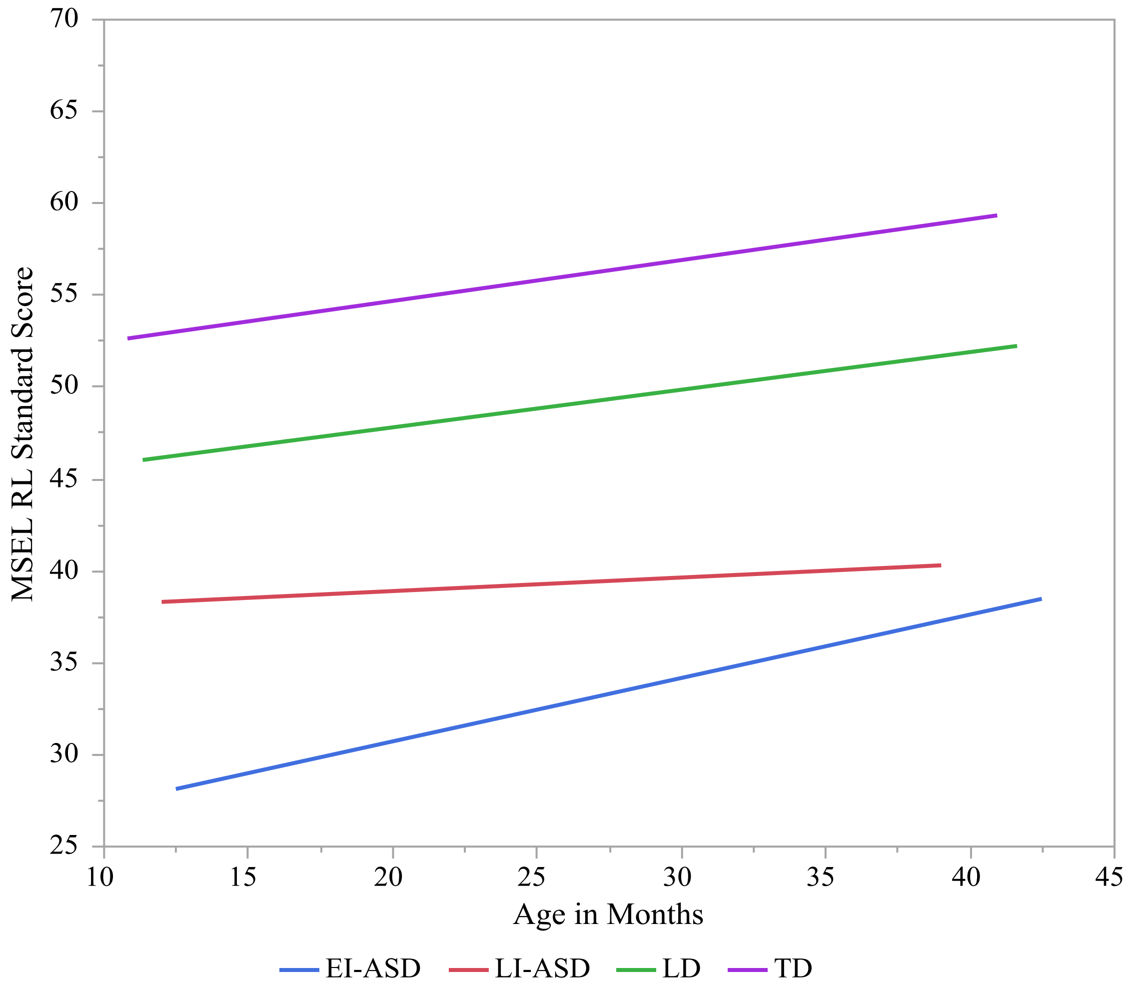

Figure 18. Trajectories of development of the MSEL receptive language domain across diagnostic groups. Standard scores are depicted.

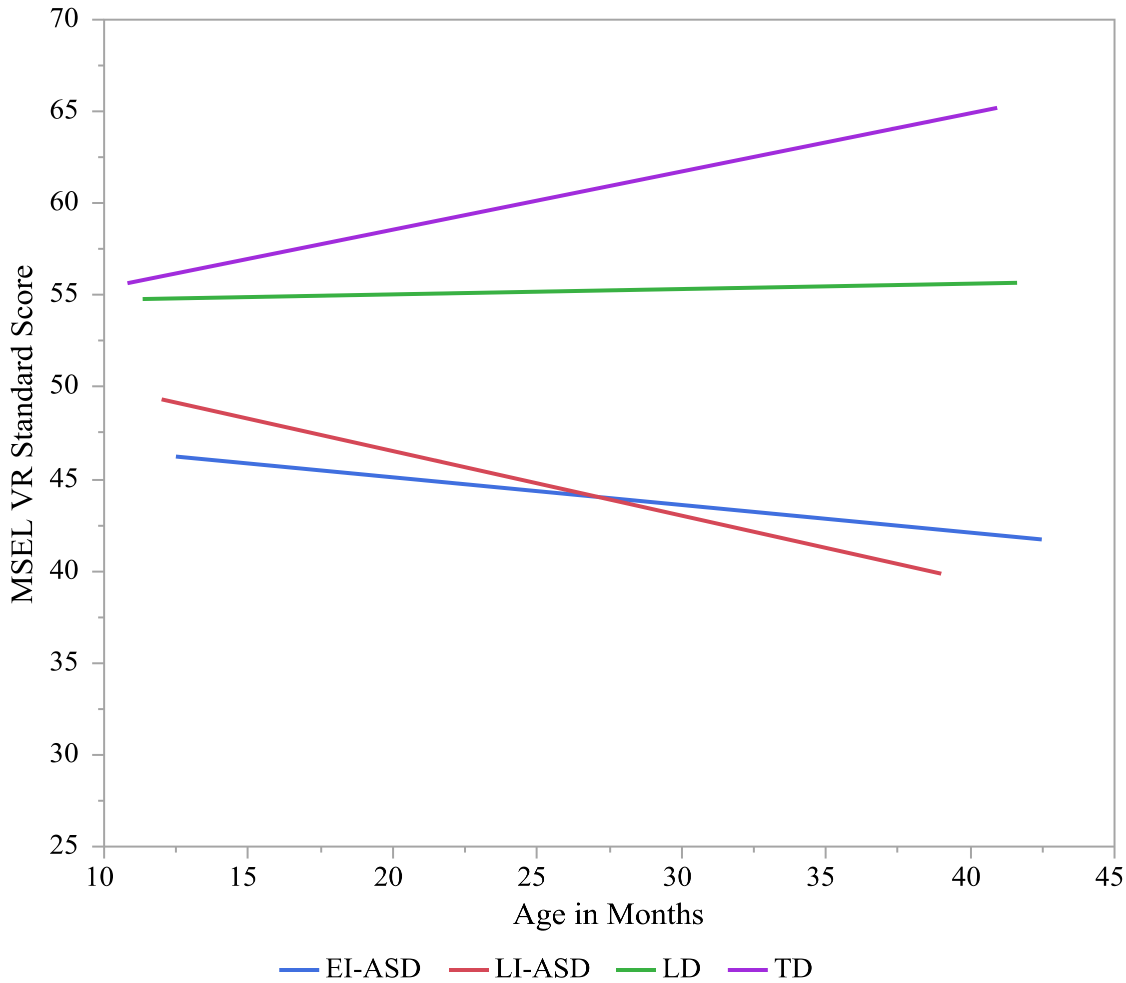

Figure 19. Trajectories of development of the MSEL visual reception domain across diagnostic groups. Standard scores are depicted.

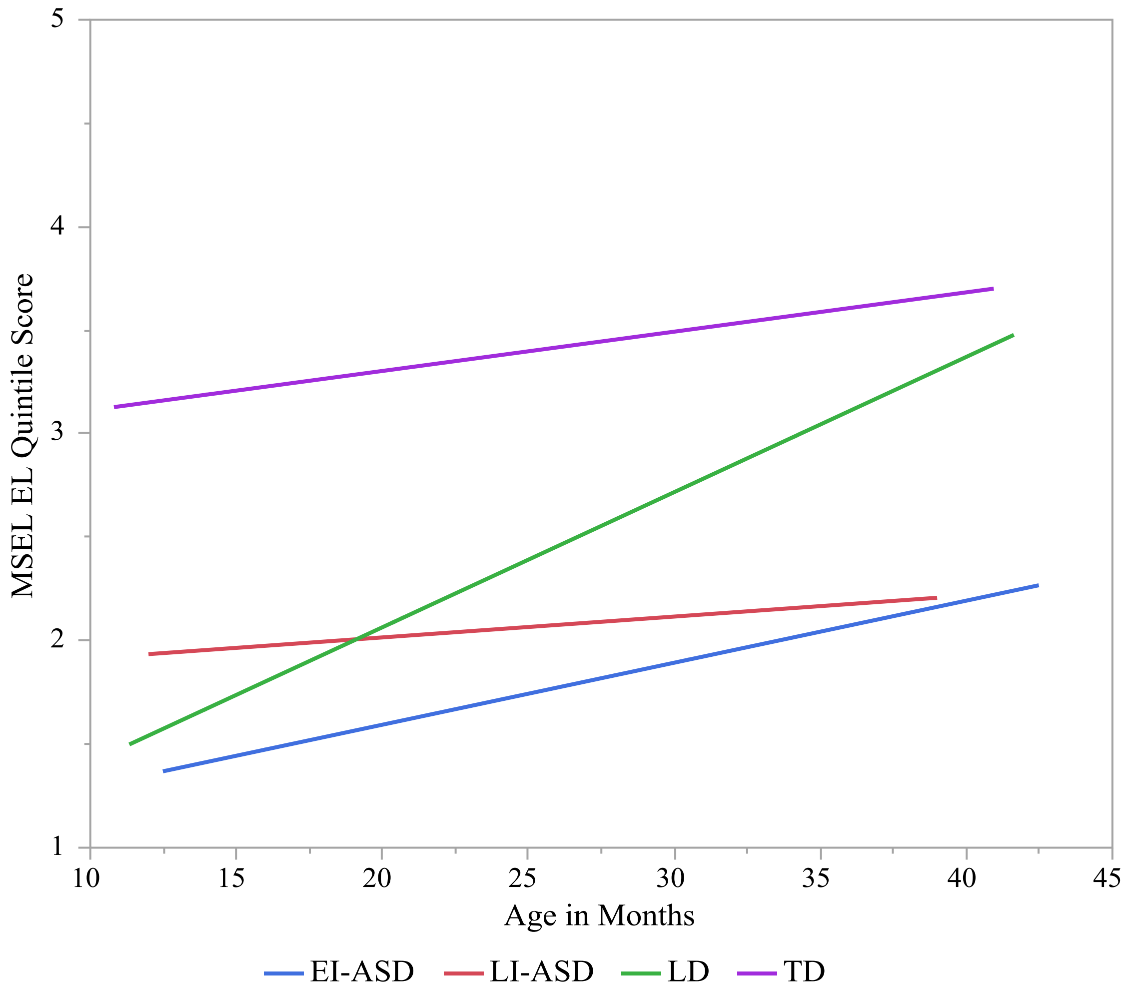

Figure 20. Trajectories of development of the MSEL expressive language domain across diagnostic groups using quintile scoring.

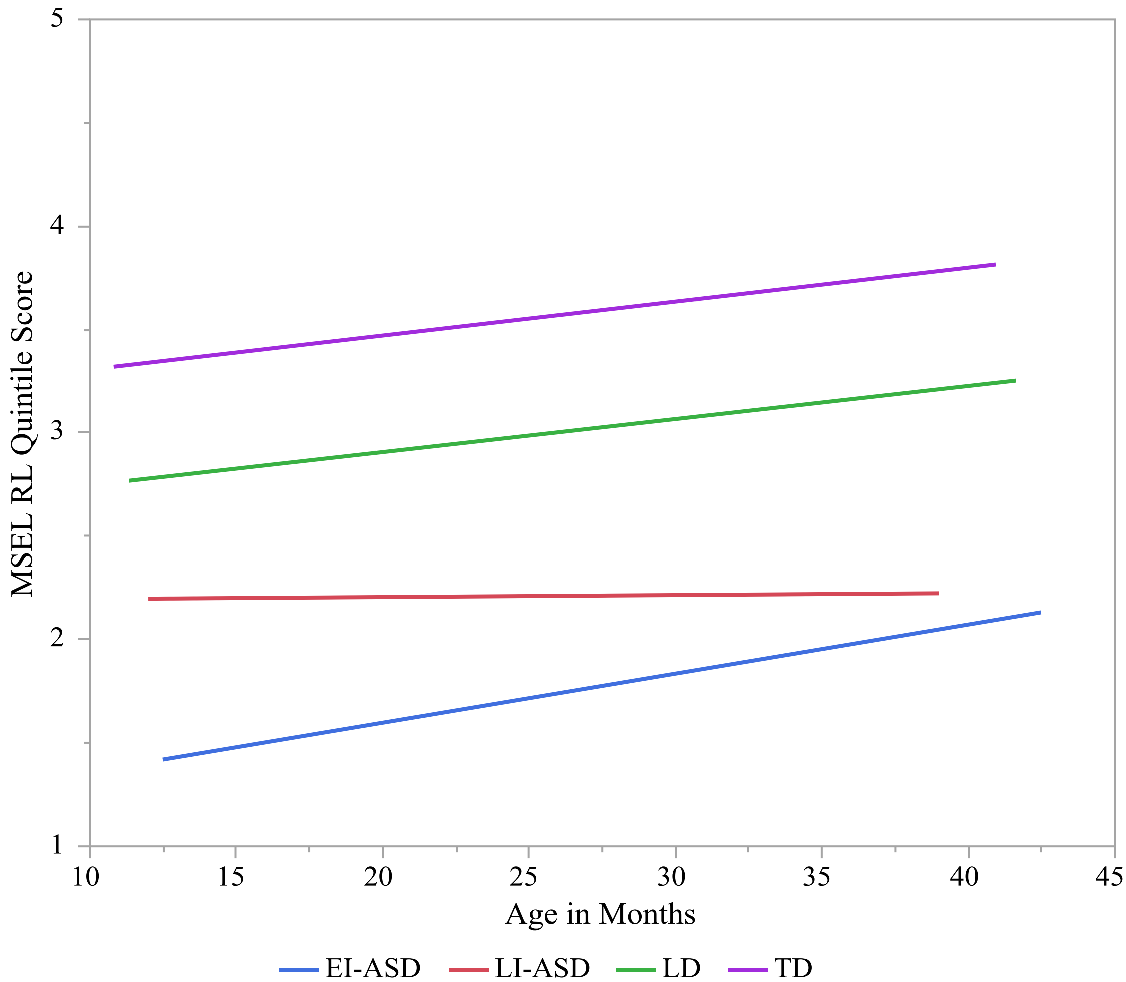

Figure 21. Trajectories of development of the MSEL receptive language domain across diagnostic groups using quintile scoring.

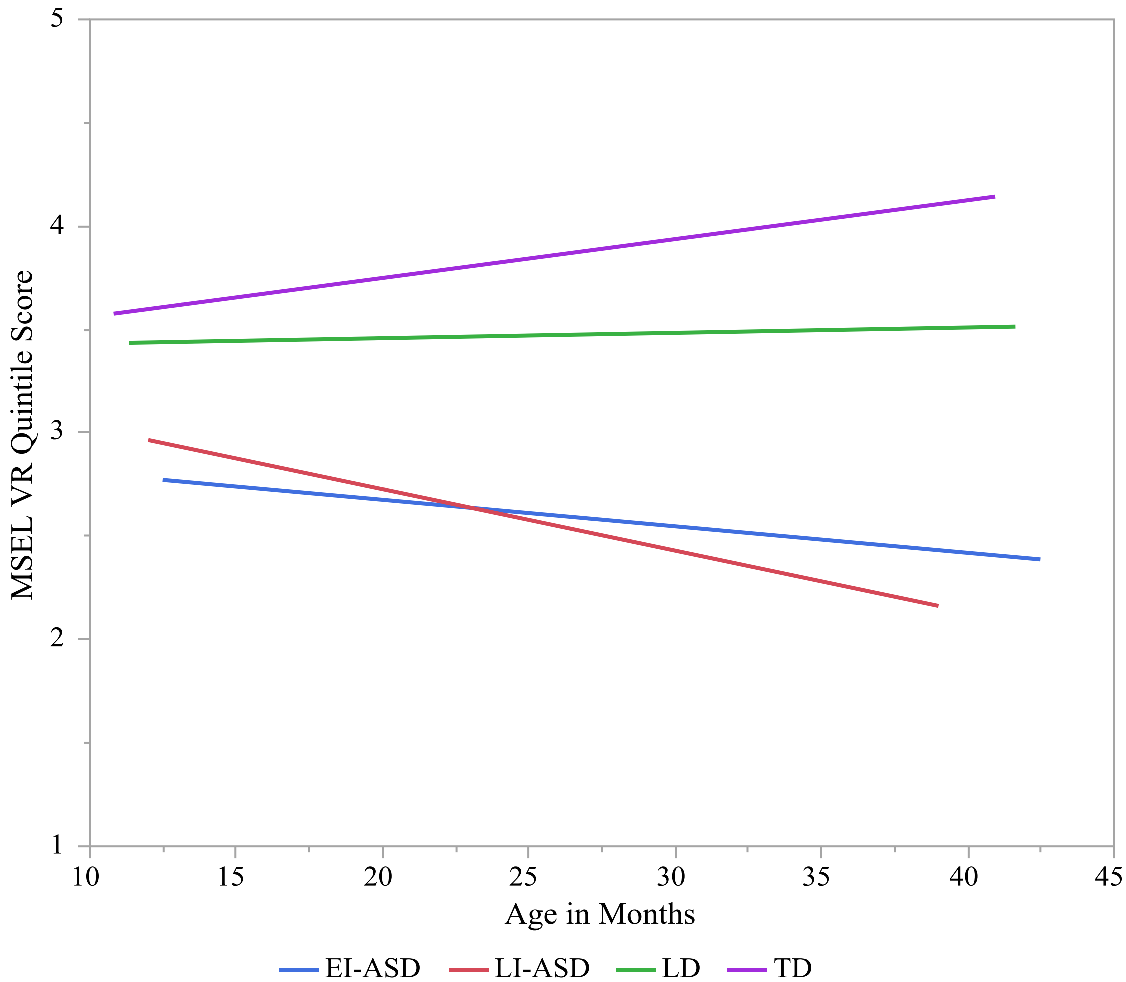

Figure 22. Trajectories of development of the MSEL visual reception domain across diagnostic groups using quintile scoring.

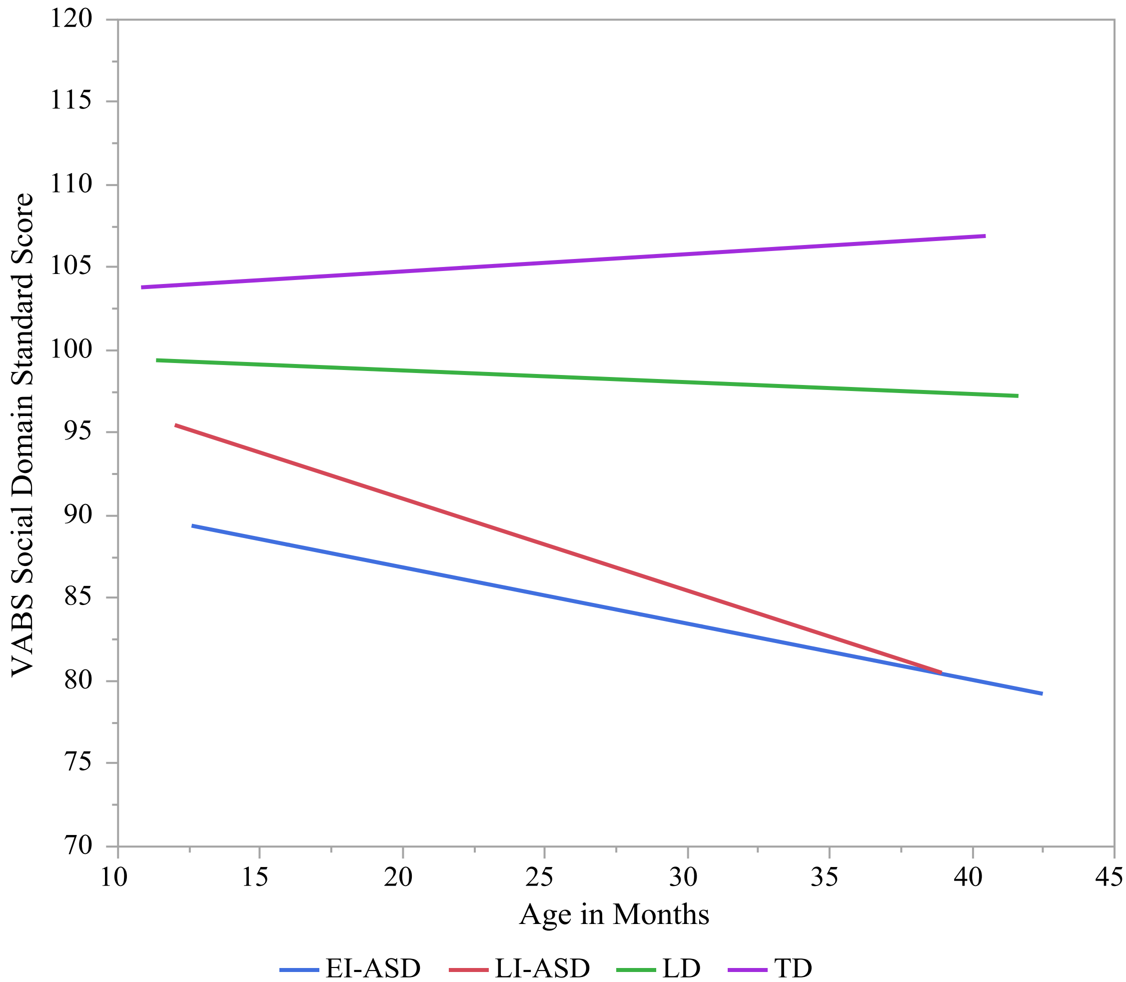

Figure 23. Trajectories of development of the VABS socialization domain across diagnostic groups. Standard scores are depicted.

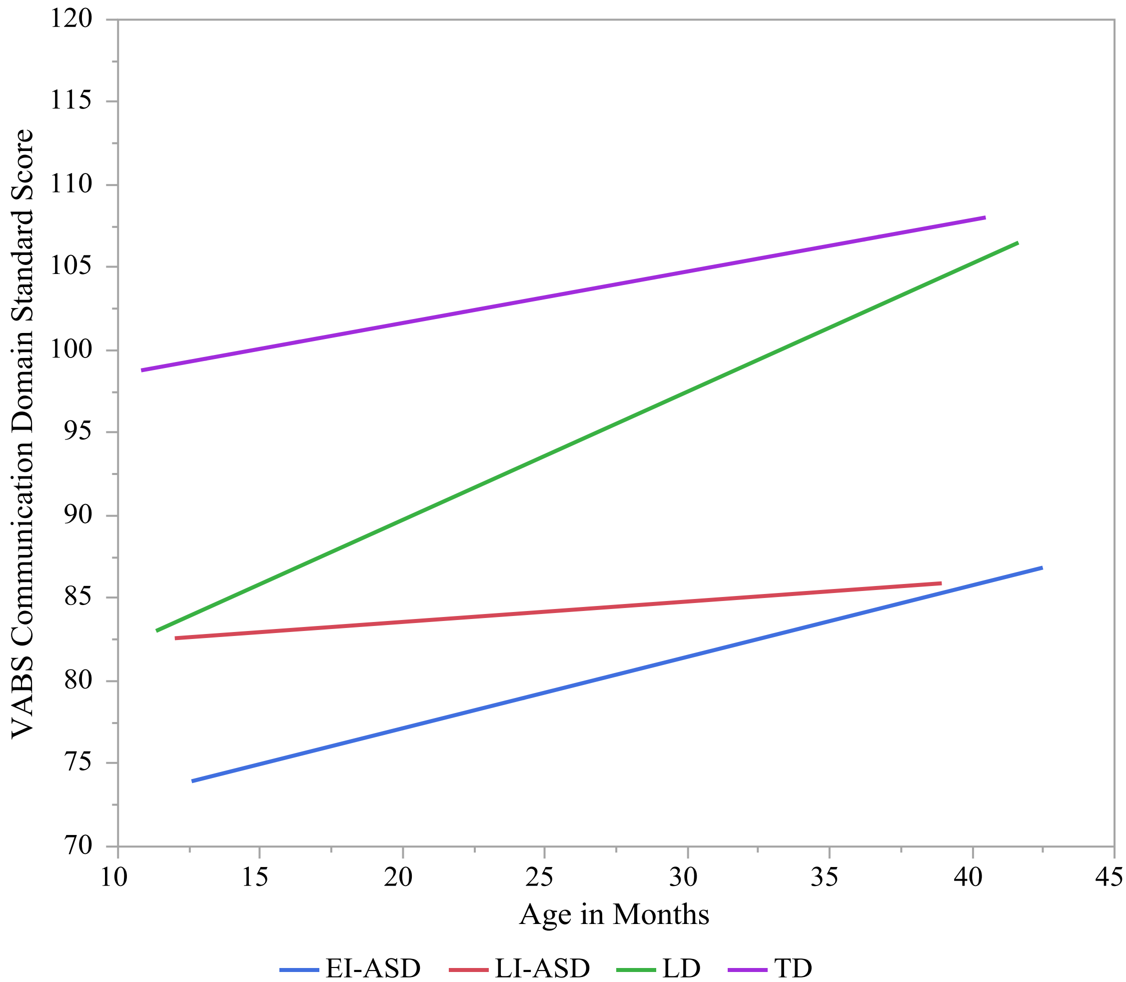

Figure 24. Trajectories of development of the VABS communication domain across diagnostic groups. Standard scores are depicted.

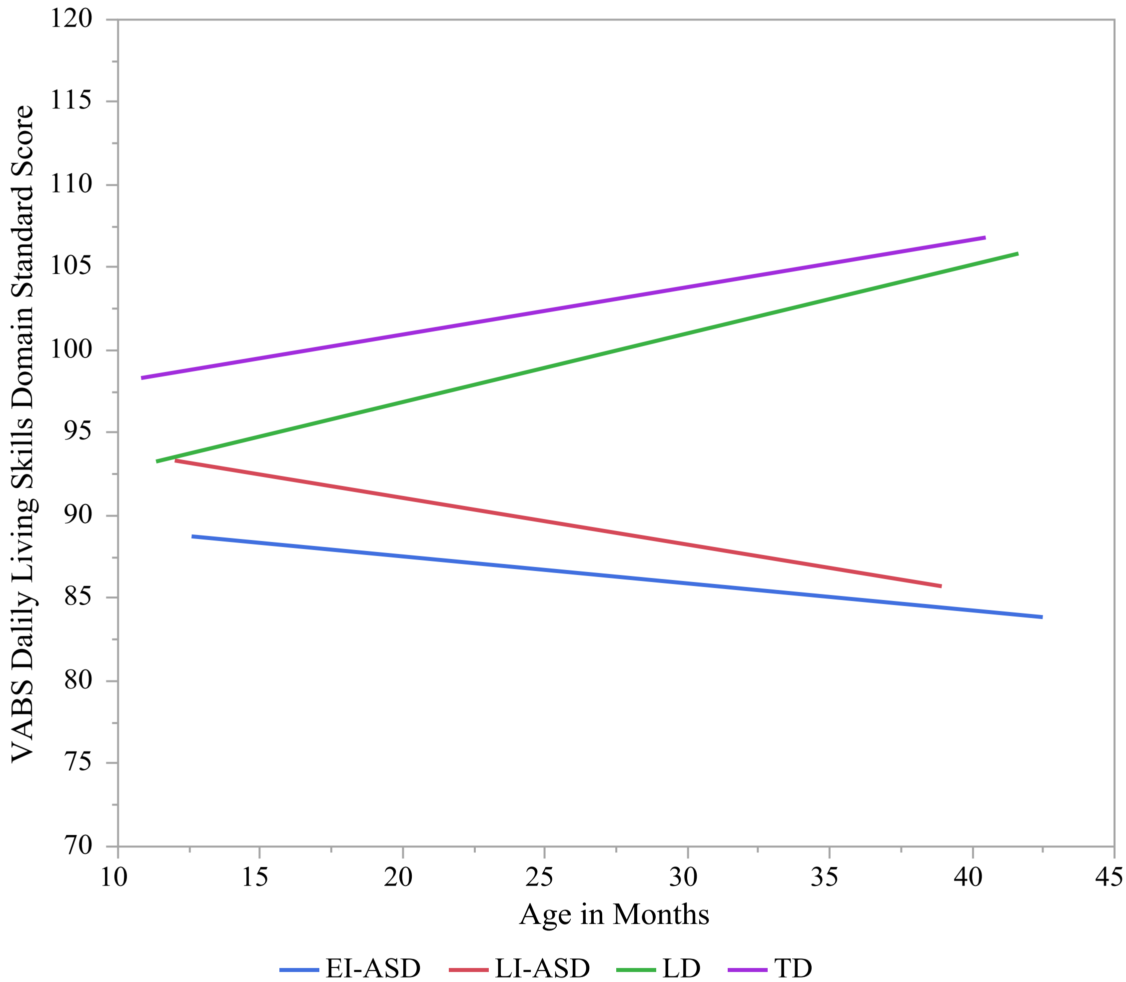

Figure 25. Trajectories of development of the VABS daily living skills domain across diagnostic groups. Standard scores are depicted.

APPENDICES

Appendix A. CSBS Screener

CSBS DP

CSBS DP Infant-Toddler Checklist

Child's name: ______________ *Date of birth:* ______________ *Date filled out:* ______________

Was birth premature? ______________ *If yes, how many weeks premature?* ______________

Filled out by: ______________ *Relationship to child:* ______________

Instructions for caregivers: This Checklist is designed to identify different aspects of development in infants and toddlers. Many behaviors that develop before children talk may indicate whether or not a child will have difficulty learning to talk. This Checklist should be completed by a caregiver when the child is between **6 and 24 months of age** to determine whether a referral for an evaluation is needed. The caregiver may be either a parent or another person who nurtures the child daily. Please check all the choices that best describe your child's behavior. If you are not sure, please choose the closest response based on your experience. **Children at your child's age are not necessarily expected to use all the behaviors listed.**

Emotion and Eye Gaze					
1. Do you know when your child is happy and when your child is upset?		☐ Not Yet	☐ Sometimes	☐ Often	
2. When your child plays with toys, does he/she look at you to see if you are watching?		☐ Not Yet	☐ Sometimes	☐ Often	
3. Does your child smile or laugh while looking at you?		☐ Not Yet	☐ Sometimes	☐ Often	
4. When you look at and point to a toy across the room, does your child look at it?		☐ Not Yet	☐ Sometimes	☐ Often	
Communication					
5. Does your child let you know that he/she needs help or wants an object out of reach?		☐ Not Yet	☐ Sometimes	☐ Often	
6. When you are not paying attention to your child, does he/she try to get your attention?		☐ Not Yet	☐ Sometimes	☐ Often	
7. Does your child do things just to get you to laugh?		☐ Not Yet	☐ Sometimes	☐ Often	
8. Does your child try to get you to notice interesting objects—just to get you to look at the objects, not to get you to do anything with them?		☐ Not Yet	☐ Sometimes	☐ Often	
Gestures					
9. Does your child pick up objects and give them to you?		☐ Not Yet	☐ Sometimes	☐ Often	
10. Does your child show objects to you without giving you the object?		☐ Not Yet	☐ Sometimes	☐ Often	
11. Does your child wave to greet people?		☐ Not Yet	☐ Sometimes	☐ Often	
12. Does your child point to objects?		☐ Not Yet	☐ Sometimes	☐ Often	
13. Does your child nod his/her head to indicate *yes*?		☐ Not Yet	☐ Sometimes	☐ Often	
Sounds					
14. Does your child use sounds or words to get attention or help?		☐ Not Yet	☐ Sometimes	☐ Often	
15. Does your child string sounds together, such as *uh oh, mama, gaga, bye bye, bada*?		☐ Not Yet	☐ Sometimes	☐ Often	
16. About how many of the following consonant sounds does your child use: *ma, na, ba, da, ga, wa, la, ya, sa, sha*?	☐ None	☐ 1–2	☐ 3–4	☐ 5–8	☐ over 8
Words					
17. About how many different words does your child use meaningfully that you recognize (such as *baba* for bottle; *gaggie* for doggie)?	☐ None	☐ 1–3	☐ 4–10	☐ 11–30	☐ over 30
18. Does your child put two words together (for example, *more cookie, bye bye Daddy*)?		☐ Not Yet	☐ Sometimes	☐ Often	
Understanding					
19. When you call your child's name, does he/she respond by looking or turning toward you?		☐ Not Yet	☐ Sometimes	☐ Often	
20. About how many different words or phrases does your child understand without gestures? For example, if you say "where's your tummy," "where's Daddy," "give me the ball," or "come here," without showing or pointing, your child will respond appropriately.	☐ None	☐ 1–3	☐ 4–10	☐ 11–30	☐ over 30
Object Use					
21. Does your child show interest in playing with a variety of objects?		☐ Not Yet	☐ Sometimes	☐ Often	
22. About how many of the following objects does your child use appropriately: cup, bottle, bowl, spoon, comb or brush, toothbrush, washcloth, ball, toy vehicle, toy telephone?	☐ None	☐ 1–2	☐ 3–4	☐ 5–8	☐ over 8
23. About how many blocks (or rings) does your child stack? **Stacks**	☐ None	☐ 2 blocks	☐ 3–4 blocks	☐ 5 or more	
24. Does your child pretend to play with toys (for example, feed a stuffed animal, put a doll to sleep, put an animal figure in a vehicle)?		☐ Not Yet	☐ Sometimes	☐ Often	

Do you have any concerns about your child's development? ☐ yes ☐ no **If yes, please describe on back.**

Appendix B. CSBS Scoring Guidelines

Cutoff Scores for the *CSBS-DP*

Infant/Toddler Checklist

Amy M. Wetherby & Barry M. Prizant © 2001 by Paul H. Brookes Publishing

		COMPOSITES			TOTAL
		Communication	**Expressive Speech**	**Symbolic**	
6 months	*No Concern*	8 to 26	2 to 14	3 to 17	13 to 57
	Concern	0 to 7	0 to 1	0 to 2	0 to 12
7 months	*No Concern*	8 to 26	2 to 14	3 to 17	14 to 57
	Concern	0 to 7	0 to 1	0 to 2	0 to 13
8 months	*No Concern*	8 to 26	4 to 14	4 to 17	16 to 57
	Concern	0 to 7	0 to 3	0 to 3	0 to 15
9 months	*No Concern*	9 to 26	4 to 14	4 to 17	18 to 57
	Concern	0 to 8	0 to 3	0 to 3	0 to 17
10 months	*No Concern*	12 to 26	5 to 14	5 to 17	23 to 57
	Concern	0 to 11	0 to 4	0 to 4	0 to 22
11 months	*No Concern*	13 to 26	5 to 14	6 to 17	25 to 57
	Concern	0 to 12	0 to 4	0 to 5	0 to 24
12 months	*No Concern*	14 to 26	6 to 14	7 to 17	28 to 57
	Concern	0 to 13	0 to 5	0 to 6	0 to 27
13 months	*No Concern*	15 to 26	6 to 14	8 to 17	29 to 57
	Concern	0 to 14	0 to 5	0 to 7	0 to 28
14 months	*No Concern*	16 to 26	7 to 14	9 to 17	33 to 57
	Concern	0 to 15	0 to 6	0 to 8	0 to 32
15 months	*No Concern*	18 to 26	7 to 14	10 to 17	35 to 57
	Concern	0 to 17	0 to 6	0 to 9	0 to 34
16 months	*No Concern*	18 to 26	7 to 14	11 to 17	36 to 57
	Concern	0 to 17	0 to 6	0 to 10	0 to 35
17 months	*No Concern*	18 to 26	7 to 14	11 to 17	37 to 57
	Concern	0 to 17	0 to 6	0 to 10	0 to 36
18 months	*No Concern*	18 to 26	8 to 14	11 to 17	38 to 57
	Concern	0 to 17	0 to 7	0 to 10	0 to 37
19 months	*No Concern*	18 to 26	8 to 14	11 to 17	38 to 57
	Concern	0 to 17	0 to 7	0 to 10	0 to 37
20 months	*No Concern*	19 to 26	8 to 14	12 to 17	39 to 57
	Concern	0 to 18	0 to 7	0 to 11	0 to 38
21 months	*No Concern*	19 to 26	9 to 14	12 to 17	40 to 57
	Concern	0 to 18	0 to 8	0 to 11	0 to 39
22 months	*No Concern*	19 to 26	9 to 14	12 to 17	40 to 57
	Concern	0 to 18	0 to 8	0 to 11	0 to 39
23 months	*No Concern*	19 to 26	9 to 14	13 to 17	42 to 57
	Concern	0 to 18	0 to 8	0 to 12	0 to 41
24 months	*No Concern*	19 to 26	9 to 14	13 to 17	42 to 57
	Concern	0 to 18	0 to 8	0 to 12	0 to 41
		Communication	**Expressive Speech**	**Symbolic**	**TOTAL**

Appendix C. Example of the Quantity of Exploration Coding Sheet

		Quadrant 1										Quadrant 2							Quadrant 3								Movement in Quadrants						
	Unscoreable	Item1: ____	Baton	Book	Ball	Nets	White Box	Tube	Yarn	Boat	Clear Box 1	Expanding Sphere	Lizards	Slinky	Stripe Box	Purple Cloth	Truck	Clear Box 2	Wild Cloth	Doll 1: Girl	Doll 2: Boy	Gold Circle	Keys	Koosh Ball	Rhino	Spin Wheel	Q1	Q2	Q3	Q4	Q5	Unscoreable	Quadrant Change
0:00-0:05																																	
0:06-0:10																																	
0:11-0:15																																	
0:16-0:20																																	
0:21-0:25																																	
0:26-0:30																																	
0:31-0:35																																	
0:36-0:40																																	
0:41-0:45																																	
0:46-0:50																																	
0:51-0:55																																	
0:56-1:00																																	
…																																	
9:01-9:05																																	
9:06-9:10																																	
9:11-9:15																																	
9:16-9:20																																	
9:21-9:25																																	
9:26-9:30																																	
9:31-9:35																																	
9:36-9:40																																	
9:41-9:45																																	
9:46-9:50																																	
9:51-9:55																																	
9:56-10:00																																	
Total Count:																																	
Number of Items Explored:																																	
Number of Containers Explored:																																	

Appendix D. Quantity of Exploration Scoring Definitions

Quantity of Exploration Scoring Definitions

List of Items/Toys

- Truck
- Book
- Nets and Ball (2 nets and 1 ball)
- Dolls (2 dolls: male and female)
- Lizards (2 lizards)
- Spinning wheel toy
- Koosh ball ®
- Keys
- Rhino
- Boat
- Purple cloth (inside striped shoe box)
- Strings (3 strings attached)
- Baton (water and glitter inside baton)
- Slinky ®
- Expanding sphere toy (multi colored, inside clear box)
- Print cloth (red, green, and black patterned, inside clear box)
- Gold disk (circle, cardboard)
- Pull tube (expandable, blue, inside white shoe box)

List of Quadrants

- Quadrant 1: northeast quarter of room closest to door
- Quadrant 2: northwest quarter of room closest to window
- Quadrant 3: southeast quarter of room closest to observation window/mirror
- Quadrant 4: southwest quarter of room without toys (view obstructed from camera)
- Quadrant 5: alcove around door (separate from quadrant 1)

Item/Toy Exploration Coding Instructions

- For each interval mark each toy the child explores.
- Mark exploration if the child is playing with the toy, mouthing the toy, or simply holding the toy. Crawling over/moving over a toy does not count if it does not appear that the child intended to touch the object (i.e. crawling over the gold circle without noticing, walking and accidentally kicking a toy).
- Mark an “x” in the interval when the child explores an item.
- If the child uses a toy to explore another toy, mark both only if the child is attending to both objects. For example, the child is hitting another toy with the baton.
- If the child or toy is not visible but there are *very clear* indicators that the child is playing with the toy (i.e. sound of the slinky, keys), mark the interval.

Movement in Space/Quadrant Coding Instructions

- Score the quadrant the child is physically in during each interval.
- Only mark one quadrant per interval, based on which quadrant the child spends the majority of their time.
- Use where body is touching the ground as a marker (i.e. feet touching ground, rather than an arm reaching over a quadrant in the air).
- If the child is resting on 2 or more quadrants, mark the quadrant the majority of the child's body is in, if equal mark the quadrant in which the child is facing toward.
- If the child moves between 2 or more quadrants in an interval, mark the interval as the quadrant they are in for the majority of the interval (i.e. more than 2.5 seconds in one quadrant) if there is no clear majority, mark the quadrant where the child started. In cases where the child is not visible for the entire interval (or at all) but their location is known (usually in quadrants 4 and 5) mark the quadrant.

Appendix E. Example of the Type of Exploration Coding Sheet

Time	Appropriate Exploration/Play		Stereotypic Play		Off-Task Behavior			Unscoreable
	Active and Engaged Exploration/ Play	Motor Exploration	Stereotypic Exploration/ Play	Stereotypic Behavior	Unengaged Exploration/ Play	Off-task	Cry/Tantrum	
0:00-0:05								
0:06-0:10								
0:11-0:15								
0:16-0:20								
0:21-0:25								
0:26-0:30								
0:31-0:35								
0:36-0:40								
0:41-0:45								
0:46-0:50								
0:51-0:55								
0:56-1:00								
…								
9:01-9:05								
9:06-9:10								
9:11-9:15								
9:16-9:20								
9:21-9:25								
9:26-9:30								
9:31-9:35								
9:36-9:40								
9:41-9:45								
9:46-9:50								
9:51-9:55								
9:56-10:00								
Time	Active and Engaged Exploration/ Play	Motor Exploration	Stereotypic Play	Stereotypic Behavior	Unengaged Exploration/ Play	Off-task	Cry/Tantrum	Unscoreable
Total Count								
% of Intervals								
Collapsed Total Count								
Collapsed % of Intervals								

Appendix F. Type of Exploration Scoring Definitions

Type of Exploration Scoring Definitions

Basic Instructions:

- Each video should be 10 minutes long, broken into 120 five-second scoring intervals.
- There are four main coding categories, with subcategories within.
- Select one scoring category that best represents each five-second interval. Score only one category per interval.
- If the child's behavior qualifies for multiple categories score the category of behavior the child is engaged in for the majority of the interval.
- If the video is shorter than 10 minutes, mark unscorable for the relevant intervals.

Appropriate Exploration/Play

Active and Engaged Exploration/Play

Score when the child is manipulating objects in an appropriate manner and attending to the object while doing so. Score for both toys and containers. Includes explorative, functional and symbolic play.

- explorative play includes touching, inspecting, mouthing, climbing on boxes, opening boxes, etc.
- functional play includes using an object in the intended way such as throwing a ball, pushing a truck, etc.
- symbolic or pretend play includes having a doll or toy animal carry out play actions, etc.

Child must be attending to the object while engaging in exploration/play to be coded as active and engaged exploration/play. If the child is not attending, then it is coded as unengaged exploration/play.

Exceptions:

- child does not need to be directly attending to the object when mouthing toys, showing toys to a caregiver (must be directly showing to the parent, not just waving in the air), or when riding vehicle toys (sometimes children try to ride the toy vehicles as if they are child-size), if the child is engaged in the activity of riding (i.e. not just sitting) then active and engaged exploration/play is coded even is the child is not directly attending to the toy.

Notes:

- Reaching for toys nearby is included in active and engaged exploration/play. If the child needs to walk/crawl a few steps to the toy, this will be coded as motor exploration instead.
- If the child sits on the containers, active and engaged exploration/play should only be coded if the child is directly attending to the container, if not inappropriate or unengaged play should be coded.

Motor Exploration

This code is used predominately to capture when a child is moving to a new toy, but this movement takes the majority of the interval. Score as motor exploration if the child moves at least 2 ft (around 2-3 steps) towards a toy, otherwise score as appropriate and engaged exploration/play if the toy is within reach of the child and he immediately picks up the toy and starts manipulating it.

Notes:

- Motor behavior must be related to object. Do not score if the child is approaching the parent, only when they are approaching a toy.
- Movement must look intentional (moving towards a toy—cannot be wandering aimlessly around the room).
- May or may not be holding a toy while engaging in motor exploration.
- If the child moves to a new item and is playing with it for the majority of the interval, then play should be coded instead.

Stereotypic Behavior and Play

Stereotypic Behavior

Any stereotypic behavior that occurs WITHOUT the use of a toy or item

Stereotypic Play

Any stereotypic behavior that occurs WITH the use of a toy or item

Stereotypy is usually characterized by intense concentration, repetition, rigidity and invariance of exact movements, as well as a tendency to be inappropriate in nature. Stereotypic behavior and play can occur with or without a toy. Regardless of the duration of the stereotypic behavior and play, check the interval in which it occurs. This category can be scored in the same interval as another category. If the self-stimulatory behavior is very brief (i.e. a brief instance of hand-flapping) mark stereotypic behavior and play as well as whatever category the child is doing for the majority of the interval. If the child is engaging in stereotypic behavior and play the majority of the interval, only mark the stereotypic behavior and play category. Please describe the behavior in the notes section (this is to gather information about whether the child is involving a toy in the self-stimulatory behavior—i.e. the child knocks the book over repeatedly and stands it back up, vs hand flapping).

Code any of the following classic examples whenever they occur, regardless of if they occur for more than 10 seconds

Classic Examples:

- Holding the truck upside down and rolling the wheels.
- Holding object very close to eye for visual inspection out of the corner of eye--flipping pages of book close to face, holding gold circle close to face
- Hand flapping
- Brushing item against body
- Smelling items

Children may also exhibit stereotypic play outside of these classic examples through

persistent fixation on parts of objects or an inflexible adherence to specific, nonfunctional routines or rituals. If the child is *repeating* the *exact behavior* for more than 10 seconds, it is considered stereotypic play. There must be no variation in the routine for this to happen. The child's attention is usually very focused. The repetitive behavior can be a functional way to use the toy, or a more unique way to use the toy, but the child plays with the toy in a very specific fashion regardless. If the child using the toy in an *appropriate* way (i.e. throwing ball or pushing truck, as opposed to playing with a toy in an unusual way such as rotating box lid 90 degrees repeatedly, lining up animals, etc.) for a long period of time, stereotypic behavior and play is coded conservatively, and only when the routine is *exactly* the same each time. Once a particular stereotypic behavior has been established for a child (i.e. you viewed a particular stereotypic behavior in the video) then the child does not have to engage in it for over 10 seconds to be marked every time. Instead the stereotypic play is immediately marked when it occurs repeatedly.
Examples:
Repetitive Singular Behaviors:

- Exact slinky movement repetition
- Rubbing tape
- Pulling strings on koosh ball
- Fanning book open

Repetitive Routines:

- Touching ball to head then throwing it (repeatedly)
- Lining up blocks in identical rows repetitively

Off-Task Behavior

Unengaged Exploration/Play

Score when play behaviors are inappropriate to object being explored or when the child is not attending to the object they are playing with.
This includes:

- Playing with an object, but not attending to the object while doing so—attention is directed elsewhere
- May simply be holding a toy or may be manipulating a toy, but not attending to the object while doing so—attention is directed elsewhere
- Looking at a toy but not touching it. Must be looking at one toy for the majority of the time. The toy must be within reach of the child. The child "surveying" the room (looking around at a lot of toys) does not count here.

Notes:

- Child may be moving about the room while doing this. If the child is moving specifically towards another object and carrying a toy with him, this should be scored as motor exploration. If the child is aimlessly moving about the room or walking towards a parent while holding a toy this is coded as inappropriate or unengaged exploration/play.

Off-Task

The child is engaged in an activity unrelated to any toy or object. Not moving directly

towards an object (this would be motor exploration). If a child is holding a toy, off-task cannot be scored.
This includes:

- Sitting down and not intently looking at any toys to play with
- Lying down, rolling around, etc.
- Staring at the wall or ceiling
- Playing with blinds, lights, cabinets, electric sockets, etc.
- Trying to open door
- Playing with parent's magazine, purse, etc.
- Pulling tape up off the floor
- Walking around but not exploring object
- If the child is interacting with mom/dad and is not engaging in other play behaviors please mark off task, but make a note of interaction in notes section

Crying/Tantrum
If the child is throwing a tantrum this is the only score they are given for any behavior that occurs during the tantrum. So in other words, if the child is crying while playing with another toy, then the exploration would not be scored, just the crying and throwing a tantrum.
This Includes:

- Crying
- Yelling
- Whining
- Hitting
- Swiping in protest or anger

Unscorable
If the child is out of the view of camera or partially out of frame and you cannot be certain what the subject is doing for the majority of the interval mark unscorable for the interval.

Hierarchy of Scoring
If a child engages in two or more behaviors during the interval and it is difficult to decide which behavior is happening for the majority of the interval (i.e. both behaviors occur for 2.5 seconds) code the highest level behavior that occurs. Use the hierarchy below to determine the highest level behavior:

1) Active and Engaged Play/Exploration
2) Motor Exploration
3) Stereotypic Behavior and Play
4) Unengaged Exploration/Play or Off-Task or Crying/Tantrum
5) Unscorable

*Off-Task and Crying/Tantrum are considered the same level of behavior, with Crying/Tantrum as a subcategory of Off-Task. If you need to decide between whether Off-Task or Crying/Tantrum occurred for the majority of the interval, always choose Crying/Tantrum.

Appendix G. Example of the PCI Coding Sheet

Time	Unscoreable	Pointing	Reaching	Showing	Giving	Approach	Social Vocalizations	Orientation	Social Referencing	Child Affect
0:00-0:05										
0:06-0:10										
0:11-0:15										
0:16-0:20										
0:21-0:25										
0:26-0:30										
0:31-0:35										
0:36-0:40										
0:41-0:45										
0:46-0:50										
0:51-0:55										
0:56-1:00										
…										
9:01-9:05										
9:06-9:10										
9:11-9:15										
9:16-9:20										
9:21-9:25										
9:26-9:30										
9:31-9:35										
9:36-9:40										
9:41-9:45										
9:46-9:50										
9:51-9:55										
9:56-10:00										
Time	Unscoreable	Pointing	Reaching	Showing	Giving	Approach	Social Vocalizations	Orientation	Social Referencing	Child Affect
Totals										

Appendix H. PCI Scoring Definitions

PCI Scoring Definitions

The following child behaviors will be coded through five-second interval coding. Mark an "X" in each interval the behavior occurs. Only mark one "X" per interval, even if the behavior occurs multiple times. If the same behavior occurs over multiple intervals, only the interval is begins in should be marked. Behavior should be coded when it begins, not when it is completed (i.e. at the start of a phrase, when a child begins moving for an approach, etc.).

Unscoreable

If the child ever moves out of the frame of the camera, mark unscoreable for that interval and do not code any other behaviors. Child must be out of frame for the majority of the interval, otherwise it should still be coded. If 3 or more of the intervals in a block used for the child affect rating are unscoreable, affect should not be scored. Otherwise affect should be scored based on the scoreable intervals.

Gestures

Pointing

The child uses an extended index finger to indicate his/her desire for an object or event or to show an object to caregiver. If reach turns into a point, code as a point. Pointing may or may not be coordinated with gaze toward the caregiver.
Examples:
Pointing at a book after the caregiver
Pointing at a toy across the room
Pointing to the locked cabinet

Reaching

The child reaches for an object the caregiver has, or an object out of reach. The child cannot simply reach and grasp an object within reach. It must be out of reach in some way.
Examples:
Caregiver is holding a ball, and the child reaches to grasp it.
Caregiver has a toy in possession and showing or enticing the child with the toy and the child reaches for the toy.
Child reaches for object up high.
Reaching for caregiver's hands during Peek-a-boo

Showing

The child raises an object upward toward the caregiver's face. The object can be held still or shaken or waved in front of the caregiver. Child may retract the object, or end the show in a give. If the child ends the show in a give, please code a show AND a give.
Examples:
Holding up an object toward the face of an caregiver
Caregiver says "show it to me!" and the child then shows the object

Giving

The child brings an object to the caregiver. This also includes pushing or throwing items to the caregiver (i.e. when playing catch or pushing train back and forth). May leave the object with caregiver and retract hands, or may continue to touch object. It the child continues to touch the object, the child must place the object so it is touching the caregiver in some way (in hands, on lap, etc.). Child must initiate handing over item. If caregiver takes an item away from the child giving should not be coded.
Examples:
Bringing an item to caregiver
Throwing the ball to the caregiver
Pushing Thomas the Train to the caregiver
Child hands caregiver a toy and watches them play with it
Gives object to caregiver after caregiver says “give it to me” or “my turn”

Approach

Any time the child directly comes up to the caregiver. Must have some movement: one or more steps toward caregiver. If child come towards caregiver, but then stops at a toy, approach cannot be coded. May or may not be looking at caregiver. Must be facing caregiver during approach.
Examples:
Walks up to caregiver
Sits/positions near caregiver while facing caregiver
Child responds to caregiver’s request to come play an activity (e.g., child comes to train set when caregiver says “come play trains!”)

Social Vocalizations

Any vocalization by the child directed toward the caregiver. Vocalizations include, babbling, sound effects, word approximations, full words, and phrases. Do not code non-language behaviors such as laughing, crying, gasping, etc. All instances must be directed toward the caregiver. If the caregiver is out of the frame, social vocalizations cannot be coded. If the child is more than arms reach away, do not code language as social vocalizations, unless clearly directed toward caregiver via child’s orientation toward the caregiver, the child is clearly asking a question, or clearly responding to the caregiver. If there is any doubt as to the interval in which the child vocalizes, mark it in the earlier interval. If the sentence is very long and could be marked in two intervals, only mark it in the initial interval (marking the number of initiations not the length).
Examples:
Babababa
Wawa (while holding baby)
Baby’s babas
Cwash!
Wooo!
Yay

No
Do you want it?
Come here
You do it
Baby is tired
I'm going to take baby to the doctor

Orientation

Orientation is coded through the child's use of eye contact with the caregiver and the child's orientation of his/her head towards the caregiver. The child only needs to be oriented in one of these ways in order for orientation to be scored (only body facing caregiver, only head facing caregiver, or looking at caregiver). The child must be facing the caregiver's face or torso, if the child is only facing the caregiver's hands, foot, leg, etc. this does not count. The child must be oriented for the **MAJORITY** (3 seconds or more) of the interval to be coded. If the child only briefly orients (less than 3 seconds) do not code. If the caregiver is not in the frame, do not score.

Social Referencing

Code each time the child references the caregivers face. Must look toward face, not hands, body, or toy that the caregiver has. Referencing will often be double coded with Orientation. Caregiver does not have to be looking at child, but child must be looking at caregiver's face. Code the duration of the reference. So, if the child references over multiple intervals, the behavior should be marked in all the intervals. If the caregiver's face is out of the frame, do not score.
Examples:
Child looks at caregiver's face during play.
Child looks at caregiver when caregiver calls child's name.
Caregiver asks a question and child looks at caregiver's face and responds.

Child Affect

The amount of positive affect displayed by the child through positive facial expression, vocalization, tone of voice, gestures, and behavior, versus the amount of negative affect displayed through negative facial expression, vocalizations, tone of voice, gestures, and behavior. Ratings of child affect will be made every 30 seconds. Ratings of child affect will be coded using a likert scale defining negative to positive displays of affect.

Negative Affect:
High intensity examples: tantrumming, screaming, crying, aggressive behavior (hitting, kitting, throwing objects),
Low intensity examples: negative facial expressions (frowning), whining, negative statements in regards to caregiver, toys or assessment (stop it, I want to go, bye-bye, all done, no)
Positive Affect:
High intensity examples: excitatory responses (laughing, jumping, singing, dancing, any sort of excitatory vocalization that is clearly not words)
Low intensity examples: positive facial expressions (smiling), positive statements in

regards to caregiver, toys or assessment (yay, wow, I like it, more)

*note: if low and high intensity behaviors co-occur, code for the higher intensity behavior.
Example: child says negative statements while tantrumming, code as high intensity negative affect.

*note: if positive and negative behaviors both occur during one interval, determine the score for all the positive behaviors, then determine the score for all the negative behaviors, and average the 2 scores.

1 High negative affect. High intensity negative affect for the majority of the interval (4 or more intervals).

2 Moderate negative affect. 2-3 short instances of high intensity negative affect and/or multiple instances (more than 3) of low intensity negative affect throughout the interval.

3 Minimal negative affect. 2-3 short instances of low intensity negative affect and/or 0-1 instances of high intensity negative affect.

4 **Neutral affect.** Absence of any positive or negative affect throughout the interval. 1 brief instance of low intensity negative or positive affect may occur and still be coded as neutral.

5 Minimal positive affect. 2-3 short instances of low intensity positive affect and/or 0-1 instances of high intensity positive affect.

6 Moderate positive affect. 2-3 short instances of high intensity positive affect and/or multiple instances (more than 3) of low intensity positive affect throughout the interval.

7 High positive affect. High intensity positive affect for the majority of the interval. (4 or more intervals).

Affect Notes

Use this section to make brief notes about any affect you see. You can then use these notes to help you make your rating of affect at the end of each 30-second block.

REFERENCES

Akshoomoff, N. (2006). Use of the mullen scales of early learning for the assessment of young children with autism spectrum disorders. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence, 12*(4-5), 269-277. doi:Q778PT217V25017U

Baranek, G. T., Barnett, C. R., Adams, E. M., Wolcott, N. A., Watson, L. R., & Crais, E. R. (2005). Object play in infants with autism: Methodological issues in retrospective video analysis. *The American Journal of Occupational Therapy : Official Publication of the American Occupational Therapy Association, 59*(1), 20-30.

Bishop, S. L., Guthrie, W., Coffing, M., & Lord, C. (2011). Convergent validity of the mullen scales of early learning and the differential ability scales in children with autism spectrum disorders. *American Journal on Intellectual and Developmental Disabilities, 116*(5), 331-343. doi:10.1352/1944-7558-116.5.331

Boyd, B. A., Odom, S. A., Humphreys, B. P., & Sam, A. M. (2010). Infants and toddlers with autism spectrum disorder: early identification and early intervention. *Journal of Early Intervention, 32,* 75-98. doi: 10.1177/1053815110362690

Bryson, S. E., Zwaigenbaum, L., Brian, J., Roberts, W., Szatmari, P., Rombough, V., & McDermott, C. (2007). A prospective case series of high-risk infants who developed autism. *Journal of Autism and Developmental Disorders, 37*(1), 12-24. doi:10.1007/s10803-006-0328-2

Charman, T., Taylor, E., Drew, A., Cockerill, H., Brown, J., & Baird, G. (2005). Outcome at 7 years of children diagnosed with autism at age 2: Predictive validity of assessments conducted at 2 and 3 years of age and pattern of symptom change over time. *Journal of Child Psychology and Psychiatry, 46*(5), 500-513. doi:10.1111/j.1469- 7610.2004.00377.x

Chawarska, K., Klin, A., Paul, R., Macari, S., & Volkmar, F. (2009). A prospective study of toddlers with ASD: Short-term diagnostic and cognitive outcomes. *Journal of Child Psychology and Psychiatry, 50*(10), 1235-1245. doi:10.1111/j.1469-7610.2009.02101.x

Chawarska, K., Shic, F., Macari, S., Campbell, D. J., Brian, J., Landa, R., . . . Bryson, S. (2014). 18-month predictors of later outcomes in young siblings of children with autism spectrum disorder: A baby siblings research consortium study. *Journal of American Academy of Child and Adolescent Psychiatry,* doi: 10.1016/j.jaac.2014.09.015

Christensen, L., Hutman, T., Rozga, A., Young, G. S., Ozonoff, S., Rogers, S. J., . . . Sigman, M. (2010). Play and developmental outcomes in infant siblings of children with autism. *Journal of Autism and Developmental Disorders, 40*(8), 946-957. doi:10.1007/s10803-010-0941-y

Cornew, L., Dobkins, K. R., Akshoomoff, N., McCleery, J. P., & Carver, L. J. (2012). Atypical social referencing in infant siblings of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(12), 2611-2621. doi:10.1007/s10803-012-1518-8

Corsello, C. M. (2005). Early intervention in autism. *Infants & Young Children, 18*(2), 74-85. doi:10.1097/00001163-200504000-00002

Curran, P. J., Obeidat, K., Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development, 11*(2), 122-136. doi:10.1080/15248371003699969

Daniels, A. M., Rosenberg, R. E., Law, J. K., Lord, C., Kaufmann, W. E., & Law, P. A. (2011). Stability of initial autism spectrum disorder diagnoses in community settings. *Journal of Autism and Developmental Disorders, 41*(1), 110-121. doi:10.1007/s10803-010-1031-x

Dawson, G. (2008). Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. *Development and Psychopathology, 20*(3), 775-803. doi: 10.1017.S0954579408000370

Dawson, G., Osterling, J., Meltzoff, A. N., & Kuhl, P. (2000). Case study of the development of an infant with autism from birth to two years of age. *Journal of Applied Developmental Psychology, 21*(3), 299-313. doi: 10.1016/S0193-3973(99)00042-8

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikseseth, 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. doi: 10.1080/15374410902851739

Fountain, C., Winter, A. S., & Bearman, P. S. (2012). Six developmental trajectories characterize children with autism. *Pediatrics, 129*(5), e1112-e1120. doi: 10.1542/peds.2011-1601

Georgiades, S., Szatmari, P., Zwaigenbaum, L., Bryson, S., Brian, J., Roberts, W.. . . . Garon, N.(2013). A prospective study of autistic-like traits in unaffected siblings of probands with autism spectrum disorder, *JAMA Psychiatry, 70*(1), 42-48. doi: 10.1001/2013.jamapsychiatry.1

Guthrie, W., Swineford, L. B., Nottke, C., & Wetherby, A. M. (2013). Early diagnosis of autism spectrum disorder: Stability and change in clinical diagnosis and symptom presentation. *Journal of Child Psychology and Psychiatry, and Allied Disciplines,* *54*(5), 582-590. doi:10.1111/jcpp.12008

Ibanez, L. V., Grantz, C. J., & Messinger, D. S. (2013). The development of referential communication and autism symptomatology in high-risk infants. *Infancy: The Official Journal of the International Society on Infant Studies, 18*(5), 10.1111/j.1532-7078.2012.00142.x. doi:10.1111/j.1532-7078.2012.00142.x

Itzchak, E. B. & Zachor, D. A. (2011). Who benefits form early intervention in autism spectrum disorders? *Research in Autism Spectrum Disorders, 5*(1), 345-350. Doi: 10.1016/j.rasd.2010.04.018

Jonsdottir, S. L., Saemundsen, E., Asmundsdottir, G., Hjartardottir, S., Asgeirsdottir, B. B., Smaradottir, H. H., . . . Smari, J. (2007). Follow-up of children diagnosed with pervasive developmental disorders: Stability and change during preschool years. *Journal of Autism and Developmental Disorders, 37*(7), 1361-1374. doi: 10.1007/s10803-006-0181-z

Kleinman, J. M., Ventola, P. E., Pandey, J., Verbalis, A. D., Barton, M., Hodgson, S., . . . Fein, D. (2008). Diagnostic stability in very young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*(4), 606-615. doi:10.1007/s10803- 007-0427-8

Landa, R., & Garrett-Mayer, E. (2006). Development in infants with autism spectrum disorders: A prospective study. *Journal of Child Psychology and Psychiatry, 47*(6), 629-638. doi: 10.1111/j.1469-7610.2006.01531.x

Landa, R. J., Gross, A. L., Stuart, E. A., & Bauman, M. (2012). Latent class analysis of early developmental trajectory in baby siblings of children with autism. *Journal of Child Psychology and Psychiatry, 53*(9), 986-996. doi:10.1111/j.1469-7610.2012.02558.x

Lord, C., Luyster, R. J., Gotham, K., & Guthrie, W. (2012). *Autism diagnostic observation schedule, Second edition, Toddler module.* Torrance, CA: 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 Psychiatry, 63*(6), 694-701. doi:10.1001/archpsyc.63.6.694

Lord, C., Rutter, M., DiLavore, P.C., & Risi, S. (2002). *Autism diagnostic observation schedule.* Los Angeles, CA: Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism diagnostic observation schedule, Second edition, Modules 1-4.* Los Angeles, CA: Western Psychological Services.

Macari, S. L., Campbell, D., Gengoux, G. W., Saulnier, C. A., Klin, A. J., & Chawarska, K. (2012). Predicting developmental status from 12 to 24 months in infants at risk for autism spectrum disorder: A preliminary report. *Journal of Autism and Developmental Disorders, 42*(12), 2636-2647. doi:10.1007/s10803-012-1521-0

Mazefsky, C. A., & Oswald, D. P. (2006). The discriminative ability and diagnostic utility of the ADOS-G, ADI-R, and GARS for children in a clinical setting. *Autism, 10*(6), 533-549. doi: 10.1177/1362361306068505

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

Nadig, A. S., Ozonoff, S., Young, G. S., Rozga, A., Sigman, M., & Rogers, S. J. (2007). A prospective study of response to name in infants at risk for autism. *Archives of Pediatrics & Adolescent Medicine,161*(4), 378-383. doi: 101001/archpedi.161.4.378

Ozonoff, S., losif, A., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., . . . Young, G. S. (2010). A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child & Adolescent Psychiatry, 49*(3), 256-266. doi: 10.1097/00004583-201003000-00009

Ozonoff, S., Macari, S., Young, G. S., Goldring, S., Thompson, M., & Rogers, S. J. (2008). Atypical object exploration at 12 months of age is associated with autism in a prospective sample. *Autism : The International Journal of Research and Practice, 12*(5), 457-472. doi:10.1177/1362361308096402

Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., . . . Stone, W. L. (2011). Recurrence risk for autism spectrum disorders: A baby siblings research consortium study. *Pediatrics, 128*(3), e488-495. doi: 10.1543/peds.2010-2825

Paul, H. (2000). Predicting outcomes of early expressive language delay: Ethical implication. In D. V. M. Bishop & L. B. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (pp. 195-209). Hove, U.K.: Psychology Press.

Paul, R., Chawarska, K., Cicchetti, D., & Volkmar, F. (2008). Language outcomes of toddlers with autism spectrum disorders: a two year follow-up. *Autism Research, 1*(2), 97-107.

Pierce, K., Carter, C., Weinfeld, M., Desmond, J., Hazin, R., Bjork, R., & Gallagher, N. (2011). Detecting, studying, and treating autism early; the one-year well-baby check-up approach. *The Journal of Pediatrics, 159*(3), 458-465

Pierce, K., Conant, D., Hazin, R., Stoner, R., & Desmond, J. (2011). Preference for geometric patterns early in life as a risk factor for autism. *Archives of General Psychiatry, 68*(1), 101- 109.

Pierce, K. & Courchesne, E. (2001). Evidence for a cerebellar role in reduced exploration and stereotyped behavior in autism. *Behavioral Psychiatry*, *49*: 655-664.

Redcay, E., & Courchesne, E. (2005). When is the brain enlarged in autism? A meta-analysis of all brain size reports. *Biological Psychiatry, 58*(1), 1-9. doi:S0006-3223(05)00369-0

Rozga, A., Hutman, T., Young, G. S., Rogers, S. J., Ozonoff, S., Dapretto, M., & Sigman, M. (2011). Behavioral profiles of affected and unaffected siblings of children with autism: Contribution of measures of mother-infant interaction and nonverbal communication. *Journal of Autism and Developmental Disorders, 41*(3), 287-301. doi:10.1007/s10803-010-1051-6

Rutter, M., LeCouteur, A., & Lord, C. (2003). *Autism diagnostic interview - revised.* Los Angeles, CA: Western Psychological Services.

Scambler, D. J., Hepburn, S. L., & Rogers, S. J. (2006). A two-year follow-up on risk status identified by the checklist for autism in toddlers. *Journal of Developmental and Behavioral Pediatrics, 27*(Suppl2; Supplement), S104-S110. doi: 10.1097/00004703-200604002-00008

Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., . . . Wigler, M. (2007). Strong association of de novo copy number mutations with autism. *Science, 316,* 445-449. doi: 10.1126/science.1138659

Shumway, S., Thurm, A., Swedo, S. E., Deprey, L., Barnett, L. A., Amaral, D. G., . . . Ozonoff, S. (2011). Brief report: Symptom onset patterns and functional outcomes in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 41*(12), 1727-1732. doi: 10.1007/s10803-011-1203-3

Sparrow, S. S., Balla, A. D., & Cicchetti, V. D. (1984). *Vineland adaptive behavior scales.* Circle Pines, MN: American Guidance Service.

Sparrow, S. S., Cicchetti, V. D., & Balla, A. D. (2005). *Vineland adaptive behavior scales, Second Edition.* Circle Pines, MN: Pearson.

Stevens, M. C., Fein, D., Dunn, M., Allen, D., Waterhouse, L. H., Feinstein, C., & Rapin, I. (2000). Subgroups of children with autism by cluster analysis: A longitudinal examination. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*(3), 346-352. doi: 10.1096/00004583-200003000-00017

Tager-Flusberg, H. (2010). The origins of social impairments in autism spectrum disorder: Studies of infants at risk. *Neural Networks : The Official Journal of the International Neural Network Society*, *23*(8-9), 1072-1076. doi:10.1016/j.neunet.2010.07.008

Talbott, M. R., Nelson, C. A., & Tager-Flusberg, H. (2013). Maternal gesture use and language development in infant siblings of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders,* doi:10.1007/s10803-013-1820-0

Tobii T120 Eye Tracker [Apparatus and software]. (2013). Dandeyrd, Sweden: Tobii Technology.

Tobii Studio 1.3 [Apparatus and software]. (2013). Dandeyrd, Sweden; Tobii Technology.

Turner, L. M., & Stone, W. L. (2007). Variability in outcome for children with an ASD diagnosis at age 2. *Journal of Child Psychology and Psychiatry*, *48*(8), 793-802. doi:10.1111/j.1469-7610.2007.01744.x

Turner, L. M., Stone, W. L., Pozdol, S. L., & Coonrod, E. E. (2006). Follow-up of children with autism spectrum disorders from age 2 to age 9. *Autism*, *10*(3), 243-265. doi: 10.1177/1362361306063296

van Daalen, E., Kemner, C., Dietz, C., Swinkels, S. H. N., Buitelaar, J. K., & van Engeland, H. (2009). Inter-rater reliability and stability of diagnoses of autism spectrum disorder in children identified through screening at a very young age. *European Child & Adolescent Psychiatry, 18*(11), 663-674. doi:10.1007/s00787-009-0025-8

Ventola, P., Saulnier, C. A., Steinberg, E., Chawarska, K., & Klin, A. (2014). Early-emerging social adaptive skills in toddlers with autism spectrum disorders: An item analysis. *Journal of Autism and Developmental Disorders, 44*(2), 283-293. doi:10.1007/s10803-011-1278-x

Wan, M. W., Green, J., Elsabbagh, M., Johnson, M., Charman, T., Plummer, F., & BASIS Team. (2012). Parent-infant interaction in infant siblings at risk of autism. *Research in Developmental Disabilities, 33*(3), 924-932. doi:10.1016/j.ridd.2011.12.011

Wan, M. W., Green, J., Elsabbagh, M., Johnson, M., Charman, T., Plummer, F., & BASIS Team. (2013). Quality of interaction between at-risk infants and caregiver at 12-15 months is associated with 3-year autism outcome. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 54*(7), 763-771. doi:10.1111/jcpp.12032

Wechsler, D. (2012). *Wechsler preschool and primary scale of intelligence, Fourth edition.* Circle Pines, MN: Pearson.

Weiss, M. J. (1999). Differential rates of skill acquisition and outcomes of early intensive behavioral intervention for autism. *Behavioral Interventions, 14*(1), 3-22. doi: 10.1002/(SICI)1099-078X(199901/03)14:1<3::AID-BIN25>3.0.CO;2-F

Wenig, M. (Creator), & Landon, T. (Director). (2004). Yoga kids 3 [DVD]. USA: Gaiam. Available from http://gaiam.com and http://yogakids.com

Wetherby, A., & Prizant, B. (2002). *Communication and symbolic behavior scales developmental profile infant-toddler checklist*. Baltimore, MD: Paul H. Brookes.

Woolfenden, S., Sarkozy, V., Ridley, G., & Williams, K. (2012). A systematic review of the diagnostic stability of autism spectrum disorder. *Research in Autism Spectrum Disorders, 6*(1), 345-354. doi: 10.1016/j.rasd.2011.06.008

Young, G. S., Rogers, S. J., Hutman, T., Rozga, A., Sigman, M., & Ozonoff, S. (2011). Imitation from 12 to 24 months in autism and typical development: A longitudinal rasch analysis. *Developmental Psychology, 47*(6), 1565-1578. doi:10.1037/a0025418

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience.Special Issue: Autism: Modeling Human Brain Abnormalities in Developing Animal Systems, 23*(2-3), 143-152. doi: 10.1016/j.ijdevneu.2004.05.001

Zwaigenbaum, L., Thurm, A., Stone, W., Baranek, G., Bryson, S., Iverson, J., . . . Sigman, M. (2007). Studying the emergence of autism spectrum disorders in high-risk infants: Methodological and practical issues. *Journal of Autism and Developmental Disorders, 37*(3), 466-480. doi:10.1007/s10803-006-0179-x