6-30-2016

# Autism Spectrum Disorders and Low Mental Age: Diagnostic Stability and Developmental Outcomes in Infants and Toddlers

Alexander J. Hinnebusch
*University of Connecticut - Storrs*, alexander.hinnebusch@uconn.edu

Autism Spectrum Disorders and Low Mental Age: Diagnostic Stability and Developmental Outcomes in Infants and Toddlers

Alexander Joseph Hinnebusch, Ph.D.

University of Connecticut, 2016

## ABSTRACT

Autism Spectrum Disorders (ASDs) involve deficits in social and communication abilities and the presence of repetitive and restricted behaviors, and affect approximately 1 in 68 children. Early identification and intervention significantly improve the prognosis of children with ASDs. Research has demonstrated that screening and diagnosis of ASDs is reliable in children as young as 18 months. Cognitive impairment affects up to 70% of individuals with ASDs, and there is diagnostic overlap between ASDs and Intellectual Disability (ID). A subset of children with ASDs present with low mental ages (Low MA), defined as an age equivalent of below 12 months in all developmental domains, at the time of their first evaluation. While research has demonstrated that ASDs can be distinguished from ID diagnostically, the validity of utilizing ASD-specific screening and diagnostic tools in the presence of Low MA has not been investigated. In this study we investigated the diagnostic stability, developmental outcomes, and ASD symptomatology in a sample of children that were initially diagnosed with ASD Low-MA ($n$ = 25), Autistic Disorder (AD; $n$ = 111), and PDD-NOS ($n$ = 82). The majority of children with ASD Low-MA (96%) remained on the autism spectrum at follow-up, compared to AD (86.5%) and PDD-NOS (73.2%). The ASD Low-MA group had the lowest mean ratio IQ scores and made the least developmental progress in all subdomains of the Mullen Early

Scales of Development; the PDD-NOS had higher scores than the AD group in all domains. This trend was observed for the Vineland Adaptive Behavior Scales as well, with exceptions in the Domestic and Community domains; each group had comparable ratio IQ's initially, but experienced a decrease in ratio IQ across time. The ASD Low-MA group had the highest total scores on the CARS, ADOS, and DSM-IV at both time points, indicating high and stable ASD symptom severity. The AD group demonstrated a higher rate of symptom improvement than the PDD-NOS and ASD Low-MA groups. These results suggest the validity of diagnosing ASDs in children that present with Low MA, as well as the high stability of diagnosis and symptoms across all ASDs.

# Autism Spectrum Disorders and Low Mental Age: Diagnostic Stability and Developmental Outcomes in Infants and Toddlers

Alexander Joseph Hinnebusch

B.S., The College of William and Mary, 2010

M.A., The University of Connecticut, 2014

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2016

APPROVAL PAGE

Doctor of Philosophy Dissertation

Autism Spectrum Disorders and Low Mental Age: Diagnostic Stability and Developmental Outcomes in Infants and Toddlers

Presented by

Alexander Joseph Hinnebusch, B.A., M.A.

Major Advisor_____

Deborah Fein, Ph.D.

Associate Advisor_____

Chi-Ming Chen, Ph.D.

Associate Advisor_____

Kimberly Cuevas, Ph.D.

University of Connecticut

2016

**ACKNOWLEDGEMENTS**

Committee

Deborah Fein, Ph.D., Chi-Ming Chen, Ph.D., Kimberly Cuevas, Ph.D.

Readers

Dr. Marianne Barton, Ph.D., Dr. Letitia Naigles, Ph.D.

Early Detection Team

Deborah Fein, Ph.D., Thyde Dumont-Mathieu, MD, Sarah Hodgson, Ph.D., Laura Brennan, Lauren Haisley, Dasal Jashar, Kate Bradbury, Emily Moulten, Lauren Miller, Cara Cordeux, Julia Chen, Christine Sutter, Courtney Manning

Undergraduate Research Assistants

Participating Families, Pediatricians, and Early Intervention Providers

**TABLE OF CONTENTS**

**Title**: Autism Spectrum Disorders and Low Mental Age: Diagnostic Stability and Developmental Outcomes in Infants and Toddlers

## Introduction

### *Autism Spectrum Disorders: Description, Diagnosis, and Intervention*

Autism Spectrum Disorders (ASDs) are characterized by behavioral deficits in social relatedness and communication, along with the presence of restricted interests and/or repetitive behaviors (RRBs). The autism spectrum includes several diagnoses that share these symptom clusters, such as Autistic Disorder (AD) and Pervasive Developmental Disorder – Not Otherwise Specified (PDD-NOS) (American Psychiatric Association, 2000). ASDs have received heightened attention among researchers and healthcare professionals primarily due to increasing rates of prevalence. In 2012, the Centers for Disease Control and Prevention estimated a rate of 1 in 88 children have an ASD (CDC, 2012), and by March 2014 this rate had increased to 1 in 68 (CDC, 2014). Caring for children with ASDs enacts a tremendous societal cost, with an estimated $9 billion spent in the United States on healthcare and therapeutic, educational, and familial services in 2011 (Lavelle, Weinstein, Newhouse, Munir, Kuhlthau, & Prosser, 2014).

Research has demonstrated that early identification and intervention are key in improving the prognosis of children with ASDs, and that intervention significantly lowers long-term societal costs (Chasson, Harris, & Neely, 2007). Instruments like Baron-Cohen's CHecklist for Autism in Toddlers (CHAT) (Baron-Cohen et al., 2000), Stone's Screening Tool for Autism in Two-Year-Olds (STAT) (Stone, McMahon, & Henderson, 2008), and the Modified Checklist for Autism in Toddlers (M-CHAT) (Robins et al., 2001) screen for autism in children as young as 16 months. The use of these screening

tools allow for subsequent diagnosis and appropriate implementation of intervention services. Literature suggests that 25-40 hours per week is an optimal amount of intervention for children with an autism spectrum disorder (ASD) (Howard, Sparkman, Cohen, Green, & Stanislaw, 2005), and foundational studies assert the effectiveness of behaviorally-based interventions in increasing positive outcomes in young children with autism (Dawson et al., 2010; Green, Brennan, & Fein, 2002; McGee, Morrier, & Daly, 1999).

*Autism Spectrum Disorders: Diagnostic Stability*

Research has demonstrated that the ASD diagnosis is reliable in children as young as 18 to 24 months (Charman et al., 2005; Moore & Goodson, 2003; Lord 1995). Filipek et al., in their 2000 review of screening and diagnostic practices for ASDs, cite several studies that found that difficulties with establishing eye contact, responding to name, establishing joint attention, engaging in pretend play and imitation, and delays in nonverbal communication and language are measurable by 18 months of age. This review also references retrospective studies that analyzed the home video tapes of children who had been diagnosed with a pervasive developmental disorder and/or ASD. Observers were able to distinguish children, some as young as eight months old, with an ASD from children with other developmental disabilities due to observed deficits in social and communicative skills, particularly joint attention (Mars, Mauk, & Dowrick, 1998) and a failure to orient toward name when it was called (Werner, Dawson, Osterling, & Dinno, 2000).

In addition to these observable, behavioral differences, several studies have established the stability of ASDs across time. Lord (1995) re-evaluated 16 children at age

three that had been diagnosed with an ASD at age two, finding that 14 of them continued to meet diagnostic criteria. Baron-Cohen et al. (1996) found that ten children diagnosed at 18-20 months all met criteria for ASD diagnoses at 40 months. Stone et al. (1999) followed a larger cohort of 65 children, 37 of whom were diagnosed with an ASD at around 30 months of age; a majority of these children continued to meet criteria for an ASD diagnosis in later childhood. Eaves and Ho (2004) re-evaluated 49 children at age four who had received a diagnosis of an ASD at age two, finding that 88% continued to meet criteria for an ASD. Charman et al. (2005) found that in a cohort of 26 children, ASD diagnosis based on standardized, autism-specific assessments at age three was highly consistent with diagnosis at age seven.

More recently, Sutera et al. (2007) found that only 13 of 73 (18%) children originally diagnosed with an ASD at an initial evaluation no longer met the criteria for an ASD diagnosis at follow-up evaluation, while 60 the children (82%) retained an ASD diagnosis at re-evaluation. Kleinman et al. (2008) found that 15 of 77 children (19%) initially evaluated and diagnosed with ASD no longer met diagnostic criteria for an ASD at a follow-up evaluation. Using partly the same sample as the Kleinman et al. (2008) study, in a poster presentation Hinnebusch et al. (2012) assessed a larger sample of 132 children diagnosed with an ASD at the initial evaluation, and found that upon reevaluation, 28 children (21%) no longer met diagnostic criteria for an ASD.

*Autism Spectrum Disorders: Intelligence and Low Mental Age*

Estimated rates of cognitive impairment in individuals with ASDs across the lifespan ranges from 24% (Chakrabarti & Fombonne, 2001), to as high as 70% (Fombonne, 2003). In Fombonne's (2003) review of 32 epidemiological studies, 30% of

individuals with an ASD presented with mild to moderate cognitive impairment, while 40% were severe to profoundly impaired. A recent prevalence study conducted by Charman et al. (2010) involving 75 children aged 10 to 14 with an ASD, demonstrated that 55% had an IQ below 70, 16% had an IQ below 50, 28% had an average intelligence (IQ between 85 and 115), and 3% had an IQ above 115. Within this sample, adaptive skill levels were significantly lower than IQ. However, intellectual ability varies heavily among individuals with ASDs, as well as among the different diagnoses that are included on the autism spectrum. Additionally, in recent decades the range of IQ in individuals with ASDs has expanded in tandem with increasing ASD prevalence, with this increase in prevalence attributed to improvements in identification and diagnostic practices, and a broadening of case definitions (Joseph, 2011).

In previous years, the reliability of measuring IQ in children with ASDs had been scrutinized due to the social-communicative impairments and related behavioral issues common to ASDs that may confound the testing process (e.g., language delays, lack of imitative and reciprocal interactive skills, preoccupations or repetitive behaviors). Contemporary research has demonstrated, however, adequate predictability and stability of nonverbal intelligence in children with ASDs (Lord & Schopler, 1989a, 1989b; Sigman & Ruskin, 1999). In their 1989 study, Lord and Schopler demonstrated that among children with ASDs, IQ becomes more predictive as the minimum age at first assessment increases. In their sample, assessments of IQ at age two were not predictive of IQ at age seven ($r$ = .00), but were predictive by age three ($r$ = .52). These children also experienced sizeable changes in their absolute scores in intellectual ability across time; preschool children who initially scored in the mildly impaired range were as likely to

move into the unimpaired range (35%) as they were to remain in the impaired range (39%) at five year follow-up, and a portion of children in the severely impaired range moved into the mildly impaired range (38%). These findings have since been replicated in several studies (Sigman & Ruskin, 1999; Howlin, Goode, Hutton, & Rutter, 2004; Charman et al., 2005). Another study, conducted by Mayes and Calhoun (2003), assessed IQ data for 164 children with autism from age 3 to 15. They found that IQ increased with age; 67% of younger children in the sample demonstrated delayed speech abilities, while by school-age verbal and non-verbal IQs were no longer significantly different. It is essential to note, however, that the authors did not distinguish whether all of these children met criteria for an ASD (i.e., AD, PDD-NOS, etc.) or if by autism they were referring to the specific diagnosis of autistic disorder (AD).

Lord (1995) demonstrated the reliability of assigning an ASD diagnosis to a group of cognitively typical, 18- to 24-month-old children, finding that clinical diagnosis was highly stable between initial diagnosis at age two and re-evaluation at age three. This study also demonstrated that diagnosis using formal measures changed significantly between initial and follow-up evaluations among children that were initially diagnosed with an ASD and were also significantly developmentally delayed. Lord implicated the emergence of identifiable repetitive behaviors in children with autism, and improvements in rudimentary social skills in the children judged not to be autistic as key factors that allow the differentiation of children with autism from developmentally delayed children. A significant amount of research regarding diagnosis, intervention efficacy, and outcome has been conducted in persons with ASDs with cognitively normal profiles (Hurley & Levitas, 2007).

However, a subgroup of children with ASDs present with a low mental age (Low-MA), meaning the age equivalent at which a child is performing intellectually. For example, a child who receives an evaluation receives a standard score on a developmental measure, which allows comparison to the normative sample of children the same age who took this test. The raw score a child receives on an assessment can be converted to an age equivalent score, which allows a clinician to determine at what "age" the child is performing intellectually (e.g., a 2-year-old may be functioning at the level of an 18-month-old in his/her expressive language). As stated previously, intellectual disability (ID), defined by having an IQ below 70, commonly co-occurs in children with ASDs at a rate of between 50-70% (Fombonne, 2003). The diagnosis of intellectual disability, however, is not assigned until age six; prior to age six, a child demonstrating clinically significant delays is typically diagnosed with Developmental Delay (DD) or Global Developmental Delay (GDD).

There is a degree of diagnostic overlap between ASDs and ID: both are characterized by cognitive, adaptive, and social skill deficits, and often involve challenging and stereotyped behaviors (Matson & Shoemaker, 2009). Despite this, Osterling, Dawson, and Munson (2002) demonstrated that children who were later diagnosed with ASDs could be distinguished behaviorally from typically developing children and those with intellectual disabilities (IDs) by 12 months of age. They rated the social, communicative, and repetitive behaviors observed in home videos of infants, observing that children with ASDs looked at others less often and responded to their names less frequently than children with ID. Children with ID demonstrated a wide range of impairment across multiple developmental domains, while children with ASDs

demonstrated more severe deficits in communication. Other than the aforementioned studies, there is a paucity of literature regarding the reliability of an ASD diagnosis within the Low-MA subgroup. There is a lack of understanding of the diagnostic stability, symptomatology, and developmental profile of children with ASD and Low-MA, as well as how ASD-specific screening tools work within this sub-population.

While research has demonstrated that ASDs can be distinguished from intellectual disability (Osterling, Dawson, & Munson, 2002), the validity of using common, ASD-specific diagnostic tools in diagnosing autism in children with concurrent low mental age has not been demonstrated in the literature. For example, the Autism Diagnostic Interview-Revised (ADI-R) is only valid for children with a mental age above 24 months (Rutter, Le Couteur, Lord, & Faggioli, 2005). The Autism Diagnostic Observation Schedule (ADOS) has a Toddler module, which facilitates assessment in children as young as 12 months (Luyster et al., 2009), but not with mental ages below this point. However, other research claims that a more appropriate age of diagnosis is actually closer to 24 to 36 months, which would exclude all children with a low mental age, even if their chronological age were 24 to 36 months at time of evaluation. This calls into question the reliability of an early diagnosis of an ASD with co-occurring low mental age.

Low mental age also influences the efficacy and validity of autism-specific screening tools. The M-CHAT (Robins et al., 2001) assesses for the presence of behaviors that Inada, Kamio, and Koyama (2010) demonstrated were typically present in normal development primarily after 12 months of age. If the M-CHAT were administered to the parents of a child with a mental age below 12 months, these behaviors would likely be absent and potentially affect the validity of such a screener. Similarly, Dietz et al.

(2006), in a study regarding the efficacy of the Early Screening of Autistic Traits Questionnaire (ESAT), which included 31,724 children aged 14 to 15 months, demonstrated a high false positive rate. Of the 73 children that failed the ESAT and received an evaluation, only 18 (25%) received a diagnosis of an ASD, while 13 (18%) received a diagnosis of mental retardation (intellectual disability), 18 (25%) a language disorder, and 11 (15%) another DSM-IV diagnosis. This is potentially suggestive of a similar effect observed by Inada et al. (2010) regarding the M-CHAT, which is that some typical children do not demonstrate these behavioral milestones until shortly after 14 months of age. Matson and Shoemaker (2009) highlight a call for research that allows better understanding of the relationship and overlap between ID and ASD, citing a need for diagnostic instruments specific to co-occurring ID and ASD.

*ASD Symptom Severity and Developmental Progress: PDD-NOS and AD*

Under the DSM-IV-TR, the definition of AD inherently indicates a higher ASD symptom severity than PDD-NOS. To receive a diagnosis of AD a child must demonstrate deficits in Social and Communication abilities, as well as the presence of RRBs; PDD-NOS requires impairment in social interaction skills, and either impairment in communication or the presence of repetitive and stereotyped behaviors. As a result, some argue that PDD-NOS functions as an ambiguous and "catchall" diagnosis for children that do not fit the criteria of another developmental disorder, or who demonstrate some symptoms of autism, but also symptoms of other disorders (Filipek et al., 1999). More recently, new DSM-5 criteria have been adopted, which requires impairment in all three social-communication symptoms and two restricted/repetitive symptoms to qualify for an ASD diagnosis (American Psychiatric Association, 2013). It was designed on a

severity continuum, in which individuals with ASDs are rated from mild to severe according to their symptomatology, and the previous, individual diagnoses have been replaced with a single, overarching diagnostic category of Autism Spectrum Disorder. Previously, repetitive and restricted behavior patterns were not required to receive a diagnosis of PDD-NOS. Critics of the DSM-V criteria suggest that not only is there a loss in diagnostic nuance with the elimination of these single category diagnoses, but as many as 30% of children that would have previously received an ASD diagnosis by meeting criteria for the DSM-IV-TR's PDD-NOS entry, would no longer receive a diagnosis under the DSM-V criteria (Barton et al., 2013; Jashar, 2014).

Past studies have shown measurable differences in the cognitive, communication, and social relatedness skills between children with PDD-NOS and AD; children with PDD-NOS consistently measure at higher levels of ability and lower levels of impairment than children with AD (Cohen et al., 1986; Sevin et al., 1995). One such study (Bolte & Poustka, 2002) assessed the association between general cognitive level and adaptive behavior in 67 children with AD and PDD-NOS, utilizing the Vineland Adaptive Scales and the Wechsler Intelligence Scales. Across all participants, both in the AD and PDD-NOS groups, intellectual functioning was measured at a higher level than adaptive behavior skills. However, among the higher functioning individuals (n = 34) with an IQ greater than 70, including both AD and PDD-NOS, IQ and adaptive behavior level differed significantly, while among the lower functioning individuals (n =33) with an IQ less than 70, adaptive behavior level was also low. These findings suggest that IQ mediates adaptive functioning; individuals of high intellectual ability have significantly lower adaptive skills, while individuals with lower intellectual ability have equally low

adaptive skills. The authors also observed that individuals with AD and PDD-NOS scored highest on Daily living, and lowest on Socialization, with Communication in between these. These authors grouped their participants by IQ, however, so differences in IQ and adaptive functioning by diagnosis (AD vs. PDD-NOS) were not explored.

Several studies have demonstrated that children that received a diagnosis of PDD-NOS had a higher likelihood of no longer demonstrating symptoms of an ASD at follow-up, also referred to as "optimal outcome," than children with AD (Lord et al., 2006, Sutera et al., 2007; Helt et al., 2008; Berry, 2009). Specifically, Berry (2009) demonstrated that among these children with PDD-NOS at an initial evaluation, those with higher motor ability early in development, low ASD symptom severity, lower levels of repetitive behaviors, and higher adaptive and expressive language abilities were more likely to move off the autism spectrum at follow-up. However, more recent studies have suggested that PDD-NOS as a diagnostic category is ambiguous and may involve several subtypes with distinct profiles (Paul et al., 2004; Walker et al. 2004). Brennan (2014) described three subsets of children with PDD-NOS: the first cluster included those children with few autism symptoms and high cognitive scores at initial evaluation, 60% of whom no longer met criteria for PDD-NOS at follow-up; the second cluster inlcuded those children with higher autism symptoms and lower cognitive scores at initial, 89.5% of whom met criteria for an ASD at follow-up; the third cluster included those children with the lowest cognitive scores and highest autism symptom severity (social and communication), but without RRB's, 60% of whom receive a diagnosis of AD at follow-up.

Walker et al. (2004), in a study of 216 children with AD, and 21 with PDD-NOS, suggested that children with PDD-NOS demonstrate stronger cognitive and adaptive functioning and exhibited fewer repetitive and stereotyped behaviors than children with AD. Children in the PDD-NOS group received higher scores on the Communication, Social, and DLS subdomains than the AD group, and received lower scores on every domain of the ADI and ABC measuring severity of ASD symptoms. However, they also suggested that the PDD-NOS group could be divided into a high functioning group with mild symptoms of autism and mild cognitive delay (24%), a low functioning group that resembled AD with severe cognitive delays, but who had an onset of symptoms after 36 months or had too few symptoms (< 6) to meet criteria for AD, and a final group (52%) lacking stereotyped and repetitive behaviors. The finding within this last subgroup is similar to the findings in other research studies, particularly in Mandy et al. (2011), which looked at a sample of 256 children with PDD-NOS, and found that only 3% of these children presented with repetitive or stereotyped behaviors (Buitelaar et al., 1999; Allen et al., 2001). These findings suggest that while there may be trends regarding ASD symptomatology within the PDD-NOS group, there is a relative degree of heterogeneity in their initial presentation and outcomes.

*Current Study*

At present, children are diagnosed with autism due to observed social and communicative deficits. Children that present with a Low MA may express similar core deficits, but not because they have an ASD, but because that child's mental age is lower than the age that these prosocial and communicative behaviors typically emerge. Autism and Low MA may in fact co-occur, but it is also a possibility that these behavioral

deficits are solely the result of a severe developmental delay. Studies that asses the diagnostic stability of children that receive an ASD diagnosis and also have Low MA are required to help demonstrate if an ASD diagnosis is reliable with accompanying Low MA.

The current study aims to evaluate the diagnostic stability, symptomatology, and developmental profiles of children with ASDs and a Low-MA (ASD-Low MA). The Early Detection study has evaluated the effectiveness of the autism-specific screening instruments, the M-CHAT (Robins et al., 2001) and M-CHAT-R (Robins et al., 2014). Children that fail these screeners are offered an initial evaluation conducted when the child is between the ages of 18 and 30 months, which involves the administration of developmental and diagnostic measures, and provision of a diagnosis. Children are reassessed approximately two years later to assess current developmental functioning and diagnostic presentation. Children included in this sample received a diagnosis of Autistic Disorder (AD), Pervasive Developmental Disorder, Not Otherwise Specified (PDD-NOS), or ASD-Low MA at their initial evaluation and received a follow-up evaluation (diagnostic criteria for these disorders is specified in Methods).

We predicted that some of the children who initially received an ASD diagnosis with a Low-MA would not retain the diagnosis at the follow-up evaluation. We hypothesized that a low mental age would cause the child to present with autism-like symptoms and delays in development. However, with the development of language and prosocial behavior between the initial and follow-up evaluations, these children might no longer express deficits consistent with an ASD diagnosis. If a child's developmental delays persisted, as would be expected, we predicted that they would then show more

classic intellectual disability, or developmental delay, without features of an ASD. Additionally, the gold standard diagnostic instruments, the Autism Diagnostic Observation Schedule (Lord et al., 1999) and the Autism Diagnostic Interview (Lord, Rutter & Le Couteur, 1994) are not designed to evaluate autism in children with mental ages below the age of about 18 months, which suggests that the validity of an ASD diagnosis in Low-MA children needs to be assessed. We planned to compare the ASD-Low MA group's performance on the M-CHAT and M-CHAT-R, diagnostic stability, and symptoms and developmental levels at both time points to the Autistic Disorder (AD) and Pervasive Developmental Disorder, Not Otherwise Specified (PDD-NOS) groups. While there is robust research assessing diagnostic stability in AD, PDD-NOS has received less attention, which will also be addressed in this study.

There were several hypotheses for the current study:

1. We predicted that a majority of children with an initial diagnosis of ASD-Low MA would no longer meet diagnostic criteria for an ASD, but continue to demonstrate significant developmental delay. We employed a qualitative, descriptive analysis of the 25 children in the ASD Low-MA group to evaluate this hypothesis.
2. We predicted that children who received an initial diagnosis of AD or PDD-NOS would demonstrate a higher rate of diagnostic stability and likelihood of receiving an ASD diagnosis at follow-up than children who received an ASD-Low MA diagnosis. We also predicted that children who received an initial diagnosis of AD would have the highest diagnostic retention rate when compared to the PDD-NOS and ASD Low-MA groups.

a. We evaluated the differences in diagnostic stability among the AD (n = 111), PDD-NOS (n = 82), and ASD Low-MA (n = 25) groups utilizing a series of chi-square analysis, comparing all three groups by: diagnostic outcome with all diagnostic possibilities included, diagnostic outcome differentiating ASD versus non-ASD diagnostic outcomes, and finally diagnostic outcome differentiated by retain original diagnosis, other ASD diagnosis, and other diagnosis. We compared two groups at a time if a significant difference was identified for each hypothesis. If the assumptions of the chi-square analysis were violated, we utilized a Fisher's Exact Test to determine significant differences between groups.

    i. The sample size for the analysis including all diagnostic outcomes (N= 218) provided sufficient power (power= .80, alpha= .05) to detect an effect size of .27 (based on our chi-square design for degrees of freedom equal to 10, conventionally, phi of .07 is considered small, .21 medium, and .35 large), which is between medium and large (see Cohen, 1988).

    ii. The sample size for the analysis including ASD versus non-ASD diagnostic outcomes (N= 218) provided sufficient power (power= .80, alpha= .05) to detect an effect size of .19 (based on our chi-square design for degrees of freedom equal to 1, conventionally, phi of .10 is considered small, .30 medium, and .50 large), which is between small and medium.

iii. The sample size for the analysis including retain original diagnosis, other ASD diagnosis, and other diagnosis study (N= 218) provided sufficient power (power= .80, alpha= .05) to detect an effect size of .21, which is between small and medium.

3. We predicted that children who initially met criteria for ASD-Low MA would have lower developmental age-equivalents across all domains of functioning at both time points than children who initially met criteria for AD and PDD-NOS. Additionally, children with an initial diagnosis of ASD-Low MA would show significantly fewer developmental gains across time than children with AD and PDD-NOS.

    a. A series of univariate analysis of variance (ANOVA) were conducted to determine if there were differences between the AD, PDD-NOS, and ASD Low-MA groups in the amount of developmental gain made between the two evaluations in the following subdomains of the Mullen Early Scales of Learning: Visual Reception, Fine Motor, Receptive Language, and Expressive Language, and the following subdomains of the Vineland Adaptive Behavior Scales: Communication Domain: Expressive Language, Receptive Language; Daily Living Skills Domain: Personal, Domestic, Community, Interpersonal Relationships; Socialization Domain: Play and Leisure Time, and Coping Skills. This progress was measured by a ratio of change in age equivalent (mental age) divided by change in chronological age for each subdomain. If the overall effect of group on developmental progress was significant, post-hoc tests (LSD when equality of variances

is not violated, Games-Howell if it is violated) were conducted to specify the significance of the differences between each group.

b. The power for the analyses regarding the Mullen varied slightly between subdomains due to missing data points. The Visual Reception and Receptive Language subdomains had sample sizes of N = 165, which provided sufficient power (power= .80, alpha= .05) to detect an effect size of .24. The Fine Motor subdomain had an overall sample size of N = 165, which provided sufficient power to detect an effect size of .24; Expressive Language had a sample size N = 163, which allowed enough power to detect an effect size of .25. The power for these analyses allowed for the detection of effect sizes that are between small and medium.

c. The power for the analyses and effect size detection threshold regarding the Vineland are as follows: Expressive Language (N = 174), effect size of .24; Receptive Language (N = 174), effect size of .24; Personal (N = 169), effect size of .24; Domestic (N = 128), effect size of .28; Community (N = 130), effect size of .28; Interpersonal Relationships (N = 169), effect size of .24); Play and Leisure Time (N = 167), effect size of .24); Coping (N =119), effect size of .29. The power for these analyses allowed for the detection of effect sizes that are between small and medium.

4. We predicted that children with an initial diagnosis of ASD-Low MA would show significantly fewer ASD symptoms at follow-up than children who initially met criteria for AD and PDD-NOS. We predicted that the AD group would

demonstrate the most symptoms of an ASD at both Time 1 and Time 2, and have the smallest rates of symptom change across time when compared to the PDD-NOS and ASD Low-MA groups.

a. A series of mixed model design analyses were conducted to investigate the relationship between initial diagnosis and ASD symptomatology at Time 1 and Time 2 on the CARS, ADOS, and DSM-IV. If an interaction was present, a one-way ANOVA and post-hoc analyses (LSD or Games Howell) were performed to specify the significance of the differences between each group.
b. The overall sample size for analyses regarding the CARS (N= 200) provided sufficient power (power= .80, alpha= .05) to detect an effect size of .22, which is between small and medium. The overall sample size for analyses regarding the ADOS (N= 175) provided sufficient power (power= .80, alpha= .05) to detect an effect size of .24, which is between small and medium. The overall sample size for analyses regarding the DSM-IV (N= 192) provided sufficient power (power= .80, alpha= .05) to detect an effect size of .23, which is between small and medium.

## Methods

### *Participants*

Participants were recruited through the Early Detection study, which aims to assess the sensitivity and specificity of an ASD-specific screening questionnaire, the Modified Checklist for Autism in Toddlers (M-CHAT- Robins et al., 2001), and a revised version of the questionnaire, the M-CHAT-Revised (Robins, et al., 2014). Children were

enrolled in the study through two main referral sources: a pediatrician and an Early Intervention service provider. After enrollment, children received screening at an 18- or 24-month well-child visit at with their pediatrician, or through screening with an Early Intervention staff member. The majority of study participants were residents of Connecticut, Massachusetts, and/or Rhode Island at the time of their enrollment, which resulted in mostly rural and suburban, and les urban, living situations represented among participants. Children were screened for significant sensory and motor impairment prior to the initial evaluation. Children with such impairments were excluded from the study due to the interference this would present in the administration and interpretation of the standardized measures used to development and adaptive skills. Data included in the current study represent the subsection of the total sample collected for the Early Detection study that received an initial diagnosis of AD, PDD-NOS, and ASD-Low MA.

Diagnoses were based upon a child's performance on the Mullen Scales of Early Learning (Mullen) (Mullen, 1994), the Vineland Adaptive Behavior Scales (Vineland) (Sparrow, Balla, & Cicchetti, 1984), the Autism Diagnostic Observation Schedule (ADOS) (Lord et al., 2000), and the Childhood Autism Rating Scale (CARS) (Schloper et al., 1980) (see below for descriptions).

*Diagnostic Criteria*

The diagnostic criteria for Autistic Disorder (AD) used in the study was consistent with the DSM-IV-TR definition (Appendix A, Appendix C). Children who received an AD diagnosis demonstrated impairment in all three ASD symptom domains (communication, socialization, and repetitive and restricted interests and behaviors) with a total of six or more symptoms at their initial evaluation, with symptom onset before the

age of three. Specifically, a child must have presented with at least two symptoms in the social cluster, one symptom in the communication cluster, and one symptom in the restricted interests and repetitive behaviors cluster. At least one of the child's age equivalent scores on the Mullen Visual Reception, Receptive Language, and Expressive Language subscales must have been above 12 months.

The diagnostic criteria for Pervasive Developmental Disorder, Not Otherwise Specified (PDD-NOS) used in the study was consistent with the DSM-IV-TR definition (Appendix A, Appendix C). To receive a diagnosis of PDD-NOS a child presented with one symptom in the social cluster, other than a failure to develop peer relationships appropriate to the child's developmental level for initial evaluations only. The child also presented with at least one symptom in the communication cluster and/or one symptom in the repetitive and restricted interests and behaviors cluster. These children did not meet criteria for AD, Asperger's Disorder, or Rett's Syndrome, and their symptomatology could not have been better accounted for by another DSM-IV-TR disorder. At least one of the child's age equivalent scores on the Mullen Visual Reception, Receptive Language, and Expressive Language subscales must have been above 12 months.

Children that received an Autism Spectrum Disorder-Low MA diagnosis at the time of their evaluation presented with at least one symptom in the social cluster *other than* a failure to develop peer relationships appropriate to the child's developmental level, at least one symptom from the communication cluster, and/or at least one symptom in the repetitive and restricted interests and behaviors cluster (Appendix C). Additionally, these children received age equivalent scores on the Mullen subscales of Visual Reception,

Receptive Language, and Expressive Language subscales that were all less than or equal to 12 months.

To date, 219 children received an ASD diagnosis at their initial evaluation and later received a follow-up evaluation. Of these 219, 111 children received an initial diagnosis of AD, 83 received a diagnosis of PDD-NOS, and 25 received a diagnosis of ASD-Low MA through the study. For analyses, to assess for differences among these groups regarding all diagnostic outcomes, participants were coded into six categories based on their Time 2 diagnosis: Autistic Disorder, PDD-NOS, ASD Low-MA, Developmental Delay, Other Diagnosis, and No Diagnosis. Then, to better assess any significant differences found between the AD, PDD-NOS, and ASD Low-MA groups, participants were recoded based on whether their Time 2 diagnosis was an ASD (e.g., AD, PDD-NOS, or ASD Low-MA) or not (e.g., Developmental Delay, Other Diagnosis, No Diagnosis). Finally, participants from the AD, PDD-NOS, and ASD-Low MA groups were recoded into three groups: Retain Original Diagnosis, Other Autism Spectrum Disorder, and Other Diagnosis. Those participants in the Retain Original Diagnosis group received the same diagnosis at Time 2 as they did at Time 1. Those participants in the Other Autism Spectrum Disorder group received an ASD diagnosis at Time 2 that was different from their Time 1 diagnosis, but still on the autism spectrum. Those participants in the Other Diagnosis group received any other diagnostic outcome (e.g., Developmental Delay, Other Diagnosis, No Diagnosis).

*Demographic Information*

Demographic information for the children included in this study can be found in Table 1. The mean age of the children diagnosed with AD at their initial evaluation was

27.09 months (SD= 4.57 months) with a range of 17.74 months to 36.66 months. The mean age of the children diagnosed with PDD-NOS at their initial evaluation was 25.88 months (SD= 4.04) with a range of 18.03 months to 34.43 months. The mean age of the children diagnosed with ASD Low-MA at their initial evaluation was 23.60 months (SD= 4.50) with a range of 15.70 months to 31.80 months. A one-way multivariate analysis of variance (MANOVA) revealed a statistically significant difference in age based on a participant's initial diagnosis, $F$ (4, 428) = 3.67, $p < .006$; Wilk's $\Lambda = 0.935$, partial $\eta^2 = .033$. A Least Significant Difference (LSD) post-hoc analysis revealed significant differences in Time 1 age between the AD and ASD-Low MA ($p < .0005$) and the PDD-NOS and ASD-Low MA ($p = .023$), but not the AD and PDD-NOS groups ($p = .058$). The mean age of the children diagnosed with AD at their follow-up evaluation was 53.27 months (SD= 9.41 months) with a range of 41.38 months to 113.48 months. The mean age of the children diagnosed with PDD-NOS at their follow-up evaluation was 51.49 months (SD= 10.38) with a range of 38.12 months to 106.52 months. The mean age of the children diagnosed with ASD Low-MA at their follow-up evaluation was 49.37 months (SD= 6.48) with a range of 34.36 months to 62.13 months. A Least Significant Difference (LSD) post-hoc analysis revealed no significant differences in Time 2 age between the AD and ASD-Low MA ($p = .066$), the PDD-NOS and ASD-Low MA ($p = .33$), and the AD and PDD-NOS groups ($p = .2$).

Of the AD participants, 93 were male (83.8%) and 18 were female (16.2%). Of the PDD-NOS participants, 65 were male (78.3%) and 18 were female (21.7%). Of the ASD Low-MA participants, 20 were male (80%) and 5 were female (20%). There were

no significant differences in gender between the AD, PDD-NOS, and ASD Low-MA groups, $x^2$ (2, $N$ = 219) = .964, $p$ = .617.

Child ethnicity was collected through parent report using the following categories: White/European American, Hispanic/Latino- not Puerto Rican, Puerto Rican, African American, Caribbean or Caribbean American, Asian or Asian American, Native Hawaiian or Pacific Islander, Native American Indian, or Other. Participant responses were recoded into 5 categories: White/Caucasian, African American, Hispanic/Latino, Asian, Biracial. Of the AD participants, ethnicity information was available for 106 children. 87 (82.1%) of these children were White/European American, 2 (1.9%) children were African American, 9 (8.5%) children were Hispanic/Latino, 4 (3.8%) were Asian, and 4 (3.8%) were biracial. Ethnicity information was available for 79 PDD-NOS participants; 67 (84.8%) of these children were White/European American, 4 (5.1%) children were African American, 5 (6.3%) children were Hispanic/Latino, and 3 (3.8%) were Asian. Ethnicity information was available for all 25 ASD Low MA participants; 19 (76%) of these children were White/European American, 2 (8%) children were African American, 3 (12%) children were Hispanic/Latino, and 1 (4%) were biracial. Chi-square tests revealed no significant differences in ethnicity between the AD, PDD-NOS, and ASD Low MA groups, $x^2$ (8, $N$ = 210) = 7.450, $p$ = .490, however this result violated the assumptions of the chi-square test because 10 of the cells (66.7%) had an expected count that was less than five. Fisher's exact test revealed that the AD, PDD-NOS, and ASD Low MA groups were not significantly different in ethnicity, ($p$ =.365).

Maternal education was self-reported in the following categories: no degree or diploma, high school diploma/GED, vocational or technical degree, Associate's Degree,

Bachelor's Degree, Master's degree, and Graduate or Professional Degree (M.D., J.D., or Ph.D.). Maternal education information was available for 154 of the participants, in both the AD (n = 70), PDD-NOS (n = 65), and ASD Low-MA (n = 19) groups. The modal level of maternal education for the entire sample was High School Diploma/GED; the median level was Associates Degree. Chi-square tests did not reveal significant differences between the maternal education of the AD, PDD-NOS, and ASD Low-MA groups, $x^2$ (12, $N$ = 154) = 9.6746, $p$ = .645, though this result violated the assumptions of the chi-square test as 11 of the cells (52.4%) had an expected count that was less than five. Fisher's exact test revealed that the AD, PDD-NOS, and ASD Low MA groups were not significantly different in maternal education, ($p$ =.814). For analyses, maternal education was recoded as an ordinal variable, with a number (1-7) assigned to each tier in ascending order.

Family income was determined through self-report by indication of annual household income. Annual household income was stratified in $10,000 intervals, ranging from between less than $10,000 to greater than $100,000 (i.e., $10,000-20,000, $20,000-30,000, etc.). Family income information was available for 143 participants, in the AD (n = 63), PDD-NOS (n = 63), and ASD Low MA (n = 17) groups. Parents of the entire sample represented the full range of yearly incomes, the modal annual income was $10,000-20,000, and the median annual income for the entire sample was $50,000-60,000. Chi-square tests revealed no significant differences between the AD, PDD-NOS, and ASD Low-MA groups, $x^2$ (20, $N$ = 143) = 23.166, $p$ = .281, though the assumptions of this test were violated, as 17 of the cells (51.5%) had expected counts that were less than five. A Fisher's Exact Test, utilizing 2-sided Monte Carlo standards due to a high number

of data points, revealed no significant differences between groups (p = .332; lower bound = .328, upper bound = .336) based on 100,000 sampled tables. For analyses, family income was coded as an ordinal variable, with a number (1-11) assigned to each income tier in ascending order.

*Measures*

*ASD screening and diagnostic measures and developmental level*

Children received measures assessing developmental level and adaptive skills as part of a standardized battery. Study personnel administered measures to assess ASD symptomatology, including the Autism Diagnostic Observation Schedule, the Autism Diagnostic Interview, Revised, and the Childhood Autism Rating Scale (Lord, Risi, Lambrecht, Cook, Leventhal & DiLavore et al., 2000; Schopler, Reichler, & Renner, 1988; Rutter, Le Couteur, &Lord, 2003). The Modified Checklist for Autism in Toddlers was used to screen for ASD, and has demonstrated exceptional psychometric properties (Kleinman, Robins, Ventola, Pandey, 2008; Robins, Fein, Barton, Green, 2001).

Modified Checklist for Autism in Toddlers (M-CHAT) (Robins et al., 2001)

The Modified Checklist for Autism in Toddlers (M-CHAT) is a tool that screens for behaviors in children consistent with those observed in children with an ASD. The M-CHAT is a 23-item questionnaire in which parents respond "yes" or "no" answer to questions regarding their child's behavior (Robins et al., 2001). The measure was developed from the Checklist for Autism in Toddlers, which identifies children aged 18 months who are at risk for autism (CHAT- Baron-Cohen, Allen, & Gillberg, 1992; Baron-Cohen, Cox, & Baird, 1996). Of the questionnaire's 23 items, four are reverse-scored; for a typically developing child a parent would most likely answer "no," reducing

response bias (e.g., "Does your child ever seem oversensitive to noise?"). If a child fails three out of 23 total items, or two out of six "critical items," it is considered screening positive on the M-CHAT. If a child screens positive, their caregivers receive a follow-up phone screening, in which failed items are re-assessed, in more detail. If a child continues to fail the M-CHAT after phone screening, they qualify for a free initial developmental evaluation, and a subsequent follow-up evaluation two years afterward. The M-CHAT's internal reliability was demonstrated to be adequate for the 23-item checklist ($\alpha = .85$), as well as six "critical items" ($\alpha = .83$), in both the original study sample (Robins et al., 2001) and in an additional study ($\alpha = .85$, $\alpha = .83$; Kleinman et al., 2008). The majority of children included in the sample for this study were screened using the M-CHAT (n = 196); 100 from the AD group, 74 from the PDD-NOS group, and 22 from the ASD Low-MA group.

Modified Checklist for Autism in Toddlers-Revised (M-CHAT-R) (Robins et al., 2014)

The M-CHAT-R is the current measure used in the Early Detection Study, composed of 20 yes/no parent-report items that were reworded to improve comprehension. Additionally, the order of items was revised to counteract a tendency of parents to endorse "yes" for all items, examples were provided to increase the clarity of items, and three low-performing M-CHAT items were removed. As with the M-CHAT, children who screened positive (failing two of seven "best 7" items, or any three items) on the M-CHAT-R were given a follow-up phone interview. Children that continued to screen positive on the M-CHAT-R on the phone interview were offered free diagnostic evaluations. Published findings show that the M-CHAT-R is an effective screening tool when used in a low-risk, pediatric sample (with a cut-off of two failed items, sensitivity

= .94, specificity = .83). Eleven children from the AD group, nine from the PDD-NOS, and three from the ASD Low-MA group included in the sample for this study were screened using the M-CHAT-R.

Mullen Scales of Early Learning

The Mullen Scales of Early Learning (Mullen, 1994) is a standardized test of cognitive ability, intended to evaluate children between birth and age 68 months. Of its five subtests, Gross Motor, Visual Reception, Fine Motor, Expressive Language, and Receptive Language, all but the Gross Motor scale were administered in this study. The Early Learning Composite (ELC) is a score that is considered an overall estimate of a child's developmental age, and is generated by summing a child's performance across all four domains administered in this study. In each subtest, T-scores, percentile ranks, and age equivalents are produced, which reflect the child's current level of development in comparison to same-aged peers. The Mullen was normed on a nationally representative sample of 1,849 children (48.7% female, 51.3% male). It is a frequently used measure of developmental level and cognitive functioning in both typically developing children and children with developmental delays, and has demonstrated good reliability and validity. The Mullen demonstrates satisfactory internal consistency of .75 to .83. The test re-test reliability of the Mullen is .84 for younger children, and .76 for older children (Mullen, 1994).

For analyses, Time 1 age equivalent scores for each tested subdomain of the Mullen were subtracted from Time 2 age equivalent scores. This difference can be considered the amount of developmental progress, in months, made between Time 1 and Time 2. A positive difference indicates that a child made developmental gains within an

area between evaluations, a difference of zero would indicate no developmental gains, and a negative difference would indicate a lower developmental level at Time 2 compared to Time 1. The number of months between the Time 1 and Time 2 evaluations was calculated for each child.

The ratio estimate used in these analyses to assess developmental progress between evaluations (mental age divided by chronological age) was based upon a ratio used in a research study evaluating developmental progress in children with ASDs, and is common to outcome literature where standardized scores have restricted range (Sallows & Graupner, 2005). For example, for each child, the difference between Time 2 and Time 1 age equivalent scores from the Expressive Language subtest was divided by the number of months that passed between Time 1 and Time 2 for each specific child. This quotient represents the change in mental age, or the proportion of expected developmental gain over the actual time elapsed between evaluations. For example, if a child demonstrated 2 years of developmental gain as measured by the Mullen and 2 chronological years had elapsed between evaluations, this ratio is equal to 1. A similar quotient was generated for the Visual Reception, Fine Motor, and Receptive Language subdomains. In summary: a quotient greater than 1 indicates developmental progress in months greater than the amount of chronological time that had passed; a quotient equal to 1 indicates the same amount of developmental progress in months as the number of actual months that had passed; a quotient less than 1 indicates less developmental progress in months than the actual number of months that had passed.

Additionally, a ratio IQ estimate, which is commonly utilized in statistical analysis for studies involving participants with low mental ages/intelligence quotients

(Kanne, et al., 2011), was generated for each participant. This IQ estimate was generated for each of the Mullen subdomains, at both time points, and was equal to the quotient of mental age (represented by age equivalent score) divided by chronological age, multiplied by 100. This allows for more sensitive comparison of data than utilizing the T-scores of the Mullen, which are affected by a floor effect due to the low mental age and level of development of the participants across groups.

Vineland Adaptive Behavior Scales- Interview Edition

The Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchetti, 1984) is a standardized parent report interview that assesses a child's adaptive skills. It includes the domains of Communication, Daily Living, Socialization, and Motor Skills. The measure yields domain scores, standard scores for individual subscales, and an overall Adaptive Behavior Composite (ABC), which allows for comparison to a child's skills to same-aged peers. The Vineland has established reliability and validity (Sparrow, Balla, &Cicchetti, 1984) and it is frequently used with varied clinical populations. The Vineland is considered valid and commonly used in the assessment of children with developmental delays and ASDs, in both research and clinical applications (Klin, Carter, & Sparrow, 1997). For the range of ages included in the Early Detection sample, the Vineland demonstrates high internal consistency for its adaptive behavioral composite (.90) and domain scores (.80-.90). Test-retest reliability for the subdomains was adequate (ICC of .85 and higher), and inter-rater reliability for the adaptive composite score (.87) and domain scores (.75) were acceptable (Sparrow, Cicchetti, Balla 2005).

For analyses, the same method (Sallows & Graupner, 2005) to calculate a progress value for the Mullen subdomains (i.e., change in mental age, or age equivalent

score, in each subdomain divided by change in chronological age between Time 1 and Time 2) was utilized to create a progress value for the Vineland subdomains. Likewise, the same method to calculate a ratio IQ for the Mullen subdomains (the quotient of mental age divided by chronological age, multiplied by 100) was utilized to create a ratio IQ for the Vineland subdomains.

Autism Diagnostic Observation Schedule

The *Autism Diagnostic Observation Schedule - Generic* (ADOS) (Lord et al., 2000) is a semi-structured assessment designed to measure potential ASDs. Only behaviors viewed during test administration are scored on this measure. The ADOS-G includes four modules, one of which is administered depending on the child's expressive language level and chronological age. The current study used modules one and two during the initial evaluations (Time 1 and 2). For analyses, an ADOS calibrated severity score (CSS) was calculated for Time 1 and Time 2 according to the algorithms developed by Gotham, Pickles, & Lord (2008).

Childhood Autism Rating Scale

The *Childhood Autism Rating Scale* (CARS) (Schopler et al., 1980) is a behavior rating scale that consists of 15 subscales for rating aspects of autistic behavior. The scale is based on a clinician's direct observation of the child and parent report of behaviors. The scale yields a numerical score of ASD symptom severity. The score can be used to label a child's symptoms as non-autistic, mild, moderate, and severe. The CARS total score was used in analyses.

DSM-IV Summary Sheet

A summary sheet generated by the clinicians on the Early Detection Study, based on the official DSM-IV-TR criteria for ASDs, was completed for each participant at Time 1 and Time 2 (Appendix C). Participants received score of 1 for each symptom of an ASD they exhibited at the time of their evaluation. Symptoms were grouped into three sub-domains: Social, Communication, and Repetitive and Restricted Behaviors. The sum of all sub-domains is considered the Total score for this summary sheet. The total scores for the three sub-domains, as well as the Total score for the summary sheet, was utilized in analyses.

*Procedures*

When the child was between 16 and 30 months, their parent or caregiver completed the M-CHAT or the M-CHAT-R through the pediatrician's office or Early Intervention provider (Robins et al., 2001). If a child failed the M-CHAT or M-CHAT-R, their parents received a follow-up interview over the phone. Upon phone follow-up, if the child continued to fail, the family was offered a free developmental evaluation at the University of Connecticut. This evaluation was conducted by a licensed psychologist or a developmental pediatrician, and a graduate student. Participants lacking transportation were provided a free taxi service from their homes to the study. In some cases, study staff traveled to conduct evaluations at participating pediatric offices in two large towns with a high proportion of low SES families. The diagnosis of AD, PDD-NOS, and ASD Low-MA was made based on meeting cut-off scores on ADOS and CARS, as well as the Mullen and Vineland, derived by the clinicians on the Early Detection study (Appendix C). This first evaluation will be referred to as the initial evaluation, or Time 1.

Children became eligible to receive a follow-up evaluation, at Time 2, when they were 42 months or older, and were invited back to the University of Connecticut. This follow-up evaluation included the same measures assessing developmental and adaptive skills as at Time 1, and a diagnosis was made based on meeting cut-off scores on the ADOS and CARS, and the Mullen and the Vineland.

All of the children in this study, in the AD, PDD-NOS, and ASD-Low MA groups, failed (screened positive on) the M-CHAT or M-CHAT-R, as well as the phone interview, and received an initial and follow-up evaluation.

## Results

*Diagnostic Stability of AD, PDD-NOS and ASD Low-MA: All Diagnostic Outcomes*

A chi-square analysis was conducted to compare possible differences in the diagnostic stability of the AD (N = 111), PDD-NOS (N = 83), and ASD Low-MA (N = 25) groups, when considering every possible diagnostic outcome. A Time 1 diagnosis was missing for one child in the PDD-NOS group, so that participant was not included in these analyses.

The Time 2 diagnostic outcomes for the 111 children that received a diagnosis of AD at Time 1 are as follows: AD, n = 75 (67.6%); PDD-NOS, n = 21 (18.9%); ASD Low-MA, n = 0; Developmental Delay, n = 2 (1.8%); Other Diagnosis, n = 1 (.9%); No Diagnosis, n = 12 (10.8%). The Time 2 diagnostic outcomes for the 82 children that received a diagnosis of PDD-NOS at Time 1 are as follows: AD, n = 28 (34.1%); PDD-NOS, n = 32 (39%); ASD Low-MA, n = 0; Developmental Delay, n = 9 (11%); Other Diagnosis, n = 3 (3.7%); No Diagnosis, n = 10 (12.2%). The Time 2 diagnostic outcomes for the 25 children that received a diagnosis of ASD Low-MA at Time 1 are as follows:

AD, n = 18 (72%); PDD-NOS, n = 1 (4%); ASD Low-MA, n = 5 (20%); Developmental Delay, n = 1 (1.4%); Other Diagnosis, n = 0; No Diagnosis, n = 0 at Time 2 (Table 2). This initial chi-square analysis revealed that there was an overall difference in the diagnostic stability based on Time 1 diagnosis of either AD, PDD-NOS, and ASD Low-MA, $x^2$ (10, $N$ = 218) = 74.83, $p$ = >.0005, Phi = .586. The assumptions of this chi-square analysis were violated, as nine of the cells had expected counts that were less than five. Fisher's Exact Test (Monte Carlo Significance, 2-sided) was significant for differences between groups (p <.0005) based on 100,000 sampled tables.

*Diagnostic Stability: ASD vs. Non-ASD Time 2 Diagnosis*

Of the 111 children that received an AD diagnosis at Time 1, 96 (86.5%) received an ASD diagnosis, while 15 (13.5%) received a non-ASD diagnosis. Of the 82 children that received a PDD-NOS diagnosis at Time 1, 60 (73.2%) received an ASD diagnosis, while 22 (26.8%) received a non-ASD diagnosis. Of the 25 children that received an ASD Low-MA diagnosis at Time 1, 24 (96%) received an ASD diagnosis, while 1 (4%) received a non-ASD diagnosis. Chi-square analysis revealed a significant difference between the AD, PDD-NOS, and ASD Low-MA groups in their likelihood to remain on the autism spectrum at Time 2, $x^2$(2, $N$ = 218) = 9.35 , $p$ = .009, Phi = .207 (Table 3).

Follow-up chi-square analyses were then conducted to specify the significant differences between these groups by individually comparing the AD and PDD-NOS, the AD and ASD Low-MA, and the PDD-NOS and ASD Low-MA groups on whether they retain an ASD diagnosis or receive a non-ASD diagnosis at Time 2. The AD and PDD-NOS groups were significantly different, as participants from the AD group (86.5%) were more likely to receive an ASD diagnosis at Time 2 than those in the PDD-NOS group

(73.2%), $x^2(1, N = 193) = 5.39$ , $p = .026$, Phi = .167. The AD and ASD Low-MA groups were not significantly different, as participants from the AD group (86.5%) and ASD Low-MA (96%) had a similar likelihood of receiving an ASD diagnosis at Time 2, $x^2(1, N = 136) = 1.78$ , $p = .304$, Phi = -.114. The PDD-NOS and ASD Low-MA groups were significantly different, as participants from the ASD Low-MA group (96%) were more likely to receive an ASD diagnosis at Time 2 than those in the PDD-NOS group (73.2%), $x^2(1, N = 107) = 5.92$ , $p = .013$, Phi = -.235.

*Diagnostic Stability: Retain Original Diagnosis, Other ASD, Other Diagnosis*

Of the 111 children in the AD group, 75 (67.6%) retained that diagnosis, 21 (18.9%) received another ASD diagnosis, and 15 (13.5%) received another diagnosis at Time 2. Of the 82 children in the PDD-NOS group, 32 (39%) retained that diagnosis, 28 (34.1%) received another ASD diagnosis, and 22 (26.8%) received another diagnosis at Time 2. Of the 25 children in the ASD Low-MA group, 5 (20%) retained that diagnosis, 19 (76%) received another ASD diagnosis, and 1 (4%) received another diagnosis at Time 2. Chi-square analysis revealed a significant difference between the AD, PDD-NOS, and ASD Low-MA groups in their likelihood to retain their initial diagnosis, receive a different autism spectrum diagnosis, and receive another diagnosis at Time 2, $x^2(4, N = 218) = 42.29$ , $p = >.0005$, Phi = .440 (Table 4).

Follow-up chi-square analyses were then conducted to specify the significant differences between these groups by individually comparing the AD and PDD-NOS, the AD and ASD Low-MA, and the PDD-NOS and ASD Low-MA groups on whether they retained their Time 1 diagnosis, received another ASD diagnosis, or receive another diagnosis at Time 2. The AD and PDD-NOS groups were significantly different, as

participants from the AD group (67.6%) were more likely to retain their Time 1 diagnosis than those in the PDD-NOS group (39%), $x^2(2, N = 193) = 15.59$ , $p < .0005$, Phi = .284. The AD and ASD Low-MA groups were significantly different, as participants from the AD group (67.6%) were more likely to retain their Time 1 diagnosis than those in the ASD Low-MA group (20%) at Time 2, $x^2(2, N = 136) = 32.02$ , $p < .0005$, Phi = .485. The PDD-NOS and ASD Low-MA groups were significantly different, as participants from the ASD Low-MA group (76%) were more likely to receive an ASD diagnosis at Time 2 than those in the PDD-NOS group (34.1%), $x^2(2, N = 107) = 14.291$ , $p = .001$, Phi = .365.

*Developmental Progress by Diagnosis – Mullen (AD, PDD-NOS, ASD Low-MA): Visual Reception*

Time 1 and Time 2 Visual Reception data were available for 80 children from the AD group, 66 children from the PDD-NOS group, and 19 children from the ASD Low-MA group. At Time 1, the mean Visual Reception ratio IQ for the AD group was 66.81 (SD = 19.86), PDD-NOS was 73.33 (SD = 17.09), and ASD Low-MA was 40.28 (SD = 12.08). At Time 2, the mean Visual Reception ratio IQ for the AD group was 74.26 (SD = 28.68), PDD-NOS was 85.92 (SD = 27.74), and ASD Low-MA was 41.72 (SD = 20.32) (Table 5). A mixed model design analysis did not demonstrate a significant interaction between Time 1 diagnosis and time-point (Time 1 and Time 2), $F(2, 162) = 1.655$, $p = 1.94$.

The mean mental growth rates for Visual Reception, which were .86 (SD = .59) for the AD group, .99 (SD = .51) for the PDD-NOS group, and .48 (SD = .43) for the ASD Low-MA group, were significantly different by group, $F(1, 162) = 6.69$, $p = .002$. A

post-hoc LSD analysis revealed significant differences in VR progress between the AD and ASD-Low MA ($p = .007$) and the PDD-NOS and ASD-Low MA ($p < .0005$) groups, but not the AD and PDD-NOS groups ($p = .129$) (Table 6).

*Fine Motor*

Time 1 and Time 2 Fine Motor data were available for 78 children from the AD group, 66 children from the PDD-NOS group, and 20 children from the ASD Low-MA group. At Time 1, the mean Fine Motor ratio IQ for the AD group was 73.81 (SD = 26.36), PDD-NOS was 76.74 (SD = 13.99), and ASD Low-MA was 56.73 (SD = 14.00). At Time 2, the mean Fine Motor ratio IQ for the AD group was 70.37 (SD = 22.82), PDD-NOS was 75.61 (SD = 20.89), and ASD Low-MA was 39.75 (SD = 13.05) (Table 5). A mixed model design analysis demonstrated a significant interaction between Time 1 diagnosis and time-point (Time 1 and Time 2), $F(2, 161) = 3.831$, $p = .024$.

The mean mental growth rates for Fine Motor, which were .69 (SD = .43) for the AD group, .74 (SD = .41) for the PDD-NOS group, and .27 (SD = .24) for the ASD Low-MA group, were significantly different by group, $F(1, 161) = 10.73$, $p < .0005$. Levene's test for equality of variances was found to be violated for the present analysis, $F(2,161) = 3.28$, $p = .04$. Due to this, a Games-Howell post-hoc analysis was conducted in lieu of an LSD analysis, which revealed significant differences in FM progress between the AD and ASD-Low MA ($p < .0005$) and the PDD-NOS and ASD-Low MA ($p < .0005$), but not the AD and PDD-NOS groups ($p = .74$) (Table 6).

*Receptive Language*

Time 1 and Time 2 Receptive Language data were available for 79 children from the AD group, 66 children from the PDD-NOS group, and 20 children from the ASD

Low-MA group. At Time 1, the mean Receptive Language ratio IQ for the AD group was 43.78 (SD = 23.25), PDD-NOS was 54.55 (SD = 22.15), and ASD Low-MA was 30.00 (SD = 16.14). At Time 2, the mean Receptive Language ratio IQ for the AD group was 63.10 (SD = 28.62), PDD-NOS was 77.01 (SD = 29.05), and ASD Low-MA was 32.05 (SD = 17.93) (Table 5). A mixed model design analysis demonstrated a significant interaction between Time 1 diagnosis and time-point (Time 1 and Time 2), $F(2, 162) = 3.988$, $p = .02$.

The mean mental growth rates for Receptive Language, which were .88 (SD = .60) for the AD group, .99 (SD = .51) for the PDD-NOS group, and .34 (SD = .34) for the ASD Low-MA group, were significantly different by group, $F(1, 162) = 11.37$, $p < .0005$. Levene's test for equality of variances was found to be violated for the present analysis, $F(2,162) = 4.31$, $p = .015$. A Games-Howell post-hoc analysis revealed significant differences in RL progress between the AD and ASD-Low MA ($p < .0005$) and the PDD-NOS and ASD-Low MA ($p < .0005$), but not the AD and PDD-NOS groups ($p = .467$) (Table 6).

*Expressive Language*

Time 1 and Time 2 Expressive Language data were available for 79 children from the AD group, 65 children from the PDD-NOS group, and 19 children from the ASD Low-MA group. At Time 1, the mean Expressive Language ratio IQ for the AD group was 48.11 (SD = 23.09), PDD-NOS was 58.34 (SD = 18.64), and ASD Low-MA was 33.55 (SD = 12.98). At Time 2, the mean Expressive Language ratio IQ for the AD group was 60.63 (SD = 26.67), PDD-NOS was 71.87 (SD = 24.72), and ASD Low-MA was 33.49 (SD = 20.36) (Table 5). A mixed model design analysis did not demonstrate a

significant interaction between Time 1 diagnosis and time-point (Time 1 and Time 2), $F(2, 160) = 2.34$, $p = .094$.

The mean mental growth rates for Expressive Language, which were .77 (SD = .52) for the AD group, .85 (SD = .47) for the PDD-NOS group, and .35 (SD = .34) for the ASD Low-MA group, were significantly different by group, $F(1, 160) = 8.26$, $p < .0005$. A post-hoc LSD analysis revealed significant differences in EL progress between the AD and ASD-Low MA (p = .001) and the PDD-NOS and ASD-Low MA (p < .0005), but not the AD and PDD-NOS groups (p = .33) (Table 6).

*Developmental Progress by Diagnosis – Vineland (AD, PDD-NOS, ASD Low-MA): Communication: Expressive and Receptive Language*

Time 1 and Time 2 Expressive Language data were available for 83 children from the AD group, 72 children from the PDD-NOS group, and 19 children from the ASD Low-MA group. At Time 1, the mean Expressive Language ratio IQ for the AD group was 38.33 (SD = 18.22), PDD-NOS was 49.37 (SD = 16.72), and ASD Low-MA was 32.27 (SD = 12.07). At Time 2, the mean Expressive Language ratio IQ for the AD group was 57.67 (SD = 60.49), PDD-NOS was 62.50 (SD = 26.52), and ASD Low-MA was 28.47 (SD = 13.39) (Table 7). A mixed model design analysis did not demonstrate a significant interaction between Time 1 diagnosis and time-point (Time 1 and Time 2), $F(2, 171) = 2.18$, $p = .116$.

The mean mental growth rates for Expressive Language, which were .80 (SD = 1.32) for the AD group, .76 (SD = .48) for the PDD-NOS group, and .24 (SD = .98) for the ASD Low-MA group, were not significantly different by group, $F(1, 171) = 2.63$, $p = .075$. However, a post-hoc LSD analysis revealed significant differences in EL progress

between the AD and ASD-Low MA (p = .026) and the PDD-NOS and ASD-Low MA (p = .04), but not the AD and PDD-NOS groups (p = .82) (Table 8).

Time 1 and Time 2 Receptive Language data were available for 83 children from the AD group, 72 children from the PDD-NOS group, and 19 children from the ASD Low-MA group. At Time 1, the mean Receptive Language ratio IQ for the AD group was 44.27 (SD = 22.13), PDD-NOS was 56.73 (SD = 23.09), and ASD Low-MA was 33.11 (SD = 13.30). At Time 2, the mean Receptive Language ratio IQ for the AD group was 60.39 (SD = 31.53), PDD-NOS was 72.39 (SD = 33.92), and ASD Low-MA was 36.50 (SD = 24.74) (Table 7). A mixed model design analysis did not demonstrate a significant interaction between Time 1 diagnosis and time-point (Time 1 and Time 2), $F(2, 171) = 1.26$, $p = .286$.

The mean mental growth rates for Receptive Language, which were .77 (SD = .62) for the AD group, .91 (SD = .7) for the PDD-NOS group, and .41 (SD = .47) for the ASD Low-MA group, were significantly different by group, $F(1, 171) = 4.78$, $p = .01$. A post-hoc LSD analysis revealed significant differences in RL progress between the AD and ASD-Low MA (p = .026) and the PDD-NOS and ASD-Low MA (p = .002), but not the AD and PDD-NOS groups (p = .18) (Table 8).

*Daily Living Skills: Personal, Domestic, Community*

Time 1 and Time 2 Personal data were available for 81 children from the AD group, 70 children from the PDD-NOS group, and 18 children from the ASD Low-MA group. At Time 1, the mean Personal ratio IQ for the AD group was 59.54 (SD = 13.44), PDD-NOS was 59.77 (SD = 17.65), and ASD Low-MA was 56.11 (SD = 18.49). At Time 2, the mean Personal ratio IQ for the AD group was 54.63 (SD = 17.62), PDD-NOS

was 57.09 (SD = 17.49), and ASD Low-MA was 40.07 (SD = 8.53) (Table 7). A mixed model design analysis demonstrated a significant interaction between Time 1 diagnosis and time-point (Time 1 and Time 2), F(2, 166) = 4.138, p =.018.

The mean mental growth rates for Personal, which were .50 (SD = .35) for the AD group, .56 (SD = .32) for the PDD-NOS group, and .23 (SD = .26) for the ASD Low-MA group, were significantly different by group, $F(1, 166) = 7.04$, $p = .001$. A post-hoc LSD analysis revealed significant differences in Personal progress between the AD and ASD-Low MA (p = .002) and the PDD-NOS and ASD-Low MA (p < .0005), but not the AD and PDD-NOS groups (p = .28) (Table 8).

Time 1 and Time 2 Domestic data were available for 59 children from the AD group, 57 children from the PDD-NOS group, and 12 children from the ASD Low-MA group. At Time 1, the mean Domestic ratio IQ for the AD group was 66.60 (SD = 22.17), PDD-NOS was 70.17 (SD = 26.40), and ASD Low-MA was 65.46 (SD = 19.05). At Time 2, the mean Domestic ratio IQ for the AD group was 63.50 (SD = 32.17), PDD-NOS was 58.68 (SD = 27.69), and ASD Low-MA was 33.47 (SD = 16.65) (Table 7). A mixed model design analysis demonstrated a significant interaction between Time 1 diagnosis and time-point (Time 1 and Time 2), F(2, 125) = 5.62, p = .005.

The mean mental growth rates for Domestic, which were .63 (SD = .66) for the AD group, .47 (SD = .53) for the PDD-NOS group, and .02 (SD = .23) for the ASD Low-MA group, were significantly different by group, $F(1, 125) = 5.76$, $p = .004$. Levene's test for equality of variances was found to be violated for the present analysis, $F(2,125) = 3.75$, p = .026. A Games-Howell post-hoc analysis revealed significant differences in

Domestic progress between the AD and ASD-Low MA ($p < .0005$) and the PDD-NOS and ASD-Low MA ($p < .0005$), but not the AD and PDD-NOS groups ($p = .34$) (Table 8).

Time 1 and Time 2 Community data were available for 61 children from the AD group, 58 children from the PDD-NOS group, and 11 children from the ASD Low-MA group. At Time 1, the mean Community ratio IQ for the AD group was 39.91 (SD = 25.89), PDD-NOS was 40.72 (SD = 27.85), and ASD Low-MA was 22.11 (SD = 23.57). At Time 2, the mean Community ratio IQ for the AD group was 58.20 (SD = 72.08), PDD-NOS was 52.72 (SD = 24.40), and ASD Low-MA was 28.20 (SD = 24.67) (Table 7). A mixed model design analysis did not demonstrate a significant interaction between Time 1 diagnosis and time-point (Time 1 and Time 2), $F(2, 127) = .323$, $p = .725$.

The mean mental growth rates for Community, which were .82 (SD = 1.61) for the AD group, .67 (SD = .58) for the PDD-NOS group, and .37 (SD = .53) for the ASD Low-MA group, were not significantly different by group, $F(1, 127) = 1.06$, $p = .48$.

*Socialization: Interpersonal Relationships, Play and Leisure Time, Coping Skills*

Time 1 and Time 2 Interpersonal Relationships data were available for 81 children from the AD group, 70 children from the PDD-NOS group, and 18 children from the ASD Low-MA group. At Time 1, the mean Interpersonal Relationships ratio IQ for the AD group was 34.64 (SD = 14.14), PDD-NOS was 44.66 (SD = 17.00), and ASD Low-MA was 33.33 (SD = 10.53). At Time 2, the mean Interpersonal Relationships ratio IQ for the AD group was 44.04 (SD = 25.95), PDD-NOS was 54.12 (SD = 27.79), and ASD Low-MA was 22.54 (SD = 9.86) (Table 7). A mixed model design analysis demonstrated a significant interaction between Time 1 diagnosis and time-point (Time 1 and Time 2), $F(2, 166) = 4.719$, $p = .01$.

The mean mental growth rates for Interpersonal Relationships, which were .54 (SD = .55) for the AD group, .66 (SD = .58) for the PDD-NOS group, and .13 (SD = .15) for the ASD Low-MA group, were significantly different by group, $F(1, 166) = 6.94$, $p = .001$. Levene's test for equality of variances was found to be violated for the present analysis, $F(2,166) = 5.86$, $p = .003$. A Games-Howell post-hoc analysis revealed significant differences in Interpersonal Relationships progress between the AD and ASD-Low MA ($p < .0005$) and the PDD-NOS and ASD-Low MA ($p < .0005$), but not the AD and PDD-NOS groups ($p = .42$) (Table 8).

Time 1 and Time 2 Play and Leisure data were available for 81 children from the AD group, 69 children from the PDD-NOS group, and 17 children from the ASD Low-MA group. At Time 1, the mean Play and Leisure ratio IQ for the AD group was 36.13 (SD = 16.79), PDD-NOS was 50.45 (SD = 20.81), and ASD Low-MA was 34.22 (SD = 21.19). At Time 2, the mean Play and Leisure ratio IQ for the AD group was 40.51 (SD = 25.72), PDD-NOS was 48.57 (SD = 24.12), and ASD Low-MA was 22.10 (SD = 13.27) (Table 7). A mixed model design analysis did not demonstrate a significant interaction between Time 1 diagnosis and time-point (Time 1 and Time 2), $F(2, 164) = 2.85$, $p = .061$.

The mean mental growth rates for Play and Leisure, which were .46 (SD = .57) for the AD group, .47 (SD = .51) for the PDD-NOS group, and .14 (SD = .25) for the ASD Low-MA group, were significantly different by group, $F(1, 164) = 3.06$, $p = .049$. Levene's test for equality of variances was found to be violated for the present analysis, $F(2,164) = 3.55$, $p = .031$. A Games-Howell post-hoc analysis revealed significant differences in Play and Leisure progress between the AD and ASD-Low MA ($p = .001$)

and the PDD-NOS and ASD-Low MA (p = .001), but not the AD and PDD-NOS groups (p = .99) (Table 8).

Time 1 and Time 2 Coping data were available for 56 children from the AD group, 52 children from the PDD-NOS group, and 11 children from the ASD Low-MA group. At Time 1, the mean Coping ratio IQ for the AD group was 42.99 (SD = 23.42), PDD-NOS was 40.34 (SD = 18.33), and ASD Low-MA was 41.47 (SD = 16.33). At Time 2, the mean Coping ratio IQ for the AD group was 45.78 (SD = 26.94), PDD-NOS was 59.07 (SD = 30.05), and ASD Low-MA was 22.38 (SD = 13.62) (Table 7). A mixed model design analysis demonstrated a significant interaction between Time 1 diagnosis and time-point (Time 1 and Time 2), F(2, 116) = 7.48, p =.001.

The mean mental growth rates for Coping, which were .50 (SD = .64) for the AD group, .82 (SD = .63) for the PDD-NOS group, and .06 (SD = .27) for the ASD Low-MA group, were significantly different by group, $F(1, 116) = 8.49$, $p < .0005$. Levene's test for equality of variances was found to be violated for the present analysis, $F(2,116) = 3.35$, p = .038. A Games-Howell post-hoc analysis revealed significant differences in Domestic progress between the AD and ASD-Low MA (p = .002) and the PDD-NOS and ASD-Low MA (p < .0005), and the AD and PDD-NOS groups (p = .025) (Table 8).

*Autism Symptomatology by Diagnosis – CARS:*

A CARS Total score was available at both Time 1 and Time 2 for 103 participants in the AD group, 76 in the PDD-NOS group, and 21 in the ASD Low-MA group.

At Time 1, the mean CARS Total score was 35.33 (SD = 4.41) for the AD group, 28.2 (SD = 3.59) for the PDD-NOS group, and 35.12 (SD = 5.69) for the ASD Low-MA group. At Time 2, the mean CARS Total score was 30.95 (SD = 7.08) for the AD group,

26.76 (SD = 5.79) for the PDD-NOS group, and 36.33 (SD = 4.89) for the ASD Low-MA group. A mixed model design analysis demonstrated a significant interaction between Time 1 Diagnosis and time-point (Time 1 and Time 2), $F(2, 197) = 9.12$, $p < .0005$.

Due to this significant interaction, a one-way ANOVA was performed to compare the AD, PDD-NOS, and ASD Low-MA groups using the change in CARS Total score between Time 1 and Time 2 (i.e., Time 2 CARS Total – Time 1 CARS Total = Change in CARS Total.). The mean change in CARS Total score was -4.37 (SD = 6.59) for the AD group, -1.45 (SD = 5.79) for the PDD-NOS group, and 1.21 (SD = 6.84) for the ASD Low-MA group, and was significantly different among groups, $F(2, 197) = 9.12$, $p < .0005$. A post-hoc LSD analysis revealed significant differences in CARS Total Score change between the AD and ASD-Low MA ($p < .0005$), and AD and PDD-NOS ($p = .003$), but not the PDD-NOS and ASD-Low MA ($p = .089$) groups (Table 9).

*Autism Symptomatology by Diagnosis – ADOS:*

An ADOS total score was available at both Time 1 and Time 2 for 82 participants in the AD group, 73 in the PDD-NOS group, and 20 in the ASD Low-MA group.

At Time 1, the mean ADOS CSS score was 7.89 (SD = 1.79) for the AD group, 5.66 (SD = 2.16) for the PDD-NOS group, and 7.3 (SD = 1.56) for the ASD Low-MA group. At Time 2, the mean ADOS CSS score was 5.9 (SD = 2.57) for the AD group, 5.05 (SD = 2.75) for the PDD-NOS group, and 7.10 (SD = 2.22) for the ASD Low-MA group. A mixed model design analysis demonstrated a significant interaction between Time 1 Diagnosis and time-point (Time 1 and Time 2), $F(2, 172) = 5.81$, $p = .004$).

Due to this significant interaction, a 1-way ANOVA was performed to compare the AD, PDD-NOS, and ASD Low-MA groups using the change in ADOS CSS score

between Time 1 and Time 2. The mean change in ADOS CSS score was -1.99 (SD = 2.99) for the AD group, -.60 (SD = 2.95) for the PDD-NOS group, and -.16 (SD = 2.12) for the ASD Low-MA group, and was significantly different among groups, $F(2, 171) = 5.82$, $p = .004$. A post-hoc LSD analysis revealed significant differences in ADOS CSS score change between the AD and ASD-Low MA (p = .014) and the AD and PDD-NOS (p = .003), but not the PDD-NOS and ASD-Low MA (p = .55) groups (Table 10).

*Autism Symptomatology by Diagnosis – DSM-IV-TR Ratings:*

*DSM-IV: Total*

A DSM Total score was available at both Time 1 and Time 2 for 95 participants in the AD group, 74 in the PDD-NOS group, and 23 in the ASD Low-MA group.

At Time 1, the mean DSM Total score was 7.15 (SD = 1.05) for the AD group, 4.2 (SD = .92) for the PDD-NOS group, and 6.35 (SD = 1.85) for the ASD Low-MA group. At Time 2, the mean DSM Total score was 5.88 (SD = 2.43) for the AD group, 4.96 (SD = 2.39) for the PDD-NOS group, and 6.87 (SD = 1.29) for the ASD Low-MA group. A mixed model design analysis demonstrated a significant interaction between Time 1 Diagnosis and time-point (Time 1 and Time 2), $F(2, 189) = 16.69$, $p < .0005$.

Due to this significant interaction, a 1-way ANOVA was performed to compare the AD, PDD-NOS, and ASD Low-MA groups using the change in DSM Total score between Time 1 and Time 2. The mean change in DSM Total score was -1.37 (SD = 2.57) for the AD group, .76 (SD = 2.34) for the PDD-NOS group, and .52 (SD = 1.83) for the ASD Low-MA group, and was significantly different among groups, $F(2, 191) = 18.03$, $p < .0005$. A post-hoc LSD analysis revealed significant differences in DSM Total

score change between the AD and ASD-Low MA (p = .001), and AD and PDD-NOS (p < .0005), but not the PDD-NOS and ASD-Low MA (p = .68) groups (Table 11).

*DSM-IV: Social*

A DSM Social score was available at both Time 1 and Time 2 for 96 participants in the AD group, 74 in the PDD-NOS group, and 23 in the ASD Low-MA group.

At Time 1, the mean DSM Social score was 3.33 (SD = .69) for the AD group, 2.04 (SD = .85) for the PDD-NOS group, and 2.91 (SD = .96) for the ASD Low-MA group. At Time 2, the mean DSM Social score was 2.4 (SD = 1.28) for the AD group, 1.65 (SD = 1.22) for the PDD-NOS group, and 2.91 (SD = .73) for the ASD Low-MA group. A mixed model design analysis demonstrated a significant interaction between Time 1 Diagnosis and time-point (Time 1 and Time 2), $F(2, 189) = 6.37$, $p < .0005$.

Due to this significant interaction, a 1-way ANOVA was performed to compare the AD, PDD-NOS, and ASD Low-MA groups using the change in DSM Social score between Time 1 and Time 2. The mean change in DSM Social score was -.94 (SD = 1.37) for the AD group, -.39 (SD =1.25) for the PDD-NOS group, and .00 (SD = 1.17) for the ASD Low-MA group, and was significantly different among groups, $F(2, 190) = 6.61$, $p = .002$. A post-hoc LSD analysis revealed significant differences in DSM Social score change between the AD and ASD-Low MA (p = .002), and AD and PDD-NOS (p = .007), but not the PDD-NOS and ASD-Low MA (p = .21) groups (Table 12).

*DSM-IV: Communication*

A DSM Communication score was available at both Time 1 and Time 2 for 65 participants in the AD group, 74 in the PDD-NOS group, and 23 in the ASD Low-MA group.

At Time 1, the mean DSM Communication score was 1.96 (SD = .50) for the AD group, 1.46 (SD = .53) for the PDD-NOS group, and 1.7 (SD = .47) for the ASD Low-MA group. At Time 2, the mean DSM Communication score was 1.95 (SD = .95) for the AD group, 1.88 (SD = .92) for the PDD-NOS group, and 2 (SD = .60) for the ASD Low-MA group. A mixed model design analysis demonstrated a significant interaction between Time 1 Diagnosis and time-point (Time 1 and Time 2), $F(2, 189) = 3.82$, $p = .024$.

Due to this significant interaction, a 1-way ANOVA was performed to compare the AD, PDD-NOS, and ASD Low-MA groups using the change in DSM Communication score between Time 1 and Time 2. The mean change in Communication score was -.04 (SD = 1.11) for the AD group, .42 (SD =1.03) for the PDD-NOS group, and .30 (SD = .70) for the ASD Low-MA group, and was significantly different among groups, $F(2, 190) = 4.65$, $p = .015$. A post-hoc LSD analysis revealed significant differences in DSM Communication score change between the AD and PDD-NOS (p = .005) groups, but not the AD and ASD-Low MA (p = .16) and PDD-NOS and ASD-Low MA (p = .646) groups (Table 13).

*DSM-IV: Restricted and Repetitive Behaviors*

A DSM Restricted and Repetitive Behaviors score was available at both Time 1 and Time 2 for 95 participants in the AD group, 74 in the PDD-NOS group, and 23 in the ASD Low-MA group.

At Time 1, the mean DSM Restricted and Repetitive Behaviors score was 1.85 (SD = .82) for the AD group, .7 (SD = .74) for the PDD-NOS group, and 1.74 (SD = 1.10) for the ASD Low-MA group. At Time 2, the mean DSM Restricted and Repetitive

Behaviors Score was 1.55 (SD = .96) for the AD group, 1.43 (SD = .98) for the PDD-NOS group, and 1.96 (SD = .88) for the ASD Low-MA group. A mixed model design analysis demonstrated a significant interaction between Time 1 Diagnosis and time-point (Time 1 and Time 2), $F(2, 188) = 18.37$, $p < .0005$.

Due to this significant interaction, a 1-way ANOVA was performed to compare the AD, PDD-NOS, and ASD Low-MA groups using the change in DSM Restricted and Repetitive Behaviors score between Time 1 and Time 2. The mean change in DSM Restricted and Repetitive Behaviors score was -.33 (SD = 1.10) for the AD group, .73 (SD =1.13) for the PDD-NOS group, and .22 (SD = 1.04) for the ASD Low-MA group, and was significantly different among groups, $F(2, 190) = 19.42$, $p < .0005$. A post-hoc LSD analysis revealed significant differences in DSM Restricted and Repetitive Behaviors score change between the AD and ASD-Low MA ($p = .033$), and AD and PDD-NOS ($p < .0005$), and the PDD-NOS and ASD-Low MA ($p = .054$) groups (Table 14).

## Discussion

The goal of the current study was to examine the diagnostic stability of Autistic Disorder, Pervasive Developmental Disorder, Not Otherwise Specified, and Autism Spectrum Disorder, Low Mental Age, and explore the change in developmental level and autism symptomatology within each group across time. The children included in this study participated in an ongoing study conducted at the University of Connecticut, which is assessing the use of the M-CHAT, and its revised version (M-CHAT-R), as an ASD screening instrument for children between the ages of 16 to 30 months.

### *Summary of Results*

*Diagnostic Stability*

The majority of the ASD Low-MA group retained an ASD diagnosis at follow-up, with five (20%) participants retaining that diagnosis and 19 (76%) receiving an AD diagnosis. Only one child (4%) received a non-spectrum diagnosis of Developmental Delay, which indicates that despite no longer demonstrating symptoms of autism, this participant continued to exhibit significant delays in multiple domains of functioning. When including all possible diagnostic outcomes at follow-up, we found a significant difference in the diagnostic stability of each group (AD, PDD-NOS, and ASD Low-MA). Specifically, the AD group was significantly more likely (86.5%) to receive an ASD diagnosis at Time 2 than the PDD-NOS group (73.2%). The ASD Low-MA group (96%) had a higher rate of ASD retention than the AD group, but the two groups were not significantly different. The PDD-NOS group was significantly less likely than the ASD Low-MA group to receive an ASD diagnosis at follow-up. Additionally, the AD (67.6%) group was significantly more likely than the PDD-NOS (39%) and ASD Low-MA (20%) groups to retain their initial diagnosis versus receiving any other ASD diagnosis or any other diagnostic outcome. Finally, the ASD Low-MA group (76%) was significantly more likely than the AD (18.9%) and PDD-NOS (34.1%) groups to receive an ASD diagnosis different from their initial diagnosis, at follow-up.

Contrary to our hypothesis, all but one of the 25 participants in the ASD Low-MA group received an ASD diagnosis at follow-up. This suggests that the initial symptoms of autism observed were stable across time and in fact indicative of an ASD at initial evaluation, and not solely a consequence of the absence of typical behavior due to low mental age. A majority of these children also receive a diagnosis of AD (n =19, 76%) at

follow-up as opposed to ASD Low MA (n = 5, 20%), which indicates that these children made enough developmental progress between evaluations to exceed the "below 12 month age equivalent" criteria to be considered low mental age. As predicted, the AD group (86.5%) had a higher likelihood of receiving an ASD diagnosis at follow-up than the PDD-NOS group (73.2%), but the rate of ASD retention was high for all groups, ranging from 73.2% to 96%. 67.6% of participants in the AD group retained that diagnosis, which was significantly greater than the PDD-NOS (39%) and ASD Low-MA (20%) groups, and consistent with our predictions.

These findings demonstrate a high stability of ASD diagnoses across time for all groups. The high likelihood of receiving an ASD diagnosis at follow-up within the ASD Low-MA group suggests that despite the presence of a low mental age, symptoms of an ASD can be accurately identified and are indicative of the presence of a concurrent ASD. Additionally, within the ASD Low-MA group, ASD symptomatology remains at follow-up, even when mental age progress has exceeded the low mental age threshold as defined by this study. The higher rate of diagnostic stability in the AD group supports the conceptualization of AD as an ASD involving a higher degree of impairment than PDD-NOS. These results also suggest that the degree of ASD symptomatology is likely higher in the ASD Low-MA group at both time points than in the PDD-NOS group, and that the degree of ASD-related impairment may be more significant and resulting in this higher ASD retention.

*Developmental Progress: Mullen*

The ASD Low-MA group had the lowest mean ratio IQ scores in all four subdomains of the Mullen (Receptive Language, Expressive Language, Fine Motor, and

Visual Reception), at both time points. The AD and PDD-NOS groups had comparable ratio IQ equivalent scores, but the PDD-NOS group had higher mean ratio IQ scores than the AD group in all four subdomains of the Mullen, at both time points. Those children with ASD Low-MA made significantly less developmental progress in all domains than children in the AD and PDD-NOS groups. Across all subdomains, the PDD-NOS group had higher developmental progress than the AD group, but these groups did not differ significantly in any subdomain. Descriptively, all groups made less than 2 years of developmental gains in across all developmental domains during the 2 years between evaluations, though the ASD Low-MA group made the lowest gains, with the lowest developmental progress ratio in Fine Motor (.27), and highest in Visual Reception (.48). These findings supported our hypothesis, in which we had predicted that children with ASD Low-MA would make significantly lower gains in all areas of development compared to children with AD and PDD-NOS. These results suggest that participants who received a diagnosis of ASD Low-MA demonstrated the most significant level of developmental impairment at both time points, and made the least amount of progress between these time points. While the differences between the AD and PDD-NOS groups were not significant, the PDD-NOS group made the greatest amount of gains between time points, even approaching a ratio of 1 in several sub-domains (.99 in Visual Reception, .99 in Receptive Language). These ratios imply that these participants on average made the same amount of growth in mental age, measured in age equivalent change between Time 1 and Time 2, as they did in chronological age. The area in which each group made the least developmental progress was Fine Motor (AD = .69, PDD-NOS = .74, ASD Low-MA = .27), which is compelling to consider alongside prior literature

that demonstrated that children with more severely delayed motor skills are less likely to move off the spectrum (Sutera et al., 2007).

*Developmental Progress: Vineland*

The results observed on the Vineland and its various subdomains were largely consistent with our predictions, as well as the outcomes observed on the Mullen, though there were some surprising findings. On the whole, the PDD-NOS group typically received the highest ratio IQ scores at both Time 1 and Time 2, followed by the AD group, with the ASD Low-MA group receiving the lowest mean ratio IQ scores. The first exception to this was the Personal subdomain, for which all three groups had similar ratio IQ's at Time 1, and each group experienced a decrease, rather than increase, in ratio IQ across time. Similarly, in the Domestic subdomain, the AD group had the highest ratio IQ score at Time 2, and all three groups experienced a decrease in ratio IQ between time points. Likewise, in the Community subdomain, the AD group had the highest ratio IQ score at Time 2, having experienced an increase in ratio IQ score. The final unexpected result was in the Coping domain, where the AD group had a higher ratio IQ at time 1 than the PDD-NOS group, but more surprisingly, the ASD Low-MA group also had a marginally higher coping rate than the PDD-NOS group. Though these group differences are quite small, it was the only observed instance on either developmental measure in which the ASD Low-MA group had a score that indicated a higher level of functioning than either the AD or PDD-NOS group. The ASD Low-MA group had the lowest ratio of developmental progress between evaluations, across all domains of the Vineland. This supports the conceptualization of ASD Low-MA as a diagnosis involving severe, highly stable developmental delays in several areas of functioning. In some cases, the AD group

had a slightly higher progress ratio than the PDD-NOS group, which was contrary to our expectations. Past studies have demonstrated consistently that the PDD-NOS group tests higher than the AD group in cognitive skills (Cohen et al., 1986; Sevin et al., 1995; Walker et al., 2004), but one study that combined AD and PDD-NOS into groups based on high versus low IQ, found an interaction between IQ and cognitive scores on adaptive functioning (Bolte & Poustka, 2002). These results may imply that while most often there are observed differences in cognitive and adaptive ability between PDD-NOS and AD, that this is not a blanket standard to be applied to every adaptive domain.

*ASD Symptomatology*

The AD (35.33) and ASD Low-MA (35.12) groups had similar mean CARS Total Score at Time 1, with the PDD-NOS (28.2) group having a significantly lower mean score. At Time 2, however, the AD group (30.95) demonstrated a notable decrease in CARS mean score and the PDD-NOS (26.76) experienced a more modest decrease, while the ASD Low-MA group experienced a modest increase (36.33). These findings are consistent with the findings of a study conducted by Chlebowski et al. (2010), which identified a CARS Total Score of 32 as the cut-off score to distinguish between AD and PDD-NOS at Time 1, and a score of 30 at Time 2. Overall, the AD group experienced a significantly higher reduction in ASD symptomatology, as compared to the PDD-NOS and ASD Low-MA groups. This reduction in ASD symptomatology for the AD group demonstrates improvement across time, and this reduction in ASD symptoms could potentially be attributed to identification and subsequent implementation and reception of intervention services at Time 1. The small decrease in PDD-NOS CARS scores was less expected, as was the small increase in ASD Low-MA CARS scores. These changes

suggest that the ASD symptomatology in both the PDD-NOS and ASD Low-MA groups are more stable across time than the AD group.

On the ADOS, the observed trends are similar to those observed regarding the CARS scores and change across time. At follow-up, the AD group experienced the most significant change across time, the PDD-NOS group was less severe than the other groups at both time points but experiencing less change, and the ASD Low-MA group was the most severe at both time points with little change. For both the CARS and the ADOS, it may be that the children with PDD-NOS had a milder ASD symptomatology at the time of their initial evaluation, but began expressing a higher number and/or intensity of ASD symptoms, particularly the RRBs, by the time of follow-up.

The change in DSM-IV Total Score was significantly different between the AD (-1.37) and the PDD-NOS (.76) and ASD Low-MA (.52) groups. These trends are similar to those observed on the ADOS and the CARS, with the exception of the PDD-NOS experiencing a minor increase in ASD symptomatology on the DSM-IV Total. This could be attributed to the nature of the DSM-IV-TR rating sheet itself, in that it does not allow a reflection of severity as is built into the CARS and ADOS, but rather an indication of whether or not a particular symptom is present or absent.

The change in DSM-IV Social Score was significantly different between the (-.94) AD and the PDD-NOS (-.39) and ASD Low-MA (.00) groups. These trends are consistent with the previously observed trends on the CARS and ADOS, but distinct from the overall trend for the DSM-IV Total Score, in which the PDD-NOS group experienced a slight increase in symptomatology. These results demonstrate that the symptoms of ASD that affect social functioning are not increasing in the PDD-NOS group, and that

this overall increase in ASD symptom severity must be attributed to either the Communication or RRB domains.

The change in DSM-IV Communication Score was significantly different between the AD (-.04) group and the PDD-NOS (.42) group. In Communication, the PDD-NOS and ASD Low-MA groups actually experienced a relative increase in symptoms endorsed, while the AD group was stable across time. These findings are consistent with the DSM-IV Total Score, as well as with the assumption that if children with PDD-NOS are making progress with regard to Social symptoms of ASDs, that they must be experiencing a worsening of symptoms in another domain, in this case Communication.

The change in DSM-IV Restricted and Repetitive Behaviors Score was significantly different between all groups, AD (-.33), PDD-NOS (.73) and ASD Low-MA (1.04). These trends are similar to those observed on DSM-IV Communication, where PDD-NOS and ASD Low-MA are experiencing an increase in symptoms endorsed across time, while the AD group experiences a relative decrease. This is indicative that for the PDD-NOS group, despite improvements in the Social domain, these children are experiencing increased difficulties in the Communication and RRB domains.

These unexpected changes in the PDD-NOS group, increases in ASD symptomatology, may be attributed to those children in this group that receive an AD diagnosis at Time 2, as their symptomatology would have increased between time points. Particularly, at Time 1, these children may not have been demonstrating as clear deficits in Communication, especially if their language were delayed. Also, these children did not have to be exhibiting any repetitive or restricted behaviors, as it is not a required criterion

for PDD-NOS, and these behaviors may have emerged between time points and resulted in a diagnosis of AD at follow-up.

The overall trends observed across the CARS, ADOS, and DSM-IV were somewhat contrary to our predictions. The ASD Low-MA group demonstrated a high degree of stability across time on each of these measures, and experienced an increase in symptomatology, though small, on all of the aforementioned measures and their subdomains, with the exception of the DSM-IV Social subdomain, where they experienced no change. The PDD-NOS group consistently demonstrated the least severity at all time-points, across all measures, though increases in symptom severity on the DSM-IV Total, Social, and RRB subdomains was unexpected. These children may have demonstrated mild symptomatology at their initial evaluation, when in reality they were not fully expressing symptoms of ASDs that emerged sometime before their follow-up. Additionally, these children may have received less intensive treatment as a result of this milder presentation than the AD group. The AD group experienced the most significant decrease in ASD symptoms across all measures, which is surprising given the high number of participants (67.6%) that retained the AD diagnosis at follow-up. However, a number of these participants in the AD group received a PDD-NOS diagnosis (18.9%) or moved off the spectrum (13.5%) entirely, which may explain the significant decrease in ASD symptomatology.

*Implications*

The results of this study demonstrate the difference between the diagnostic stability of AD, PDD-NOS, and ASD Low-MA diagnoses between two time points (approximately 2 years). A majority of children with ASD Low-MA remain on the autism

spectrum at follow-up (96%), with some that continue to demonstrate a low mental age (20%), and only one child that no longer demonstrated symptoms of an ASD, but was still significantly developmentally delayed (4%). This finding suggests that autism can in fact be detected accurately, even in the presence of co-occurring low mental age, and that the symptoms of autism are highly stable across time in this group. It also implies that children that test with a low mental age at an early time point will likely make cognitive gains, such that they are no longer meeting this category at a later time point. A high proportion of children in the AD (86.5%) and PDD-NOS (73.5%) groups also received an ASD diagnosis at follow-up, highlighting the high rate of diagnostic retention between time-points across all ASD diagnoses, ASD Low-MA included. Children with AD had a particularly high rate of receiving that same diagnosis at follow-up (67.6%), while the PDD-NOS group had a comparable likelihood of receiving PDD-NOS (39%), another ASD (34.1%), or a non-spectrum diagnosis (26.8%). The rates of ASD maintenance within the PDD-NOS group seems to suggest that it is a conceptually distinct diagnostic group from AD, and the even split between outcomes mirrors prior research in which the PDD-NOS group can be divided into subgroups based on their symptomatology, developmental level, and outcome. The significant portion of children from the PDD-NOS group that go on to receive AD at follow-up (equal to or less than 34.1%) is similar to the rate of children that would not receive a diagnosis under DSM-V criteria. This could suggest that perhaps these children are not demonstrating the full array of ASD symptoms at initial evaluation, but would still benefit heavily from identification and reception of early intervention services. The elimination of the PDD-NOS diagnosis, and the introduction of the requirement of Restricted and Repetitive Behaviors could result in

a portion of these children that go on to meet criteria for AD at follow-up being missed and failing to receive intervention.

The results of the Mullen are largely consistent with previous research that describes children with PDD-NOS as higher cognitively than their AD counterparts. By definition, ASD Low-MA should have the lowest measured improvements in these areas, which they did in this sample. An unexpected result, however, was the observed decrease in Fine Motor Ratio IQ's at follow-up for all three groups. The developmental progress ratio was positive for each group, which suggests that despite having a low ratio IQ because of the continuing differences between mental age and chronological age, that these individuals on a whole are making some degree of progress between evaluations. However, the developmental progress ratio is less than 1.00 for each group, implying that all participants are making less progress in mental age than what would be expected by the change in actual chronological age between evaluations. This fact would potentially explain the decrease in ratio IQ observed at Time 2, since relative to their peers, these individuals are not making mental progress equal to the amount of chronological time that passes. The developmental progress ratios were relatively high for the AD group, with the lowest ratio .69 for Fine Motor, and the highest .99 for Receptive Language. This score in Receptive Language suggests that these children able to make the same years' worth of progress in mental age as their change in chronological age, which is quite promising for children with AD, a disorder that impacts communication and language. Similarly, the PDD-NOS group had a rate of .99 for both the Visual Reception and Receptive Language subdomains, with their lowest ratio .74 for Fine Motor. In contrast, the ASD Low-MA group had ratios as low as .27 for Fine Motor, .35 for

Expressive and .34 for Receptive Language, and .48 for Visual Reception. This suggests just how limited the developmental progress is for children in this diagnostic category, as these children on average are making only a quarter to one half of the progress expected of them between evaluations.

The results of the Vineland were largely consistent with trends observed on the Mullen, with some unexpected exceptions. The ASD Low-MA group is the lowest functioning at both time points for all but one domain at a single time point, which was Coping at Time 1, in which they had a higher ratio IQ score than the PDD-NOS group. Otherwise, the PDD-NOS group was typically the highest functioning group, with the AD group having higher ratio IQ's than the PDD-NOS group at Time 1 on Coping, and on Time 2 for Domestic and Community. These findings are contrary to prior research (Walker et al., 2004) in which PDD-NOS has a higher level of functioning than AD in every domain without exception. As expected, the AD group also had higher progress ratios than the PDD-NOS group in Domestic and Community, but the PDD-NOS group had higher progress ratios than AD in every other subdomain. Overall, the ASD Low-MA had the lowest progress ratios in every subdomain, which ranged from .02 in Domestic at the lowest to .41 in Receptive Language at the highest. This is comparatively low when considering the higher progress ratios for both the AD, ranging from .46 in Play and Leisure to .82 in Community, and PDD-NOS, ranging from .47 in Play and Leisure to .91 in Receptive Language, groups. The developmental progress ratios for Language were high for the AD (.80 Expressive and .77 Receptive) and PDD-NOS groups (.76 Expressive and .91 Receptive), which is higher than might be expected due to the difficulties with communication and language often present in ASDs. These trends may

be in part influenced by the fact that the Vineland is a parent-based measure, and based on subjective ratings of behaviors that the child demonstrates in an environment outside of the evaluation. The Vineland itself also increases in the level of sophistication required between items as the items progress within each subdomain; as a result it may take less development and be more forgiving at the lower end of functioning, and become more difficult to receive a higher score in the subdomains as the questions progress. These domains may also involve skills that are likely to not improve for children with ASDs easily, such as those related to Daily Living Skills and Socialization. These could also include skills that are not targeted as part of intervention services at that age, which are primarily focused on rehabilitating and teaching more rudimentary behaviors and tasks, such as following verbal commands and engaging with others socially, and not answering the telephone or observing rules while riding in the car. Other domains also have items that are likely to be at a deficit for an individual with an ASD, regardless of developmental or diagnostic progress between evaluations. For example, the Domestic subdomain included items concerning personal safety, like being careful around hot or sharp objects, which are often at a deficit in individuals with ASDs, who may have an undersensitivity to pain and poorer appreciation of danger in the immediate environment.

A few major trends regarding autism symptomatology emerged across the CARS, ADOS, and DSM-IV. Firstly, the AD group experienced a decrease in total CARS score, ADOS CSS, and DSM-IV Total score between two time points, a decrease that was significantly greater than changes in the PDD-NOS and ASD Low-MA groups. The PDD-NOS group experienced a smaller decrease on the CARS and ADOS, and an increase in symptom severity on the DSM-IV Total, specifically in the Communication

and RRB subdomains. The ASD Low-MA group, in contrast, experienced small increases on the CARS and DSM-IV Total, with a minor decrease on the ADOS. This presented an interesting trend of improvement in the AD group, compared to relative stability, and occasional worsening, of symptom profile among the PDD-NOS and ASD Low-MA groups. This high degree of symptom stability provides further support to the legitimacy of assigning an ASD diagnosis in the presence of low mental age, suggesting that the observed behaviors are in fact indicative of an ASD, and not just a product of a low mental age or an absence of developmental milestones. The trend observed earlier, that the PDD-NOS group is likely to receive another ASD diagnosis at follow-up, which is most likely AD, would help explain an increase in symptom severity on the DSM-IV Total. Additionally, at Time 2 the PDD-NOS group shows an increase in RRB's, which arguably present later in development, and may not have been present at Time 1 (since a PDD-NOS diagnosis did not require RRB's). These very individuals that began to exhibit increased RRB's at follow-up may constitute a significant number of those individuals from the PDD-NOS group remained on the ASD spectrum, but did not receive a PDD-NOS diagnosis at follow-up. The decrease in symptoms in the AD group could suggest that the observed effects are a response to being identified and receiving intervention services, which in Connecticut consists typically of 15-20 hours per week of intensive intervention services for a child with an ASD.

*Limitations*

There were a number of limitations that must be considered when interpreting the findings of this study. While the overall sample size (N=218) is considered large compared to many studies, the ASD Low-MA subgroup was significantly smaller (n =

25) than both the AD (n = 111) and PDD-NOS (n = 82) groups. This caused the power for some analyses to be smaller than anticipated. If each group had a more comparable sample size, we would be more confident in the findings that indicate no differences between groups; at current, we may not have had adequate sample sizes to detect differences for every question we were investigating.

There was also an unfortunate lack of available data for every participant at both initial and follow-up, resulting in missing data in some cases. This likely could have reduced the likelihood of detecting an effect, particularly if this data was missing for the ASD Low-MA group, due to its small sample size. This is, however, a function of the rarity of the ASD Low-MA subtype to begin with, as over the many years this study has operated we were only able to collect data at two time points on these 25 participants. Overall, these analyses could be revisited in future studies with larger sample sizes.

Additionally, across all three groups a portion of children who received a Time 1 evaluation declined a Time 2 evaluation, or the study was unable to contact for a follow-up. It is quite possible that the parents of these children were not as concerned with their child's development at the time a follow-up evaluation was offered, when compared to those that did return for follow-up. If true, this trend may have biased our data, as our sample would be comprised of children who were more likely to demonstrate delays or symptoms of autism at a follow-up evaluation.

Another limitation to consider is the time frame of the study. The data included is from two time points as part of the study design, at approximately two and four years of age. There is no data available past the follow-up evaluation. As a result, we do not have potentially informative data regarding the long-term outcomes for these children.

Previous research has demonstrated that ASD symptoms, cognitive ability, and adaptive skill level changes with age in children with ASDs, which may have occurred further in this sample.

An additional limitation in this study was the inability to account for the likely impact and role of intervention on the diagnostic and developmental outcome for these participants. Through participation in this study, each participant in this sample received a diagnosis of an ASD at a relatively young age, around two years on average. This diagnosis is normally accompanied by a diagnostic report with lengthy recommendations for services targeting these symptoms of autism, as well as any developmental delays present at time of evaluation. Intervention services are known to make a significantly positive impact on autism symptomatology and severity of developmental delay, and it is reasonable to think that within this sample, the results at Time 2 reflect some of this progress. Within the study, families indicated on history forms at follow-up that each participant received some type of intervention, but specific information regarding the type of services, as well the intensity and frequency of services, was not always available or detailed enough to truly evaluate the impact and quality of services on a child's progress over time. Therefore, we were unable to directly assess if a child's improvement across time was directly related to the services they received, or the increased parental understanding of their child's deficits, or any other factor or combination of factors that may have contributed to a child's progress, or lack thereof, between evaluations.

The reliance upon a single measure of development (Mullen) and one measure of adaptive skills (Vineland) is also a limitation of the study. These evaluations are conducted within a relatively brief, 3-hour time frame. This allows for the assessment of

a child's skill-level and assignment of a diagnosis, but does not necessarily allow for the most comprehensive skill-level measurement. The Mullen as a diagnostic tool is not the most inclusive and thorough measure of development; there are a number of alternative assessments that require higher administration time, but produce a more precise assessment of developmental (e.g., The Differential Ability Scales, The Bayley Scales of Infant and Toddler Development, The Stanford-Binet Intelligence Skills). However, the exceptionally low developmental level of many of these children precludes the use of these measures, especially at initial evaluation.

*Future Directions*

These results support the position that AD, PDD-NOS, and now ASD Low-MA are distinct, sub-categories of ASDs. Each group demonstrated a distinct diagnostic stability, developmental progress across time, and change in ASD symptomatology. Even within these groups, particularly PDD-NOS, there seemed to be further divisions and sub-categories based on performance at initial evaluation and ultimate outcome at follow-up. The results for the PDD-NOS group were particularly compelling, as this group did not uniformly have the least impairment in adaptive skill levels and autism symptom severity at both time points as expected based on previous research. A significant portion of the group received an ASD other than PDD-NOS at follow-up, most likely AD, indicating an emergence of new, or worsening of already present, ASD symptoms. This provides further, compelling evidence, for the strength of PDD-NOS as a diagnostic category, and the potential loss of individuals under the DSM-V criteria who at initial evaluation may not be demonstrating Repetitive and Restricted Behaviors because they have yet to emerge.

The developmental progress rates and ASD symptom improvement in the AD and PDD-NOS groups suggests that despite their severe delays and symptomatology at initial evaluation, the intensive intervention services that these groups are likely receiving after their initial evaluation is helping to address these delays and providing an opportunity to make improvements. However, the very low rate of developmental progress, and the worsening of ASD symptomatology in the ASD Low-MA group signals the severity of this condition, and may merit more intensive services targeting the delays in this population. It may also be a function of the role a low mental age plays in being able to engage with intervention services and make developmental progress at all.

There is a clear need for replication of these results, particularly those regarding the ASD Low-MA sample, to ensure the reliability of diagnosing an ASD in the presence of a low mental age. If these results were to be replicated, it would be of great service to the field, and allow clinicians the confidence in diagnosing ASDs, regardless of mental age at time of evaluation.

## References


Allen, D.A., Steinberg M., Dunn M., et al. (2001), Autistic disorder versus other pervasive developmental disorders in young children: same or different? *Eur Child Adolesc Psychiatry* 10:67–78

American Psychiatric Association (Ed.). (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. American Psychiatric Pub.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-V* Washington, DC: American Psychiatric Association.

Baron-Cohen, S., Wheelwright, S., Cox, A., Baird, G., Charman, T., Swettenham, J., Drew, A., & Doehring, P. (2000). Early identification of autism by the CHecklist for Autism in Toddlers (CHAT). *Journal of the Royal Society of Medicine*, *93*(10), 521.

Baron-Cohen, S., Cox, A., Baird, G., Swettenham, J., Drew, A., Nightingale, N., Morgan, K., & Charman, T. (1996). Psychological markers of autism at 18 months of age in a large population. *British Journal of Psychiatry*, *168*(2), 158-163.

Barton, M. L., Robins, D. L., Jashar, D., Brennan, L., & Fein, D. (2013). Sensitivity and specificity of proposed DSM-5 criteria for autism spectrum disorder in toddlers. *Journal of Autism and Developmental disorders*, *43*(5), 1184-1195.

Berry, L. N. (2009). Early treatments associated with optimal outcome in children with autism spectrum disorders. (Unpublished doctoral dissertation) University of Connecticut, Storrs, Connecticut.

Bölte, S., & Poustka, F. (2002). The relation between general cognitive level and adaptive behavior domains in individuals with autism with and without co-morbid mental retardation. *Child Psychiatry and Human Development*, *33*(2), 165-172.

Brennan, L., Barton, M., Chen, C. M., Green, J., & Fein, D. (2014). Detecting Subgroups in Children Diagnosed with Pervasive Developmental Disorder-Not Otherwise Specified. *Journal of autism and developmental disorders*, 1-16.

Buitelaar, J. K., Van der Gaag, R., Klin, A., & Volkmar, F. (1999). Exploring the boundaries of pervasive developmental disorder not otherwise specified: Analyses of data from the DSM-IV autistic disorder field trial. Journal of Autism and Developmental Disorders, 29(1), 33–43.

Centers for Disease Control and Prevention (2012). Autism and Developmental Disabilities Monitoring Network Surveillance Year 2008 Principal Investigators; Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders - Autism and Developmental Disabilities Monitoring Network, 14 Sites,

United States, 2008. Morbidity and Mortality Weekly Report Surveillance Summaries. **61**(3), 1 –19.

Centers for Disease Control and Prevention (2014). Autism and Developmental Disabilities Monitoring Network Surveillance Year 2010 Principal Investigators; Centers for Disease Control and Prevention. Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2010. Morbidity and Mortality Weekly Report Surveillance Summaries. **63**(2), 1-21.

Chakrabarti, S., & Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *Jama*, *285*(24), 3093-3099.

Charman, T., Pickles, A., Simonoff, E., Chandler, S., Loucas, T., & Baird, G. (2011). IQ in children with autism spectrum disorders: data from the Special Needs and Autism Project (SNAP). *Psychological medicine*, *41*(03), 619-627.

Charman, T., Taylor, E., Drew, A., Cockerill, H., Brown, J. A., & 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.

Chasson, G. S., Harris, G. E., & Neely, W. J. (2007). Cost comparison of early intensive behavioral intervention and special education for children with autism. Journal of Child and Family Studies, 16(3), 401-413.

Chlebowski, C., Green, J. A., Barton, M. L., & Fein, D. (2010). Using the childhood autism rating scale to diagnose autism spectrum disorders. *Journal of autism and developmental disorders*, *40*(7), 787-799.

Cohen DJ, Paul R, Volkmar FR (1986), Issues in the classification of pervasive and other developmental disorders: toward *DSM-IV. J Am Acad Child Psychiatry* 25:213–220

Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., Donaldson, A., Varley, J. (2010). Randomized, controlled trail of an intervention for toddlers with autism: the early start Denver model. *Pediatrics, 125* (1), 17-23.

Dietz, C., Swinkels, S., van Daalen, E., van Engeland, H., & Buitelaar, J. K. (2006). Screening for autistic spectrum disorder in children aged 14–15 months. II: population screening with the Early Screening of Autistic Traits Questionnaire (ESAT). Design and general findings. *Journal of autism and developmental disorders*, *36*(6), 713-722.

Eaves, L. C., & Ho, H. H. (2004). The very early identification of autism: Outcome to age 41/2–5. *Journal of autism and developmental disorders*, *34*(4), 367-378.

Fein, D. (Ed.). (2011). *The neuropsychology of autism*. Oxford University Press.

Filipek, P. A., Accardo, P. J., Ashwal, S., Baranek, G. T., Cook, E. H., Dawson, G., Gordon, B., Gravel, J.S., Johnson, C.P., Kallen, R.L., Levy, S.E., Minshew, N.J., Ozonoff, S., Prizant, B.M., Rapin, I., Rogers, S.J., Stone, W.L., Teplin, S.W., Tuchman, R.F., & Volkmar, F. R. (2000). Practice parameter: Screening and diagnosis of autism Report of the Quality Standards Subcommittee of the American Academy of Neurology and the Child Neurology Society. *Neurology*,*55*(4), 468-479.

Fombonne, Eric. "The prevalence of autism." *Jama* 289.1 (2003): 87-89.

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

Green, G., Brennan, L.C., Fein, D. (2002). Intensive behavioral treatment for a toddler at high risk for autism. *Behavioral Modification, 26*(1), 69-102.

Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: A four-to six-year follow-up. *Journal of autism and developmental disorders*, *30*(2), 137-142.

Helt, M., Kelley, E., Kinsbourne, M., Pandey, J., Boorstein, H., Herbert, M., et al. (2008). Can children with autism recover? If So, How? Neuropsychology Review, 18(4), 339–366.

Hinnebusch A., Carr K., Barton M., Fein D. Diagnostic stability of autism spectrum disorders among children and siblings of children with an autism spectrum disorder diagnosis. Poster session presented at: IMFAR Annual conference of INSAR; 2012; Toronto, Canada.

Howard, J. S., Sparkman, C. R., Cohen, H. G., Green, G., & Stanislaw, H. (2005). A comparison of intensive behavior analytic and eclectic treatments for young children with autism. *Research in developmental disabilities*, *26*(4), 359-383.

Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry*, *45*(2), 212-229.

Hurley, A.D., Levitas, A.S. (2007). The importance of recognizing autism spectrum disorders in intellectual disability. *Mental Health Aspects of Developmental Disabilities, 10,* 157-161.

Inada, N., Kamio, Y., & Koyama, T. (2010). Developmental chronology of preverbal social behaviors in infancy using the M-CHAT: Baseline for early detection of atypical social development. *Research in Autism Spectrum Disorders*, *4*(4), 605-611.

Jashar, Dasal T., "DSM-5 Autism Criteria Applied to Toddlers with DSM-IV-TR Autism" (2014).*Master's Theses.* Paper 670.

Kanne, S. M., Gerber, A. J., Quirmbach, L. M., Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2011). The role of adaptive behavior in autism spectrum disorders: Implications for functional outcome. *Journal of autism and developmental disorders*, *41*(8), 1007-1018.

Kleinman, J. M., Ventola, P. E., Pandey, J., Verbalis, A. D., Barton, M., Hodgson, S., Green, J., Dumont-Mathieu, T., Robins, D., & Fein, D. (2008). Diagnostic stability in very young children with autism spectrum disorders. *Journal of autism and developmental disorders*, *38*(4), 606-615.

Lavelle, T.A., Weinstein, M.C., Newhouse, J.P., Munir, K., Kuhlthau, K.A., Prosser, L.A. (2014). Economic burden of childhood autism spectrum disorders. *Pediatrics, 133*(3), e520-529.

Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry*, *36*(8), 1365-1382.

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

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., et al. (2000). The Autism Diagnostic Observation Schedule--Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205-223.

Lord, C., & Schopler, E. (1989a). The role of age at assessment, developmental level, and test in the stability of intelligence scores in young autistic children.*Journal of autism and developmental disorders*, *19*(4), 483-499.

Lord, C., & Schopler, E. (1989b). Stability of assessment results of autistic and non-autistic language-impaired children from preschool years to early school age. *Journal of Child Psychology and Psychiatry*, *30*(4), 575-590.

Luyster, R., Guthrie, W., Gotham, K., Risi, S., DiLavore, P., & Lord, C. (2009). The Autism Diagnostic Observation Schedule–Toddler module: Preliminary findings using a modified version of the ADOS. In *International meeting for autism research* (pp. 15-17).

Mandy, W., Charman, T., Gilmour, J., & Skuse, D. (2011). Toward specifying pervasive developmental disorder—not otherwise specified. Autism Research, 4(2), 121–131.

Mars, A. E., Mauk, J. E., & Dowrick, P. W. (1998). Symptoms of pervasive developmental disorders as observed in prediagnostic home videos of infants and toddlers. *The Journal of pediatrics*, *132*(3), 500-504.

Matson, J. L., & Shoemaker, M. (2009). Intellectual disability and its relationship to autism spectrum disorders. *Research in developmental disabilities*, *30*(6), 1107-1114.

Mayes, S. D., & Calhoun, S. L. (2003). Ability profiles in children with autism influence of age and IQ. *Autism*, *7*(1), 65-80.

McGee, G.G., Morrier, M.J., Teresa, D. (1999). An incidental teaching approach to early intervention for toddlers with autism. *Research and Practice for Persons with Severe Disabilities, 24*(3), 133-146.

Moore, V., & Goodson, S. (2003). How well does early diagnosis of autism stand the test of time? Follow-up study of children assessed for autism at age 2 and development of an early diagnostic service. *Autism*, *7*(1), 47-63.

Mullen, E. M. (1995). *Mullen Scales of Early Learning*. Circle Pines, MN: American Guidance Service Inc.

Osterling, J. A., Dawson, G., & Munson, J. A. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation.*Development and psychopathology*, *14*(02), 239-251.

Paul, R., Miles, S., Cicchetti, D., Sparrow, S., Klin, A., Volkmar, F. (2004). Adaptive behavior in autism and pervasive developmental disorder—not otherwise specified: Microanalysis of scores on the Vineland adaptive behavior scales. Journal of Autism and Developmental Disorders, 34(2), 223–228.

Rutter, M., Le Couteur, A., Lord, C., & Faggioli, R. (2005). *ADI-R: Autism diagnostic interview--revised: Manual*. OS, Organizzazioni speciali.

Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *Journal Information*, *110*(6).

Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders, 10*(1), 91-103.

Sevin JA, Matson JL, Williams D, Kirkpatrick-Sanchez S (1995), Reliability of emotional problems with the Diagnostic Assessment for the Severely Handicapped (DASH). *Br J Clin Psychol* 34:93–94

Sigman, M., & Ruskin, E. (1999). Change and continuity in the social competence of children with Autism, Down syndrome, and developmental delays. *Monographs of the Society for Research in Child Development*.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984).The Vineland Adaptive Behavior Scales-Interview Edition. Circle Pines, MN: American Guidance Service.

Stone, W. L., Lee, E. B., Ashford, L., Brissie, J., Hepburn, S. L., Coonrod, E. E., & Weiss, B. H. (1999). Can autism be diagnosed accurately in children under 3 years?. *Journal of Child Psychology and Psychiatry*, *40*(02), 219-226.

Stone, W. L., McMahon, C. R., & Henderson, L. M. (2008). Use of the Screening Tool for Autism in Two-Year-Olds (STAT) for children under 24 months An exploratory study. *Autism*, *12*(5), 557-573.

Sutera, S., Pandey, J., Esser, E. L., Rosenthal, M. A., Wilson, L. B., Barton, M., Green, J., Hodgson, S., Robins, D.L., Dumont-Mathieu, T., & Fein, D. (2007). Predictors of optimal outcome in toddlers diagnosed with autism spectrum disorders. *Journal of autism and developmental disorders*, *37*(1), 98-107.

Walker, D. R., Thompson, A., Zwaigenbaum, L., Goldberg, J., Bryson, S. E., Mahoney, W. J., Strawbridge, C.P., & Szatmari, P. (2004). Specifying PDD-NOS: a comparison of PDD-NOS, Asperger syndrome, and autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, *43*(2), 172-180.

Werner, E., Dawson, G., Osterling, J., & Dinno, N. (2000). Brief report: Recognition of autism spectrum disorder before one year of age: A retrospective study based on home videotapes. *Journal of autism and developmental disorders*, *30*(2), 157-162.

Table 1. Participant Demographics

| | AD Mean Age (months), SD, and range (n=111) | PDD-NOS Mean Age (months), SD, and range (n=83) | ASD Low-MA Mean Age (months), SD, and range (n=25) | Difference Between Groups |
|---|---|---|---|---|
| Time 1 | 27.09 (4.57)<br>17.74 – 36.66 | 25.88 (4.04)<br>18.03 – 34.43 | 23.60 (4.509)<br>15.70-31.80. | $F$ (4, 428) = 3.67, $p < .006$ |
| Time 2 | 53.27 (9.41)<br>41.38-113.48 | 51.49 (10.38)<br>38.12-106.52 | 49.37 (6.48)<br>34.36-62.13 | |
| | Frequency (%) AD | Frequency (%) PDD-NOS | Frequency (%) ASD Low-MA | |
| Participant Gender | | | | |
| Male | 93(83.8%) | 65 (78.3%) | 20 (80%) | |
| Female | 18 (16.2%) | 18 (21.7%) | 5 (20%) | |
| Total | 111 | 83 | 25 | $x^2$ (2, $N$ = 219) = .964, $p$ = .617 |
| Ethnicity | | | | |
| White/European American | 87(82.1%) | 67(84.8%) | 19(76%) | |
| Black/African American | 2(1.9%) | 4(5.1%) | 2(8%) | |
| Latino/Hispanic | 9(8.5%) | 5(6.3%) | 3(12%) | |
| Asian | 4(3.8%) | 3(3.8%) | 0 | |
| Biracial | 4(3.8%) | 0 | 1(4%) | |
| Total | 106 | 79 | 25 | $x^2$ (8, $N$ = 210) = 7.450, $p$ = .490 |
| Maternal Education | | | | |
| No degree or diploma | 3 (4.3%) | 2 (3.1%) | 0 | |

| | | | | |
|---|---|---|---|---|
| High school diploma or GED | 22 (31.4%) | 20 (30.8%) | 11 (57.9%) | |
| Vocational or technical degree | 5 (7.1%) | 5 (7.7%) | 0 | |
| Associates degree | 6 (8.6%) | 6 (9.2%) | 0 | |
| College degree | 20 (28.6%) | 19 (29.2%) | 4 (21.1%) | |
| Masters Level degree | 11(15.7%) | 12(18.5%) | 4 (21.1%) | |
| Ph.D., MD, JD level degree | 3(4.3%) | 1(1.5%) | 0 | |
| Total | 70 | 65 | 19 | $x^2$ (12, $N$ = 154) = 9.6746, $p$ = .645 |
| Yearly Income | | | | |
| <$10,000 | 6 (9.5%) | 1(1.6%) | 0 | |
| $10,000-$20,000 | 10 (15.9%) | 13(20.6%) | 2 (11.8%) | |
| $20,000-$30,000 | 4 (6.3%) | 3(4.8%) | 0 | |
| $30,000-$40,000 | 3(4.8%) | 5(7.9%) | 4 (23.5%) | |
| $40,000-$50,000 | 6 (9.5%) | 7(11.1%) | 2 (11.8%) | |
| $50,000-$60,000 | 4(6.3%) | 6(9.5%) | 2 (11.8%) | |
| $60,000-$70,000 | 6(9.5%) | 6(9.5%) | 0 | |
| $70,000-$80,000 | 7(11.1%) | 7(11.1%) | 3(17.6%) | |
| $80,000-$90,000 | 7(11.1%) | 3(4.8%) | 2 (11.8%) | |
| $90,000-$100,000 | 5(7.9%) | 2(3.2%) | 2(11.8%) | |
| >$100,000 | 5 (7.9%) | 10(15.9%) | 0 | |
| Total | 63 | 63 | 17 | $x^2$ (20, $N$ = 143) = 23.166, $p$ = .281 |

Table 2. Diagnostic Stability, AD, PDD-NOS, ASD Low-MA - All Diagnostic Outcomes

| Time 1 Diagnosis | Time 2 Diagnosis | | | | | | $x^2$ | *p* | Φ (Phi) |
|---|---|---|---|---|---|---|---|---|---|
| | AD | PDD-NOS | ASD Low-MA | DD | Other | No Dx | | | |
| AD | 75 (67.7%) | 21 (18.9%) | 0 | 2(1.8%) | 1 (0.9%) | 12 (10.8%) | 74.828 | <.0005 | .506 |
| PDD-NOS | 28 (34.1%) | 32 (39.0%) | 0 | 9(11.0%) | 3 (3.7%) | 10 (12.2%) | | | |
| ASD Low-MA | 18 (72.0%) | 1 (4.0%) | 5 (20.0%) | 1 (1.4%) | 0 | 0 | | | |

Table 3. Diagnostic Stability, AD, PDD-NOS, ASD Low-MA – Autism Spectrum Disorder Diagnosis vs. non-Autism Spectrum Disorder Diagnosis

| Time 1 Diagnosis | Time 2 Diagnosis | | | | |
|---|---|---|---|---|---|
| | ASD | Non-ASD | $x^2$ | *p* | Φ (Phi) |
| AD | 96 (86.5%) | 15 (13.5%) | 9.349 | .009 | .207 |
| PDD-NOS | 60 (73.5%) | 22 (26.8%) | | | |
| ASD Low-MA | 24 (96.0%) | 1 (4.0%) | | | |

Table 4. Diagnostic Stability AD, PDD-NOS, ASD Low-MA – Retain Diagnosis (Dx), Other Autism Spectrum Disorder, Other Diagnosis

| Time 1 Diagnosis | Time 2 Diagnosis | | | | | |
|---|---|---|---|---|---|---|
| | Retain Dx | Other ASD | Other Dx | $x^2$ | *p* | Φ (Phi) |
| AD | 75 (67.6%) | 21 (18.9%) | 15 (13.5%) | 42.29 | <.0005 | .440 |
| PDD-NOS | 32 (39.0%) | 28 (34.1%) | 22 (26.8%) | | | |
| ASD Low-MA | 5 (20.0%) | 19 (76.0%) | 1 (4.0%) | | | |

Table 5. Time 1 and Time 2 Mean Ratio IQ Scores: Mullen

| | AD | | PDD-NOS | | ASD Low-MA | |
|---|---|---|---|---|---|---|
| Average Age Equivalent | Time 1 | Time 2 | Time 1 | Time 2 | Time 1 | Time 2 |
| Visual Reception | 66.81 (19.86) | 74.26 (28.68) | 73.33 (17.09) | 85.92 (27.74) | 40.28 (12.09) | 41.72 (20.32) |
| Fine Motor | 73.81 (26.36) | 70.37 (22.82) | 76.74 (13.99) | 75.61 (20.89) | 56.73 (14.01) | 39.75 (13.05) |
| Receptive Language | 43.78 (23.25) | 63.10 (28.62) | 54.55 (22.15) | 77.01 (29.05) | 30.00 (16.14) | 32.05 (17.93) |
| Expressive Language | 48.11 (23.09) | 60.63 (26.67) | 58.34 (18.64) | 71.87 (24.72) | 33.55 (12.98) | 33.49 (20.36) |

Table 6. Developmental Progress between Evaluations, AD, PDD-NOS, ASD Low-MA - Mullen

| | Mental Growth Mean (SD) | Mental Growth Mean (SD) | Mental Growth Mean (SD) | | |
|---|---|---|---|---|---|
| | AD | PDD-NOS | ASD Low-MA | *F* | *p* |
| Visual Reception | .86 (.59) | .99 (.51) | .48 (.43) | 6.693 | .002 |
| Fine Motor | .69 (.43) | .74 (.41) | .27 (.24) | 10.732 | < .0005 |
| Expressive Language | .77 (.52) | .85 (.47) | .35 (.34) | 8.264 | < .0005 |
| Receptive Language | .99 (.60) | .99 (.51) | .34 (.34) | 4.31 | .015 |

Table 7. Time 1 and Time 2 Mean Ratio IQ Scores: Vineland

| | AD | | PDD-NOS | | ASD Low-MA | |
|---|---|---|---|---|---|---|
| | Time 1 | Time 2 | Time 1 | Time 2 | Time 1 | Time 2 |
| Expressive Language | 38.33 (18.22) | 57.67 (60.49) | 49.37 (16.72) | 62.5 (26.52) | 32.27 (12.07) | 28.47 (13.39) |
| Receptive Language | 44.27 (22.13) | 60.39 (31.53) | 56.73 (23.09) | 72.39 (33.92) | 33.11 (13.30) | 36.50 (24.735) |
| Personal | 59.54 (13.44) | 54.63 (17.62) | 59.77 (17.65) | 57.09 (17.49) | 56.11 (18.49) | 40.07 (8.53) |
| Domestic | 66.60 (22.17) | 63.50 (32.17) | 70.17 (26.40) | 58.68 (27.69) | 65.46 (19.04) | 33.47 (16.65) |
| Community | 39.91 (25.89) | 58.20 (72.08) | 40.72 (27.85) | 52.72 (24.40) | 22.11 (23.57) | 28.20 (24.67) |
| Interpersonal Relationships | 34.64 (14.14) | 44.04 (25.95) | 44.66 (17.00) | 54.12 (27.79) | 33.33 (10.53) | 22.54 (9.86) |
| Play and Leisure | 36.13 (16.79) | 40.51 (25.72) | 50.45 (20.81) | 48.57 (24.12) | 34.22 (21.19) | 22.10 (13.27) |
| Coping | 42.99 (23.42) | 45.78 (26.94) | 40.34 (18.33) | 59.07 (30.05) | 41.47 (16.33) | 22.38 (13.62) |

Table 8. Developmental Progress between Evaluations, AD, PDD-NOS, ASD Low-MA - Vineland

| | Mental Growth Mean (SD) | Mental Growth Mean (SD) | Mental Growth Mean (SD) | | |
|---|---|---|---|---|---|
| | AD | PDD-NOS | ASD Low-MA | *F* | *p* |

| | | | | | |
|---|---|---|---|---|---|
| Expressive Language | .80 (1.32) | .76 (.48) | .24 (.98) | 2.628 | .075 |
| Receptive Language | .77 (.62) | .91 (.7) | .41 (.47) | 4.777 | .01 |
| Personal | .50 (.35) | .56 (.32) | .23 (.26) | 7.044 | .001 |
| Domestic | .63 (.66) | .47 (.53) | .02 (.23) | 5.764 | .004 |
| Community | .82 (1.61) | .67 (.58) | .37 (.53) | 1.057 | .475 |
| Interpersonal Relationships | .54 (.55) | .66 (.58) | .13 (.15) | 6.942 | .001 |
| Play and Leisure | .46 (.57) | .47 (.51) | .14 (.25) | 3.062 | .049 |
| Coping | .50 (.64) | .82 (.63) | .06 (.27) | 8.493 | <.0005 |

Table 9. ASD Symptoms/Severity between Evaluations, AD, PDD-NOS, ASD Low-MA – CARS

| | AD | PDD-NOS | ASD Low-MA | *F* | *p* |
|---|---|---|---|---|---|
| CARS T1 Total | 35.33 (4.4) | 28.2 (3.59) | 35.12 (5.69) | 9.12 | <.0005 |
| CARS T2 Total | 30.95 (7.08) | 26.76 (5.79) | 36.33 (4.89) | | |
| Change in CARS | -4.37 (6.59) | -1.45 (5.79) | 1.21 (6.84) | | |

Table 10. ASD Symptoms/Severity between Evaluations, AD, PDD-NOS, ASD Low-MA – ADOS

| | AD | PDD-NOS | ASD Low-MA | *F* | *p* |
|---|---|---|---|---|---|
| ADOS T1 Total | 7.89 (1.79) | 5.66 (2.16) | 7.3 (1.56) | 5.81 | .004 |
| ADOS T2 Total | 5.90 (2.57) | 5.05 (2.75) | 7.1 (2.22) | | |
| Change in ADOS | -1.99 (2.99) | -.60 (2.95) | -.16 (2.12) | | |

Table 11. ASD Symptoms/Severity between Evaluations, AD, PDD-NOS, ASD Low-MA – DSM-IV Total

| | AD | PDD-NOS | ASD Low-MA | *F* | *p* |
|---|---|---|---|---|---|
| DSM T1 Total | 7.15 (1.05) | 4.20 (.921) | 6.35 (1.85) | 16.692 | <.0005 |
| DSM T2 Total | 5.88 (2.43) | 4.96 (2.39) | 6.87 (1.29) | | |
| Change in DSM | -1.37 (2.57) | .76 (2.34) | .52 (1.83) | | |

Table 12. ASD Symptoms/Severity between Evaluations, AD, PDD-NOS, ASD Low-MA – DSM-IV Social

| | AD | PDD-NOS | ASD Low-MA | *F* | *p* |
|---|---|---|---|---|---|
| DSM T1 Total | 3.33 (.69) | 2.04 (.85) | 2.91 (.96) | 6.37 | .002 |
| DSM T2 Total | 2.40 (1.28) | 1.65 (1.22) | 2.91 (.733) | | |
| Change in DSM | -.94 (1.37) | -.39 (1.25) | .00 (1.17) | | |

Table 13. ASD Symptoms/Severity between Evaluations, AD, PDD-NOS, ASD Low-MA – DSM-IV Communication

| | AD | PDD-NOS | ASD Low-MA | *F* | *p* |
|---|---|---|---|---|---|
| DSM T1 Total | 1.96 (.5) | 1.46 (.53) | 1.70 (.47) | 3.82 | .024 |
| DSM T2 Total | 1.95 (.95) | 1.88 (.92) | 2.00 (.6) | | |
| Change in DSM | -.04 (1.11) | .42 (1.03) | .30 (.7) | | |

Table 14. ASD Symptoms/Severity between Evaluations, AD, PDD-NOS, ASD Low-MA – DSM-IV RRB's

| | AD | PDD-NOS | ASD Low-MA | *F* | *p* |
|---|---|---|---|---|---|
| DSM T1 Total | 1.85 (.82) | .70 (.74) | 1.74 (1.10) | 18.37 | <.0005 |
| DSM T2 Total | 1.55 (.96) | 1.43 (.98) | 1.96 (.88) | | |
| Change in DSM | -.33 (1.10) | .73 (1.13) | .22 (1.04) | | |

Appendix A

DSM-IV TR

**Autistic Disorder (299.00 DSM-IV)**

A. A total of six (or more) items from (1), (2), and (3), with at least two from (1), and one each from (2) and (3):

1. Qualitative impairment in social interaction, as manifested by at least two of the following:
    - Marked impairment in the use of multiple nonverbal behaviors such as eye to-eye gaze, facial expression, body postures, and gestures to regulate social interaction .
    - Failure to develop peer relationships appropriate to developmental level
    - A lack of spontaneous seeking to share enjoyment, interests, or achievements with other people (e.g., by a lack of showing, bringing, or pointing out objects of interest)
    - Lack of social or emotional reciprocity
2. Qualitative impairments in communication as manifested by at least one of the following:
    - Delay in, or total lack of, the development of spoken language (not accompanied by an attempt to compensate through alternative modes of communication such as gestures or mime)
    - In individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others
    - Stereotyped and repetitive use of language or idiosyncratic language
    - Lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level
3. Restricted repetitive and stereotyped patterns of behavior, interests, and activities, as manifested by at least one of the following:
    - Encompassing preoccupation with one or more stereotyped patterns of interest that is abnormal either in intensity or focus
    - Apparently inflexible adherence to specific, nonfunctional routines or rituals
    - Stereotyped and repetitive motor mannerisms (e.g., hand or finger flapping or twisting, or complex whole-body movements)
    - Persistent preoccupation with parts of object

B. Delays or abnormal functioning in at least one of the following areas, with onset prior to age 3 years:

- Social interaction
- Language as used in social communication
- Symbolic or imaginative play

C. The disturbance is not better accounted for by Rett's Disorder or Childhood Disintegrative Disorder.

**PDD-NOS (299.80 DSM-IV)**

The essential features of PDD-NOS are severe and pervasive impairment in the development of reciprocal social interaction or verbal and nonverbal communication skills; and stereotyped behaviors, interests, and activities. The criteria for Autistic Disorder are not met because of late age onset; atypical and/or sub- threshold symptomotology are present.

This category should be used when there is a severe and pervasive impairment in the development of reciprocal social interaction or verbal and nonverbal communication skills, or when stereotyped behavior, interests, and activities are present, but the criteria are not met for a specific Pervasive Developmental Disorder, Schizophrenia, Schizotypical Personality Disorder, or Avoidant Personality Disorder. For example, this category includes "atypical autism"– presentations that do not meet the criteria for Autistic Disorder because of late age of onset, atypical symptomatology, or sub-threshold symptomatology, or all of these.

Appendix B

DSM V

Autism Spectrum Disorder 299.00 (F84.0)

Diagnostic Criteria

A. Persistent deficits in social communication and social interaction across multiple contexts, as manifested by the following, currently or by history (examples are illustrative, not exhaustive, see text):

1. Deficits in social-emotional reciprocity, ranging, for example, from abnormal social approach and failure of normal back-and-forth conversation; to reduced sharing of interests, emotions, or affect; to failure to initiate or respond to social interactions.

2. Deficits in nonverbal communicative behaviors used for social interaction, ranging, for example, from poorly integrated verbal and nonverbal communication; to abnormalities in eye contact and body language or deficits in understanding and use of gestures; to a total lack of facial expressions and nonverbal communication.

3. Deficits in developing, maintaining, and understanding relationships, ranging, for example, from difficulties adjusting behavior to suit various social contexts; to difficulties in sharing imaginative play or in making friends; to absence of interest in peers.

*Specify* current severity:

**Severity is based on social communication impairments and restricted repetitive patterns of behavior** (see Table 2).

B. Restricted, repetitive patterns of behavior, interests, or activities, as manifested by at least two of the following, currently or by history (examples are illustrative, not exhaustive; see text):

1. Stereotyped or repetitive motor movements, use of objects, or speech (e.g., simple motor stereotypies, lining up toys or flipping objects, echolalia, idiosyncratic phrases).

2. Insistence on sameness, inflexible adherence to routines, or ritualized patterns or verbal nonverbal behavior (e.g., extreme distress at small changes, difficulties with transitions, rigid thinking patterns, greeting rituals, need to take same route or eat food every day).

3. Highly restricted, fixated interests that are abnormal in intensity or focus (e.g, strong attachment to or preoccupation with unusual objects, excessively circumscribed or perseverative interest).

4. Hyper- or hyporeactivity to sensory input or unusual interests in sensory aspects of the environment (e.g., apparent indifference to pain/temperature, adverse response to specific sounds or textures, excessive smelling or touching of objects, visual fascination with lights or movement).

*Specify* current severity:

**Severity is based on social communication impairments and restricted, repetitive patterns of behavior** (see Table 2).

C. Symptoms must be present in the early developmental period (but may not become fully manifest until social demands exceed limited capacities, or may be masked by learned strategies in later life).

D. Symptoms cause clinically significant impairment in social, occupational, or other important areas of current functioning.

E. These disturbances are not better explained by intellectual disability (intellectual developmental disorder) or global developmental delay. Intellectual disability and autism spectrum disorder frequently co-occur; to make comorbid diagnoses of autism spectrum disorder and intellectual disability, social communication should be below that expected for general developmental level.

**Note:** Individuals with a well-established DSM-IV diagnosis of autistic disorder, Asperger's disorder, or pervasive developmental disorder not otherwise specified should be given the diagnosis of autism spectrum disorder. Individuals who have marked deficits in social communication, but whose symptoms do not otherwise meet criteria for autism spectrum disorder, should be evaluated for social (pragmatic) communication disorder.

*Specify* if:

**With or without accompanying intellectual impairment**
**With or without accompanying language impairment**
**Associated with a known medical or genetic condition or environmental factor**

Appendix C

Diagnostic Criteria for AD, ASD Low-MA, PDD-NOS

# Autistic Disorder

_____ At least two symptoms in Cluster 1 (Social) DSM-IV-TR checklist relative to developmental level

**AND**

_____ At least one symptom in Cluster 2 (Communication)

**AND**

_____ At least one symptom in Cluster 3 (Repetitive and/or Restricted Interests and Behaviors)

**AND**

_____ Child displays SIX OR MORE total symptoms

**AND**

_____ Onset was before age three

**AND**

_____ Child's age equivalence must be 12 months or higher on at least one of the following: **Mullen** Visual Reception, Receptive Language, or Expressive Language

# ASD-Low MA

_____ Child displays at least 1 symptom from Cluster 1 (Social)
-Must have 1 symptom other than 1b!!

**AND**

_____ Child displays at least 1 other symptom from Cluster 2 (Communication)
and/or Cluster 3 (Repetitive and/or Restricted Interests and Behaviors)

**AND**

_____ Child's **Mullen** scores on Visual Reception, Receptive Language, and
Expressive Language are ALL less than or equal to 12 months AE

# Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS)

_____ At least one symptom in Cluster 1 (Social) DSM-IV-TR checklist relative to developmental level:

_____ CANNOT include ONLY **1b** for Time 1 evaluations

**AND**

_____ At least one symptom in Cluster 2 (Communication) and/or Cluster 3 (Repetitive and/or Restricted Interests and Behaviors)

**AND**

_____ Child does not meet criteria for Autistic Disorder, Asperger's Disorder, or Rett's Syndrome

**AND**

_____ Symptoms noted on checklist cannot be better accounted for by another disorder (e.g., reactive attachment disorder, sensory or motor impairments, etc.)

**AND**

_____ Child's age equivalence must be 12 months or greater on at least one of the following: **Mullen** Visual Reception, Receptive Language, or Expressive Language

**AND**

_____ Child displays clinically significant impairment in home, school, and/or community settings