

# Autism Spectrum Disorders and Low Mental Age: Diagnostic Stability and Developmental Outcomes in Early Childhood

Alexander J. Hinnebusch<sup>1</sup> · Lauren E. Miller<sup>2</sup> · Deborah A. Fein<sup>2</sup>

Published online: 31 August 2017  
© Springer Science+Business Media, LLC 2017

**Abstract** Some children with autism spectrum disorders (ASDs) exhibit low mental age (Low-MA; i.e., cognitive functioning below 12 months). We examined diagnosis, symptom severity, and development in children with ASD-low MA ( $n=25$ ), autistic disorder ( $n=111$ ), and PDD-NOS ( $n=82$ ) at ages two and four. We predicted that some ASD-low MA children would demonstrate just intellectual impairment and not autism symptoms on follow-up, with social deficits at age two attributable to global delays. Instead, most ASD-low MA children (96%) had an ASD at follow-up, compared to children initially diagnosed with autistic disorder (86.5%) or PDD-NOS (73.2%). They showed the least developmental progress and highest symptom severity. Results support diagnosing ASDs in children functioning below a 12-month level.

**Keywords** Autism spectrum disorders · Low mental age · Developmental growth · Diagnostic stability · Autism severity

## Introduction

Autism spectrum disorders (ASDs) are characterized by behavioral deficits in social relatedness and communication, along with the presence of restricted interests and repetitive

behaviors (RRBs). The autism spectrum, as per DSM-IV-TR, encompasses several diagnoses that share these symptoms, including autistic disorder (AD), pervasive developmental disorder-not otherwise specified (PDD-NOS), and Asperger's disorder (American Psychiatric Association 2000).

Early identification and intervention are key in improving the prognosis of children with ASDs, and intervention significantly lowers long-term societal costs (Chasson et al. 2007). Diagnosis has been shown to be reliable and stable in children ages 18–24 months (Baron-Cohen et al. 1996; Charman et al. 2005; Moore and Goodson 2003; Lord 1995; Stone et al. 1999). Instruments such as the Checklist for Autism in Toddlers (CHAT) (Baron-Cohen et al. 2000), Screening Tool for Autism in 2-year-olds (STAT) (Stone et al. 2008), and Modified Checklist for Autism in Toddlers (M-CHAT) (Robins et al. 2001) and its revision, the M-CHAT-R (Robins et al. 2014), screen for ASDs in children as young as 16 months. However, some children in that age range are expected to have developmental age equivalents below 12 months, given the developmental delays commonly observed in children with ASDs. Estimated rates of cognitive impairment or intellectual disability in individuals with ASDs range from 24 percent (Chakrabarti and Fombonne 2001) to 70 percent (Fombonne 2003). Specifically, in Fombonne's (2003) review of 32 epidemiological studies, 30 percent of individuals with ASDs presented with mild to moderate cognitive impairment, while 40 percent were severely to profoundly impaired.

The reliability of measuring IQ in children with ASDs has been scrutinized because the social communication impairments and related behavioral issues common to ASDs may confound the testing process (e.g., limited comprehension of instructions, lack of imitative and reciprocal interactive skills, preoccupations or repetitive behaviors, distractibility, lack of a desire to please). There is also a degree of

✉ Lauren E. Miller  
lauren.miller@uconn.edu

<sup>1</sup> Department of Physical Medicine and Rehabilitation, University of Michigan Health System, 325 East Eisenhower Parkway, Suite 100, Ann Arbor, MI 48108, USA

<sup>2</sup> Department of Psychological Sciences, University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, CT 06269, USA

diagnostic overlap between ASDs and intellectual disability (ID), as both are characterized by cognitive, adaptive, and social skills deficits and often involve challenging and stereotyped behaviors (Matson and Shoemaker 2009). Despite this, Osterling et al. (2002) demonstrated that children who were later diagnosed with an ASD could be distinguished behaviorally from typically developing children and those with ID by 12 months of age. Additionally, contemporary research has demonstrated adequate stability of nonverbal intelligence in children with ASDs (Lord and Schopler 1989a, b; Sigman and Ruskin 1999).

Despite previous research demonstrating stability of ASDs and ID, even in young children, the validity of autism diagnoses in children with concurrent low mental age, defined here as age equivalents below 12 months in all cognitive domains, has not yet been demonstrated. For example, the Autism Diagnostic Interview-Revised (ADI-R) is only valid for children with a mental age above 24 months (Rutter et al. 2005). The Autism Diagnostic Observation Schedule (ADOS) has a toddler module that allows for diagnostic assessment in children as young as 12 months (Luyster et al. 2009), but not at developmental levels below that point. If autism screening is implemented at 16 months or earlier (see First Year Inventory (Reznick et al. 2007) and Infant-Toddler Checklist (Wetherby et al. 2008), which are designed to screen in children as young as 12 months, and Early Screening of Autistic Traits (Swinkels et al. 2006), designed for use in 14-month-olds), developmental levels may not permit valid diagnostic assessment of autism symptoms. For example, the M-CHAT (Robins et al. 2001) assesses the presence of some behaviors that Inada, Kamio, and Koyama (2010) suggest are usually present in typical development by 12 months, but others that do not usually appear until 15 months of age. Thus, if the M-CHAT, or another measure, were used to screen a child with a mental age below 12 months, these behaviors could be absent on the basis of developmental delay alone, and this could potentially affect the validity of such a screener to detect an ASD, producing false positives. In addition to global delays, these children might show restricted and repetitive behaviors (RRBs) common in persons with ID (Oliver et al. 2012), further leading to invalid autism diagnoses. Matson and Shoemaker (2009) call for research that validates autism diagnostic procedures in children with intellectual impairment.

### Current Study

Data collection began before the publication of DSM-5; therefore, DSM-IV-TR criteria (APA, 2000) were used for AD and PDD-NOS. To our knowledge, this is the first study to follow children with the ASD-low MA presentation, which we define as having (1) sufficient symptoms of an ASD to warrant at least a DSM-IV-TR PDD-NOS diagnosis

and (2) receptive and expressive language and nonverbal reasoning age equivalent scores below a 12 month level. Children were initially evaluated at approximately 24 months of age and were then reevaluated about 2 years later to determine whether they retained an ASD diagnosis and to assess autism severity and degree of cognitive and adaptive progress.

Specifically, we predicted that:

1. Children who initially met criteria for ASD-low MA would show significantly smaller developmental gains across time than children with AD or PDD-NOS.
2. A smaller percentage of children with an initial diagnosis of ASD-low MA would remain on the autism spectrum compared to those who initially met criteria for AD or PDD-NOS, as their early symptoms and delays might be due to global delay rather than an ASD. Furthermore, we predicted that children with an initial diagnosis of AD would have the highest diagnostic stability. This hypothesis was based on the AD children achieving a mental age by which prosocial and reciprocal communication behaviors are generally apparent and exhibiting RRBs at the time of initial evaluation, as well as the fact that children with PDD-NOS demonstrate lower levels of impairment than children with AD (Cohen et al. 1986; Sevin et al. 1995) and are more likely to lose their autism diagnosis over time (Anderson et al. 2014; Lord et al. 2006; Sutera et al. 2007; Helt et al. 2008; Berry 2009).
3. The AD group would demonstrate the most severe autism symptoms at both time points and the least change in symptom severity across time compared to the PDD-NOS and ASD-low MA groups. This hypothesis was based on the fact that children diagnosed with AD on initial evaluation already showed the full syndrome, and their symptoms could not be attributed simply to global delay.

## Methods

### Participants

Participants were 219 children drawn from a larger, federally funded investigation of the early detection of ASDs. In order to be included in the current study, participants (1) screened positive on the M-CHAT (Robins et al. 2001) or M-CHAT-R (Robins et al. 2014) between the ages of 16 and 30 months, (2) were evaluated through the research project and diagnosed with an ASD (i.e., AD, PDD-NOS, or ASD-low MA) at the approximate age of 2 years (Time 1), and (3) returned for a follow-up evaluation at the approximate age of 4 years (Time 2). Exclusion criteria included

significant sensory impairments (e.g., blindness) or deficits in motor functioning (e.g., severe cerebral palsy) that would impact a child’s ability to complete testing. Given that the larger study from which participants were drawn aimed to develop an autism-specific screening measure, children were also excluded from the current project if they had a prior diagnosis of an ASD by a qualified person. Validating an autism-specific screening tool using a sample of children who already have a diagnosis of an ASD is not appropriate, as a prior diagnosis may impact parents’ approach to screening questions, and screening is not needed if a diagnosis has already been established.

*Demographic Data*

Demographic information is presented in Table 1. The mean age of children in the ASD-low MA group at Time 1 was significantly different than that of the other groups, about 2.5 months younger. Consistent with autism prevalence data, males predominated, and there were no significant differences in gender between the AD, PDD-NOS, and ASD-low MA groups. There were no significant differences between groups on ethnicity or income.

**Table 1** Participant demographics by group

Variable	AD	PDD-NOS	ASD-low MA	Statistic	<i>p</i>
Age [M (SD)]					
Time 1	27.1 (4.6)	25.9 (4.0)	23.6 (4.5)	$F(4, 428) = 3.67$	<0.006
Time 2	53.5 (9.4)	51.5 (10.4)	49.4 (6.5)	–	–
Gender [N (%)]					
Male	93 (83.8)	65 (78.3)	20 (80.0)	$\chi^2(2, N = 219) = 0.964$	0.617
Female	18 (16.2)	18 (21.7)	50 (20.0)		
Total	111	83	25		
Ethnicity [N (%)]					
White	87 (82.1)	67 (84.8)	19 (76.0)	$\chi^2(8, N = 210) = 7.450$	0.490
Black	2 (1.9)	4 (5.1)	2 (8.0)		
Hispanic/Latino	9 (8.5)	5 (6.3)	3 (12.0)		
Asian	4 (3.8)	3 (3.8)	0		
Biracial	4 (3.8)	0	1 (4.0)		
Total	106	79	25		
Maternal education [N (%)]					
No degree	3 (4.3)	2 (3.1)	0	$\chi^2(12, N = 154) = 9.675$	0.645
High school diploma/GED	22 (31.4)	20 (30.8)	11 (57.9)		
Vocational/technical degree	5 (7.1)	5 (7.7)	0		
Associate’s degree	6 (8.6)	6 (9.2)	0		
Bachelor’s degree	20 (28.6)	19 (29.2)	4 (21.1)		
Master’s degree	11 (15.7)	12 (18.5)	4 (21.1)		
Doctoral degree	3 (4.3)	1 (1.5)	0		
Total	70	65	19		
Annual income [N (%)]					
<\$10,000	6 (9.5)	1 (1.6)	0	$\chi^2(20, N = 143) = 23.166$	0.281
\$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		

Age is presented in months

### Diagnostic Criteria

DSM-IV-TR diagnostic criteria for AD and PDD-NOS were used. In addition, to meet study criteria for either of these two diagnoses, at least one of the child's age equivalent scores on the Mullen Scales of Early Learning (MSEL) visual reception, receptive language, or expressive language scales must have been at or above a 12 month level.

Children received a research diagnosis of ASD-low MA if they 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 RRBs cluster. Additionally, these children received age equivalent scores below 12 months on the MSEL visual reception, receptive language, and expressive language scales, as well as on the Vineland Adaptive Behavior Scales (VABS) communication and socialization domains (see Table 2).

At Time 1, 111 children received a diagnosis of AD, 83 received a diagnosis of PDD-NOS, and 25 received a diagnosis of ASD-low MA. To assess differences among these three groups on follow-up, participants were coded into six groups based on Time 2 diagnosis: AD, PDD-NOS, ASD-low MA, developmental delay, other diagnosis, and no diagnosis. Participants were also coded based on whether their Time 2 diagnosis was an ASD diagnosis (e.g., AD, PDD-NOS, or ASD-low MA) or not (e.g., developmental delay, other diagnosis, no diagnosis). Criteria for all diagnoses are listed in "Appendix".

### Procedures

Participants were screened with either the M-CHAT (Robins et al. 2001) or M-CHAT-R (Robins et al. 2014) at their pediatrician's office or through their early intervention (i.e.,

non-ASD services) provider between the ages of 16 and 30 months. Children who screened positive on the initial questionnaire, as well as the follow-up phone interview, were offered a free developmental and diagnostic evaluation, which was conducted by a licensed psychologist or a developmental-behavioral pediatrician and a clinical psychology doctoral student. Most evaluations took place at the research team's university clinic, and families who did not have transportation were provided with a free taxi service. In some cases, study staff traveled to conduct evaluations at participating pediatric offices in two large towns with a high proportion of low income patients. Diagnoses were based on clinical best estimate judgment of symptoms from observation, developmental history, and testing data on the Autism Diagnostic Observation Schedule (ADOS), Childhood Autism Rating Scale (CARS), MSEL, and VABS.

Children became eligible for a follow-up evaluation (Time 2) at age 42 months or older and were recontacted by letter or telephone. This evaluation included the same measures and diagnostic procedures as Time 1.

### Measures

#### *Mullen Scales of Early Learning (MSEL)*

The MSEL (Mullen 1995) is a standardized test of cognitive ability for children between birth and age 68 months. The visual reception, fine motor, expressive language, and receptive language scales were administered in this study. Each scale yields an age equivalent score, which reflects the child's current developmental level. The MSEL was normed on a nationally representative sample of 1849 children (48.7% female, 51.3% male) and is a frequently used measure of developmental level and cognitive functioning in both typically developing children and children with

**Table 2** Mean (SD) age equivalent scores by group

Variable	AD		PDD-NOS		ASD-low MA	
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
MSEL visual reception	18.1 (5.4) <i>n</i> =80	37.7 (16.0)	18.8 (4.5) <i>n</i> =66	42.7 (15.0)	8.8 (2.1) <i>n</i> =19	19.8 (8.1)
MSEL fine motor	19.7 (6.5) <i>n</i> =80	36.3 (13.9)	19.9 (4.2) <i>n</i> =66	47.8 (12.0)	12.5 (2.8) <i>n</i> =19	19.7 (5.5)
MSEL receptive language	11.6 (6.0) <i>n</i> =79	32.3 (16.3)	14.0 (6.5) <i>n</i> =66	38.3 (15.1)	6.8 (3.2) <i>n</i> =19	15.2 (8.6)
MSEL expressive language	13.1 (6.5) <i>n</i> =79	30.8 (14.4)	15.2 (5.6) <i>n</i> =65	35.8 (12.8)	7.3 (2.1) <i>n</i> =19	15.4 (9.1)
VABS receptive language	12.7 (7.4) <i>n</i> =83	30.3 (16.6)	15.2 (7.2) <i>n</i> =72	36.7 (17.8)	7.8 (2.9) <i>n</i> =19	18.0 (11.1)
VABS expressive language	10.7 (5.6) <i>n</i> =83	29.2 (31.6)	13.4 (5.9) <i>n</i> =72	31.9 (15.3)	7.8 (2.6) <i>n</i> =19	14.0 (6.1)
VABS interpersonal relationships	9.7 (4.1) <i>n</i> =81	22.6 (14.7)	12.1 (5.5) <i>n</i> =70	28.1 (14.8)	7.7 (2.1) <i>n</i> =18	11.7 (4.7)
VABS play and leisure	10.0 (4.5) <i>n</i> =81	20.7 (13.8)	13.3 (5.7) <i>n</i> =69	24.6 (14.8)	7.3 (4.2) <i>n</i> =17	11.8 (6.7)
VABS coping	11.9 (6.7) <i>n</i> =56	24.6 (15.1)	11.1 (5.2) <i>n</i> =52	31.4 (16.8)	9.9 (3.0) <i>n</i> =11	12.8 (6.5)

Age equivalents are presented in months. *Ns* indicate number of participants with data at both time points

*MSEL* Mullen Scales of Early Learning, *VABS* Vineland Adaptive Behavior Scales

developmental delays. It demonstrates satisfactory internal consistency and test–retest reliability (Mullen 1995).

Age equivalents for each group are shown in Table 2. Because the three groups differed in chronological age by several months, a developmental quotient (DQ) was calculated (Table 3). The decision to use a DQ (i.e., ratio IQ, using the formula mental age equivalent divided by chronological age, multiplied by 100) instead of a traditional deviation-type standardized score was based upon other studies of children with low mental ages showing that standardized scores have restricted ranges in this population (DiLavore et al. 1995; Kanne et al. 2011; Sallows and Graupner 2005; Schopler et al. 1989).

To estimate development between time points, Time 1 age equivalent scores for each MSEL scale were subtracted from corresponding Time 2 age equivalent scores. This difference can be considered the amount of mental growth made between evaluations. Number of months between Time 1 and Time 2 evaluations was then calculated for each child, and a ratio (i.e., mental growth divided by time elapsed) was used to assess rate of progress (Table 4). A

growth rate >1 indicates that a child made more rapid progress than expected, a rate of 1 indicates expected progress, and a number <1 indicates slower progress than expected. For example, if a child made the equivalent of 6 months' growth in receptive language over the course of a year, his growth rate would be 0.5 (i.e., half the growth expected in typical development), whereas if he made 18 months of mental growth over the course of a year, his growth rate would be 1.5 (i.e., faster growth than expected in typical development). This value is often called the learning rate (Howard et al. 2005; Eikeseth et al. 2012; Klintwall et al. 2015), although Bagnato and Neisworth (1980) called it the intervention efficiency index. Klintwall et al. (2015) summarized the advantages of using this value to gauge developmental progress: it can be compared even when different time periods have elapsed or children are of similar mental age but different chronological age, if a child makes no progress the Learning Rate equals 0 rather than a decrease as one would see if using standard scores, and it is easier to explain to parents and teachers, especially if graphed. Similar methods of indexing

**Table 3** Mean (SD) developmental quotient scores by group

Variable	AD	PDD-NOS	ASD-low MA	<i>F</i>	<i>p</i>
MSEL visual reception					
Time 1	67.0 (19.7)	73.6 (16.8)	39.7 (12.0)	<i>F</i> (2, 175)=28.284	<0.001
Time 2	73.1 (30.3)	85.7 (28.4)	41.3 (18.2)	<i>F</i> (2, 183)=21.903	<0.001
MSEL fine motor					
Time 1	73.7 (25.6)	77.3 (14.2)	55.6 (14.6)	<i>F</i> (2, 175)=9.248	<0.001
Time 2	70.4 (25.1)	75.7 (21.5)	40.8 (12.8)	<i>F</i> (2, 181)=22.042	<0.001
MSEL receptive language					
Time 1	38.9 (17.3)	49.6 (20.8)	29.7 (15.8)	<i>F</i> (2, 176)=11.705	<0.001
Time 2	63.0 (31.2)	77.0 (21.9)	31.3 (17.1)	<i>F</i> (2, 182)=22.400	<0.001
MSEL expressive language					
Time 1	48.3 (23.5)	59.1 (18.8)	33.3 (12.7)	<i>F</i> (2, 174)=13.517	<0.001
Time 2	59.9 (27.4)	72.1 (25.1)	32.1 (19.4)	<i>F</i> (2, 182)=21.777	<0.001
VABS receptive language					
Time 1	45.7 (24.8)	59.5 (28.9)	33.3 (12.7)	<i>F</i> (2, 187)=10.960	<0.001
Time 2	57.3 (30.8)	71.4 (33.6)	37.0 (22.9)	<i>F</i> (2, 200)=11.849	<0.001
VABS expressive language					
Time 1	39.1 (18.9)	52.5 (24.4)	33.9 (13.3)	<i>F</i> (2, 187)=11.344	<0.001
Time 2	50.0 (24.9)	61.8 (26.0)	28.9 (12.6)	<i>F</i> (2, 199)=17.063	<0.001
VABS interpersonal relationships					
Time 1	35.4 (14.0)	47.3 (22.3)	34.2 (11.5)	<i>F</i> (2, 184)=10.588	<0.001
Time 2	42.2 (25.3)	54.7 (27.4)	24.1 (9.7)	<i>F</i> (2, 198)=14.525	<0.001
VABS play and leisure					
Time 1	36.9 (16.6)	52.5 (23.2)	33.2 (20.2)	<i>F</i> (2, 182)=14.936	<0.001
Time 2	39.2 (24.9)	47.7 (23.7)	24.3 (14.1)	<i>F</i> (2, 198)=9.242	<0.001
VABS coping					
Time 1	43.1 (22.6)	42.3 (20.4)	43.4 (16.4)	<i>F</i> (2, 131)=0.026	0.975
Time 2	46.2 (26.6)	60.6 (30.1)	25.9 (12.9)	<i>F</i> (2, 186)=15.034	<0.001

MSEL Mullen Scales of Early Learning, VABS Vineland Adaptive Behavior Scales

**Table 4** Mean (SD) developmental progress between evaluations by group

	AD	PDD-NOS	ASD-low MA	<i>F</i>	<i>p</i>
MSEL visual reception	0.86 (0.59)	0.99 (0.51)	0.48 (0.43)	6.693	0.002
MSEL fine motor	0.69 (0.43)	0.74 (0.41)	0.27 (0.24)	10.732	<0.0005
MSEL receptive language	0.77 (0.52)	0.85 (0.47)	0.35 (0.34)	8.264	<0.0005
MSEL expressive language	0.88 (0.60)	0.99 (0.51)	0.34 (0.34)	4.31	0.015
VABS receptive language	0.80 (1.32)	0.76 (0.48)	0.24 (0.98)	2.628	0.075
VABS expressive language	0.77 (0.62)	0.91 (0.7)	0.41 (0.47)	4.777	0.01
VABS interpersonal relationships	0.54 (0.55)	0.66 (0.58)	0.13 (0.15)	6.942	0.001
VABS play and leisure	0.46 (0.57)	0.47 (0.51)	0.14 (0.25)	3.062	0.049
VABS coping	0.50 (0.64)	0.82 (0.63)	0.06 (0.27)	8.493	<0.0005

Developmental progress is a ratio of mental growth over chronological time elapsed between evaluation time points

*MSEL* Mullen Scales of Early Learning, *VABS* Vineland Adaptive Behavior Scales

developmental progress have been used previously in the literature, particularly when evaluating a child's developmental gains in response to early intervention (Bagnato and Neisworth 1985; Shonkoff et al. 1992; Wolery and Dyk 1985).

#### *Vineland Adaptive Behavior Scales, Interview Edition (VABS)*

The VABS (Sparrow et al. 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 VABS has well-established reliability and validity (Sparrow et al. 1984) and is frequently used with varied clinical populations (Klin et al. 1997). For the range of ages included in the current study sample, the VABS demonstrates high internal consistency and test–retest reliability (Sparrow et al. 2005).

Age equivalents are shown in Table 2. The same methods used to calculate DQ scores and growth between evaluations for MSEL scales was used to create similar scores for VABS socialization and communication domains, whose content is relevant to autism (Tables 3, 4).

#### *Autism Diagnostic Observation Schedule—Generic (ADOS)*

The ADOS (Lord et al. 2000) is a semi-structured assessment designed to measure symptoms of ASDs. Only behaviors viewed during test administration are scored on this measure. The ADOS includes four modules, one of which is administered based on the child's expressive language level and chronological age. The current study used Modules 1 and 2. All children in the current study met the autism spectrum cut-off on the ADOS, in addition to meeting

DSM-IV-TR diagnostic criteria as determined by a senior clinician.

#### *Childhood Autism Rating Scale (CARS)*

The CARS (Schopler et al. 1980) is a behavior rating scale that consists of 15 items measuring behaviors associated with an ASD. The scale is based on a clinician's direct observation of the child, incorporating parent report of behaviors. The CARS yields a numerical score of autism symptom severity. This score can be used to label a child's symptoms as non-autistic, mild, moderate, or severe. In addition to these established cut-offs, Chlebowski et al. (2010) defined a cut-off score for PDD-NOS.

#### **Data Analytic Plan**

A series of univariate analyses of variance (ANOVA) were conducted to determine differences on DQs between the AD, PDD-NOS, and ASD-low MA groups at each time point (Table 3), as well as in the overall level of developmental gains made between the two evaluations (Table 4). Power analyses revealed sufficient power (i.e., power = .80,  $\alpha = 0.05$ ) to detect small to medium effect sizes (i.e.,  $\eta^2 = 0.24 - 0.29$ , depending on MSEL/VABS domain).

We also evaluated differences in diagnostic stability among the AD ( $n = 111$ ), PDD-NOS ( $n = 82$ ), and ASD-low MA ( $n = 25$ ) groups with Chi-square tests, comparing the three groups by (1) diagnostic outcome, with all six Time 2 diagnostic possibilities included (Table 5), and (2) diagnostic outcome, differentiating ASD versus non-ASD (Table 6). When comparing all six diagnostic outcomes, power analyses revealed sufficient power (i.e., power = .80,  $\alpha = 0.05$ ) to detect a medium to large effect size ( $\phi = 0.27$ ), whereas comparison of autism spectrum versus non-ASD diagnoses revealed sufficient power to detect a small to medium effect ( $\phi = 0.19$ ).

**Table 5** Diagnostic stability by group: all diagnostic outcomes at time 2

Time 1 diagnosis	Time 2 diagnosis						Statistic
	AD	PDD-NOS	ASD-low MA	DD	Other Dx	No Dx	
AD	75 (67.7%)	21 (18.9%)	0	2 (1.8%)	1 (0.9%)	12 (10.8%)	$X^2 = 74.828, p < .0005, \text{Phi} = 0.506$
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	

Data are presented as frequency counts and percentages

DD developmental delay, Other Dx other DSM-IV-TR diagnosis, No Dx no DSM-IV-TR diagnosis

**Table 6** Diagnostic stability by group: ASD versus non-ASD diagnostic outcomes at time 2

Time 1 diagnosis	Time 2 diagnosis		Statistic
	ASD	Non-ASD	
AD	96 (86.5%)	15 (13.5%)	$X^2 = 9.349, p = .009, \text{Phi} = 0.207$
PDD-NOS	60 (73.5%)	22 (26.8%)	
ASD-low MA	24 (96.0%)	1 (4.0%)	

Data are presented as frequency counts and percentages. ASD diagnoses include AD, PDD-NOS, and ASD-low MA. Non-ASD diagnoses include developmental delay, other diagnosis, and no diagnosis

**Table 7** Mean (SD) ASD severity on CARS by group

	AD	PDD-NOS	ASD-low MA	F	p
Time 1 CARS	35.3 (4.4)	28.2 (3.6)	35.1 (5.7)	9.12	<0.0005
Time 2 CARS	31.0 (7.1)	26.8 (5.8)	36.3 (4.9)		
Change in CARS	-4.4 (6.6)	-1.5 (5.8)	1.2 (6.8)		

CARS Childhood Autism Rating Scale

AD N = 103, PDD-NOS N = 76, ASD-low MA N = 21

Finally, a mixed-model ANOVA was conducted to investigate the relationship between initial diagnosis and Time 1 and Time 2 autism symptom severity as measured by the CARS (Table 7). For all analyses, on a case-by-case basis, participants with missing data were excluded. Power analyses revealed sufficient power (i.e., power = .80,  $\alpha = 0.05$ ) to detect a small to medium effect size (i.e.,  $\eta^2 = 0.22$ ).

## Results

### Differential Attrition

Three hundred eighty-nine children were diagnosed with an ASD at Time 1. Of these, 44 percent ( $n = 171$ ) declined, or were unable to be contacted for, a Time 2 evaluation. Specifically, 47 percent ( $n = 100$ ) of children diagnosed with AD, 43 percent ( $n = 61$ ) with PDD-NOS, and 29 percent

( $n = 10$ ) with ASD-low MA at Time 1 did not return for a Time 2 evaluation. When compared to children who returned at follow-up, those who did not return did not significantly differ on Time 1 age [ $t(387) = -1.751, p = .081$ ], diagnosis ( $X^2(2, N = 389) = 4.655, p = .098$ ), autism severity, as measured by the CARS [ $t(372) = 1.400, p = .162$ ], or nonverbal IQ, as measured by MSEL visual reception [ $t(321) = 0.417, p = .677$ ]. However, the two groups did significantly differ on ethnicity [ $X^2(6, N = 389) = 28.817, p < .001$ ] and gender [ $X^2(1, N = 389) = 4.607, p = .032$ ]. Those who returned at Time 2 were more likely to be White and male. See Table 8 for a summary of differential attrition data.

### Developmental Growth by Group

Age equivalents and DQs on MSEL scales and VABS domains at both time points are shown in Tables 2 and 3. In general, across domains, the ASD-low MA group performed lower on measures of cognitive and adaptive functioning, though children in the PDD-NOS group demonstrated somewhat higher language ability and social and play skills than those in both other groups (Table 3). Developmental growth rate, or Learning Rate, as previously explained, is shown by group in Table 4. We predicted that children with an initial diagnosis of ASD-low MA would show significantly smaller developmental gains across time than children with AD or PDD-NOS.

Overall, all domains were significantly different by group, except that VABS expressive language was only a trend. Post-hoc LSD analyses were performed for pairwise comparisons except where Levene's test for equality

**Table 8** Differential attrition

Time 1 variable	Returned at Time 2	Did not return at Time 2	Statistic	<i>p</i>
Age [ <i>M</i> ( <i>SD</i> )]	26.7 (4.6)	25.8 (4.9)	<i>t</i> (387) = -1.751	0.081
CARS total score [ <i>M</i> ( <i>SD</i> )]	32.6 (5.5)	33.3 (5.0)	<i>t</i> (372) = 1.400	0.162
MSEL visual reception DQ [ <i>M</i> ( <i>SD</i> )]	65.5 (20.1)	66.4 (18.9)	<i>t</i> (321) = 0.417	0.677
Gender [ <i>N</i> (%)]				
Male	178 (82)	124 (73)	$X^2(1, N=389) = 4.607$	0.032
Female	40 (18)	47 (27)		
Total	218	171		
Ethnicity [ <i>N</i> (%)]				
White	173 (79)	98 (57)	$X^2(6, N=389) = 28.817$	<0.001
Black	8 (4)	15 (9)		
Hispanic/Latino	17 (8)	16 (9)		
Asian	7 (3)	8 (5)		
Biracial	5 (2)	7 (4)		
Other	0 (0)	2 (1)		
Unknown	8 (4)	25 (15)		
Total	218	171		
Diagnosis [ <i>N</i> (%)]				
AD	110 (50)	100 (58)	$X^2(2, N=389) = 4.655$	0.098
PDD-NOS	83 (38)	61 (36)		
ASD-low MA	25 (12)	10 (6)		
Total	218	171		

Age is presented in months

CARS Childhood Autism Rating Scale, MSEL Mullen Scales of Early Learning

of variances was found to be violated, in which case a Games–Howell post-hoc analysis was conducted.

A post-hoc LSD analysis revealed significant differences in MSEL visual reception 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 4). The ASD-low MA group only demonstrated approximately half of the progress that was expected from the time elapsed, while the other two groups, especially children with PDD-NOS, made progress close to what was expected.

The mean growth rates for MSEL fine motor showed the same pattern, with significant differences in progress between the AD and ASD-low MA ( $p < .0005$ ) and the PDD-NOS and ASD-low MA ( $p < .0005$ ) groups, but not the AD and PDD-NOS groups ( $p = .74$ ). Again, the ASD-low MA group made less progress than the other two groups.

On MSEL expressive language, a post-hoc LSD analysis revealed significant differences in progress between the AD and ASD-low MA ( $p = .001$ ) and the PDD-NOS and ASD-low MA ( $p < .0005$ ) groups, but not the AD and PDD-NOS groups ( $p = .33$ ). Again, the ASD-low MA group exhibited the least progress over time.

The mean growth rates for MSEL receptive language were also significantly different by group,  $F(1, 162) = 11.37$ ,  $p < .0005$ . A Games–Howell post-hoc analysis revealed significant differences in progress between the AD and ASD-low MA ( $p < .0005$ ) and the PDD-NOS and ASD-low MA ( $p < .0005$ ) groups, but not the AD and PDD-NOS groups ( $p = .467$ ). The AD and PDD-NOS groups showed progress close to what would be expected for the time elapsed, while children in the ASD-low MA group made only about one-third as much progress as expected.

VABS expressive language showed a trend for an overall group difference; a post-hoc LSD analysis revealed significant differences in expressive language progress between the AD and ASD-low MA ( $p = .026$ ) and the PDD-NOS and ASD-low MA ( $p = .04$ ) groups, but not the AD and PDD-NOS groups ( $p = .82$ ), with the ASD-low MA group making only about a quarter of the progress expected.

VABS receptive language growth was significantly different by group [ $F(1, 171) = 4.78$ ,  $p = .01$ ]. A post-hoc LSD analysis revealed significant differences in progress between the AD and ASD-low MA ( $p = .026$ ) and the PDD-NOS and ASD-low MA ( $p = .002$ ) groups, but not the AD and PDD-NOS groups ( $p = .18$ ), again with the ASD Low-MA group making the least developmental progress over time.



VABS interpersonal relationships growth also differed by group [ $F(1, 166) = 6.94, p = .001$ ]. 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 ( $p < .0005$ ) groups, but not the AD and PDD-NOS groups ( $p = .42$ ). The AD and PDD-NOS groups made less progress here than in language domains, and did not differ from each other, while the ASD-low MA group made almost no measurable progress.

The mean mental growth rates for VABS play and leisure also differed by group. 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 ( $p = .001$ ) groups, but not the AD and PDD-NOS groups ( $p = .99$ ), with slower progress than in language domains for the AD or PDD-NOS groups, and almost no progress for the ASD Low-MA group.

The mean mental growth rates for VABS coping showed an overall significant group difference [ $F(1, 116) = 8.49, p < .0005$ ]. A Games–Howell post-hoc analysis revealed significant differences in progress between the AD and ASD-low MA ( $p = .002$ ) and the PDD-NOS and ASD-low MA ( $p < .0005$ ) groups. The AD and PDD-NOS groups also differed ( $p = .025$ ). The PDD-NOS group made close to expected progress in this domain, while the AD group made only about half the expected progress, and the ASD-low MA group showed almost no progress.

Data for VABS daily living are not presented here, as they do not reflect core autism symptoms. Even so, all of the subdomains in daily living showed the same pattern outlined above, namely that progress did not differ between AD and PDD-NOS groups, but was lower than both of these in the ASD-low MA group.

### Diagnostic Stability: All Outcomes at Time 2

To assess the prediction that fewer children initially diagnosed with ASD-low MA would remain on the autism spectrum, but would instead show non-autistic global delays, 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 all six Time 2 diagnostic outcomes.

Outcomes for children with each diagnosis at Time 1 are shown in Table 5. This initial Chi-square analysis revealed that there was an overall difference in diagnostic stability based on Time 1 diagnosis ( $X^2(10, N = 218) = 74.83, p < 0.0005, \Phi = 0.586$ ). Only one of the ASD-low MA children moved off the spectrum, contrary to prediction, and most of them stayed in either the significantly delayed ASD-low MA group or moved into the AD category.

### Diagnostic Stability: ASD versus Non-ASD Outcomes at Time 2

We also predicted that diagnostic stability (i.e., ASD vs. non-ASD) would be highest for the group of children initially diagnosed with AD. Of the 111 children that received an AD diagnosis at Time 1, 96 (86.5%) received an ASD diagnosis at Time 2, 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 at Time 2, while 22 (26.8%) received a non-ASD diagnosis; this group therefore had the most positive outcome, defined as transitioning off the autism spectrum. Of the 25 children that received an ASD-low MA diagnosis at Time 1, 24 (96%) received a diagnosis of an ASD at Time 2, while 1 (4%) received a non-ASD (but still developmentally delayed) 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 [ $X^2(2, N = 218) = 9.35, p = .009, \Phi = 0.207$ ] (see Table 6).

Follow-up Chi-square analyses were then conducted to identify the significant differences between pairwise groups on whether they retained a diagnosis of an ASD or received 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%) [ $X^2(1, N = 193) = 5.39, p = .026, \Phi = 0.167$ ]. The 86.5 percent of the AD group and the 96 percent of the ASD-low MA group remaining on the spectrum were not significantly different. The PDD-NOS and ASD-low MA groups were significantly different ( $X^2(1, N = 107) = 5.92, p = .013, \Phi = -0.235$ ).

Therefore, the prediction that children initially diagnosed with AD would show greater diagnostic stability than those in the PDD-NOS group was borne out. However, the prediction that a significant number of the ASD-low MA children would move from an ASD diagnosis to global developmental delay, because their apparent autism at Time 1 was due to global delays, was *not* confirmed. Instead, all but one of these children remained on the autism spectrum. Furthermore, most of these children moved into the AD category, and five were still functioning below a 12 month cognitive level at Time 2.

### Autism Severity

CARS scores at Time 1 and 2 are shown in Table 7. 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 points. The mean change in CARS score was significantly different among groups ( $F(2, 197) = 9.12,$

$p < .0005$ ). The AD and ASD-low MA scores are in the mild-to-moderate autism range, although the AD group improved to the lower end of this range, whereas the ASD-low MA group worsened slightly. The PDD-NOS group improved slightly but remained within the PDD-NOS range (Chlebowski et al. 2010). A post-hoc LSD analysis revealed significant differences in CARS score change between the AD and ASD-low MA ( $p < .0005$ ) and the AD and PDD-NOS ( $p = .003$ ) groups, but not the PDD-NOS and ASD-low MA ( $p = .089$ ) groups. Therefore, although the AD group showed strong stability, as predicted, by largely remaining on the autism spectrum, its members did show more symptomatic improvement than the other groups. The severe disability in the ASD-low MA group is very likely contributory to the observed pattern, as are the relatively milder autism symptoms of the PDD-NOS group.

## Discussion

The goal of the current study was to examine developmental growth, diagnostic stability, and autism symptom severity, over 2 years, in children with initial diagnoses of AD, PDD-NOS, and ASD-low MA. Overall, the ASD-low MA group demonstrated less developmental growth over time (i.e., slower progress) than the AD and PDD-NOS groups. Furthermore, results suggested high stability of autism spectrum diagnoses in children with Low MA. Children in the ASD-low MA group also showed consistent and severe autism symptoms over the 2-year study period, whereas children in the AD group exhibited some improvement in symptom severity from Time 1 to Time 2, and children in the PDD-NOS group had milder symptoms at their initial evaluation. Taken together, our findings provide clear support for the diagnosis of ASDs in very young children and those with cognitive delays and are important for understanding the expected developmental trajectory of these children.

Cognitive DQs on the MSEL generally showed improvement in the AD and PDD-NOS groups, with limited gains in the ASD-low MA group. However, developmental progress ratios below one suggested that all participants made less progress in mental age than what was expected based on the actual change in chronological age between evaluations. Growth that was 75 percent or more of what would be expected from the time elapsed was shown by the AD and PDD-NOS groups in MSEL visual reception, receptive language, and expressive language, and VABS expressive and receptive language, and by the PDD-NOS group in VABS coping. No group showed this much progress in MSEL fine motor, VABS interpersonal relationships, or play and leisure, and the ASD-low MA group did not show this much progress in any cognitive or adaptive domain.

The slower progress on the MSEL fine motor subdomain could be a result of its design. This scale relies on a child having a level of receptive language proficiency high enough to understand basic verbal cues, as well as the basic skills (e.g., eye contact, reciprocal interaction, imitation) to attend to a visual demonstration and replicate and sequence the observed actions correctly. Additionally, fine motor tasks may not be as much of a focus of autism-specific intervention services as are language and social skills. This hypothesis, however, would not explain the slower progress on VABS interpersonal relationships and play. One possibility is that these items are tapping into core autism impairments that are more difficult to ameliorate, even with intervention. Additionally, the VABS play and leisure subscale is largely composed of items addressing play with peers; it is possible that lack of interest in peers, also a core symptom of ASDs, is difficult to improve. Since the children in the current study were all approaching preschool age at Time 2, it is also likely that their intervention services between evaluations had focused on basic communication and socialization (e.g., eye contact), and thus perhaps they had not yet been treated in a setting with other children.

The ASD-low MA group had quite low growth ratios for all measured skill areas; they did make some progress on the cognitive and communication items, but almost none on the VABS socialization domains. This finding indicates just how limited developmental progress is for children in this diagnostic category, as these children on average made only a quarter to one-half of the progress expected of them between evaluations, even in cognitive and adaptive domains where they made the greatest gains. Thus, our first prediction, that the ASD-low MA group would make the least developmental progress over time, was strongly confirmed.

The second prediction hypothesized that our three groups would differ in diagnostic stability, with the AD group having the highest percentage of children remaining on the autism spectrum, and the ASD-low MA group the least, at follow-up. Both when considering all six specific Time 2 diagnoses, and when considering only autism versus non-ASD outcomes, the AD group did show a higher rate of diagnostic stability than the PDD-NOS group. Two-thirds of the AD group maintained their diagnosis, with most of the rest improving to PDD-NOS, and only roughly 10 percent losing an ASD diagnosis. In the PDD-NOS group, about one-third kept that diagnosis, about one-third increased in severity to AD (likely due to the emerging presence of RRBs later in the developmental period), and another third moved off the autism spectrum. Approximately half of these children still showed other developmental delays at Time 2. In the ASD-low MA group, contrary to predictions, all but one of the children maintained an autism spectrum diagnosis, with most meeting criteria for AD, 20 percent continuing to display characteristics of ASD-low MA, one child improving

to PDD-NOS, and one child moving off of the autism spectrum but still showing significant developmental delays. Thus, our second prediction regarding diagnostic stability was only partially supported, suggesting validity of early (i.e., 2-year-old) autism diagnoses in children with concurrent low mental age.

We also predicted that the AD group would demonstrate the most severe symptoms of an ASD at both time points and exhibit the smallest rates of symptom severity change across time, when compared to the PDD-NOS and ASD-low MA groups. Contrary to our prediction, the AD group showed the most improvement in autism severity as measured by the CARS, although their average score was still solidly within the autistic range. The PDD-NOS group showed significantly less improvement than the AD group; however, at both time points, they were still on average in the PDD-NOS range, with lower severity scores than the AD group. The ASD-low MA group had the most severe scores at Time 1 and showed minimal improvement. Thus, our third prediction regarding symptom severity was only partially supported, again supporting the validity of the diagnosis of ASDs in children with low mental age.

These findings suggest that ASDs can in fact be detected accurately, even in the presence of co-occurring low mental age, and that the symptoms of autism are severe and highly stable across time in this subgroup. This high degree of symptom stability provides further support to the legitimacy of assigning an autism spectrum diagnosis in the presence of low mental age, suggesting that the observed behaviors are in most cases indicative of an ASD, and not solely a product of significantly delayed developmental (i.e., cognitive and adaptive) milestones.

A high proportion of children in the AD and PDD-NOS groups also received an autism diagnosis at follow-up, though our findings were consistent with prior research indicating that children with PDD-NOS are more likely to lose their diagnosis over time. The elimination of the PDD-NOS diagnosis, and the introduction of the requirement of RRBs in DSM-5, might result in a number of these children failing to meet DSM-5 autism spectrum disorder diagnostic criteria (Barton et al. 2013) and thereby missing out on necessary intervention services. The greater improvement shown by the AD group might be attributable to their showing the full syndrome at age 2 years and receiving more intensive, targeted intervention, resulting in some amelioration of deficits. However, it should be noted that there were almost no significant differences in learning rate on cognitive or adaptive measures between the AD and PDD-NOS groups, suggesting that although they may differ somewhat in symptom severity, they do not show notable differences in cognitive or adaptive growth, while the ASD-low MA group was significantly different from the other two autism groups on level of functioning and growth rate on every variable.

## Limitations

There were a number of limitations that must be considered. While the overall sample size is large compared to many studies of ASDs, the ASD-low MA subgroup was small ( $n=25$ ). This caused the power for some analyses to be limited.

Additionally, across all three groups, a large minority (44%) of children who received a Time 1 evaluation declined, or were unable to be contacted for, a Time 2 evaluation. Differential attrition analyses indicated that children who returned for follow-up were more likely to be male and, perhaps more significantly, White, than those who did not return at Time 2. As a result, our study may not fully capture the developmental patterns shown in children from racial and ethnic minority backgrounds, or those of lower income. Further, it is quite possible that the parents of non-returning children had fewer concerns about their child's development at the time a follow-up evaluation was offered, when compared to those that did return at Time 2. 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 an ASD at a follow-up evaluation. Of note, compared to children initially diagnosed with AD and PDD-NOS, a smaller percentage (29%) of those in the ASD-low MA group failed to return for follow-up, perhaps suggesting that parents of these children were *more* concerned, which may be a product of this group's combined deficits in autism symptoms and cognitive functioning.

Another limitation to consider is the time frame of the study. Our longitudinal study was limited to data collected at ages 2 and 4 years. As a result, we do not have potentially informative data regarding long-term outcomes for these children. Furthermore, given the time frame of the recruitment years, DSM-IV-TR criteria was used to diagnose participating children. Since ICD-10 criteria are similar to those of DSM-IV-TR, the results of a parallel study using ICD-10 are likely to be comparable to the current results. However, our findings may not be generalizable to a study using DSM-5.

Finally, we were unable to account for the likely impact of intervention on the diagnostic and developmental outcomes of children in the current study. Through participation in the larger study, each participant in our sample received an autism spectrum diagnosis at a relatively young age (i.e., on average, around age 2 years) and was subsequently referred for intensive autism-specific intervention services. At follow-up, parents were asked to report on a developmental history form whether or not their child had received some type of intervention, but specific information regarding the type and structure of services, as well as intensity, frequency, and duration, was either unavailable or not detailed enough to allow for meaningful evaluation

of the impact of these services on the child's progress over time. Therefore, we were unable to directly assess if a child's improvement across time was related to the intervention services they received between evaluations, increased parental understanding of their child's deficits, or any other factors that may have contributed to a child's progress.

### Future Directions

These results support the position that ASD-low MA is a distinct autism subtype, insofar as its developmental growth and diagnostic stability are consistently different than AD and PDD-NOS. The developmental progress rates and autism symptom improvement in the AD and PDD-NOS groups suggest that, despite their delays and symptomatology at initial evaluation, the intensive intervention services that these groups likely received after their initial evaluation helped to address these delays and provided an opportunity to make improvements. However, the very low rate of developmental progress over time, and the stability of autism spectrum symptomatology, in the ASD-low MA group signal the severity of this particular subtype, which may merit more intensive services. These findings may also suggest that children with low mental age may struggle to engage with traditional autism-specific intervention services, thus limiting opportunities for developmental gains.

There is a clear need for replication of these results to ensure the reliability and validity of diagnosing ASDs in the presence of low mental age. If replicated, such findings would allow clinicians greater confidence in diagnosing ASDs regardless of mental age at the time of evaluation. Longitudinal studies should also follow these children into later childhood, when they are more likely to receive a comorbid diagnosis of ID (i.e., at age 6 years and older),

to examine whether their autism symptoms continue to remain stable over time.

**Acknowledgments** The authors wish to thank the families and clinicians who participated in the current study, as well as the Early Detection Study research team at the University of Connecticut for their assistance with data collection.

**Funding** This study was funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (Grant number R01HD039961) and the Maternal and Child Health Bureau (Grant number R40MC00270).

**Author Contributions** AH participated in conception of the study design, data collection, performance of statistical analyses, and drafting of the manuscript; LM participated in performance of statistical analyses, interpretation of the data, and revising of the manuscript; DF participated in conception of the study design, data collection, interpretation of the data, and revising of the manuscript. All authors read and approved the final manuscript.

### Compliance with Ethical Standards

**Conflict of interest** Deborah Fein is part owner of M-CHAT-R, LLC, which receives royalties from companies that incorporate the M-CHAT-R into commercial products and charge for its use. Data reported in the current paper is from the freely available paper versions of the M-CHAT and M-CHAT-R. Alexander Hinnebusch and Lauren Miller declare that they have no conflicts of interest.

**Ethical Approval** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Informed Consent** Informed consent was obtained from all individual participants included in the study.

## Appendix

### *Diagnostic Criteria*

#### **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 years

**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 lack of interest in peers

**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 and **Vineland** scores on Communication and Social subdomains are ALL less than or equal to 12 months age equivalent

#### **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** (peer relationships) 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

**Developmental Delay**

\_\_\_\_\_ Delay of at least 1.5 standard deviations on AT LEAST ONE of the following (“non-language”):

\_\_\_\_\_ **Mullen** Visual Reception (T-score=35 or less)

\_\_\_\_\_ **Mullen** Fine Motor (T-score=35 or less)

\_\_\_\_\_ **Vineland** Motor Skills (SS=77 or less)

**AND**

\_\_\_\_\_ Delay of at least 1.5 standard deviations on AT LEAST ONE of the following (“Language”):

\_\_\_\_\_ **Mullen** Expressive Language (T-score=35 or less)

\_\_\_\_\_ **Mullen** Receptive Language (T-score=35 or less)

\_\_\_\_\_ **Vineland** Communication (SS=77 or less)

**AND**

\_\_\_\_\_ At least one from the 2 categories above must be a delay on the **Mullen**

**Other Diagnosis**

\_\_\_\_\_ DOES NOT meet criteria for an ASD diagnosis

**AND**

Meets criteria for a non-ASD DSM-IV-TR diagnosis, such as ADHD, ODD, Reactive Attachment Disorder, an anxiety disorder, a mood disorder, or some other psychological condition: Specify: \_\_\_\_\_

**No Diagnosis**

Meets criteria for a non-DSM-IV-TR disorder or condition (e.g., sensory processing difficulties): Specify: \_\_\_\_\_

**OR**

DOES NOT meet criteria for any DSM-IV-TR or study diagnoses, but DOES NOT meet criteria for Typical Development (*include children who cannot complete reliable/valid testing due to non-compliance here*)

**AND**

Child displays sub-syndromic characteristics of a condition (*reserve for clinically significant motor and/or speech-language difficulties that do not meet diagnostic criteria but present as clinically significant*)

**References**

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- Anderson, D. K., Liang, J. W., & Lord, C. (2014). Predicting young adult outcome among more and less cognitively able individuals with autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, 55(5), 485–494.
- Bagnato, S. J., & Neisworth, J. T. (1980). The intervention efficiency index: An approach to preschool program accountability. *Exceptional Children*, 46(4), 264–269.
- Bagnato, S. J., & Neisworth, J. T. (1985). Efficacy of interdisciplinary assessment and treatment for infants and preschoolers with congenital and acquired brain injury. *Analysis and Intervention in Developmental Disabilities*, 5, 107–128.
- 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.
- 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.
- 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*. Storrs: University of Connecticut.
- Chakrabarti, S., & Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *Journal of the American Medical Association*, 285(24), 3093–3099.
- 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, D. J., Paul, R., & Volkmar, F. R. (1986). Issues in the classification of pervasive and other developmental disorders: Toward DSM-IV. *Journal of the American Academy of Child Psychiatry*, 25, 213–220.
- DiLavore, P. C., Lord, C., & Rutter, M. (1995). The pre-linguistic autism diagnostic observation schedule. *Journal of Autism and Developmental Disorders*, 25(4), 355–379.
- Eikeseth, S., Klintwall, L., Jahr, E., & Karlsson, P. (2012). Outcome for children with autism receiving early and intensive behavioral intervention in mainstream preschool and kindergarten settings. *Research in Autism Spectrum Disorders*, 6, 829–835.
- Fombonne, E. (2003). The prevalence of autism. *Journal of the American Medical Association*, 289(1), 87–89.
- 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.
- Howard, J., Sparkman, C., Cohen, H., Green, G., & Stanislaw, H. (2005). A comparison of intensive behavior analytic and eclectic treatments for young children with autism. *Research in Developmental Disabilities*, 26, 359–383.
- 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.
- Kanne, S. M., Gerber, A. J., Quirnbach, 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.
- Klin, A., Carter, A. S., & Sparrow, S. S. (1997). Psychological assessment of children with autism. In D. J. Cohen & F. R. Volkmar (Eds.), *Handbook of autism and pervasive developmental disorders* (2nd ed.) (pp. 418–427). New York: Wiley.
- Klintwall, L., Eldevik, S., & Eikeseth, S. (2015). Narrowing the gap: Effects of intervention on developmental trajectories in autism. *Autism: The International Journal of Research and Practice*, 19(1), 53–63.
- Lord, C. (1995). Follow-up of 2-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*, (pp. 15–17). West Hartford: International Meeting for Autism Research.
- Matson, J. L., & Shoemaker, M. (2009). Intellectual disability and its relationship to autism spectrum disorders. *Research in Developmental Disabilities*, 30(6), 1107–1114.
- 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: The International Journal of Research and Practice*, 7(1), 47–63.
- Mullen, E. M. (1995). *Mullen scales of early learning*. Circle Pines, MN: American Guidance Service, Inc.
- Oliver, C., Petty, J., Ruddick, L., & Bacarese-Hamilton, M. (2012). The association between repetitive, self-injurious and aggressive behavior in children with severe intellectual disability. *Journal of Autism and Developmental Disorders*, 42(6), 910–919.
- 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.
- Reznick, J. S., Baranek, G. T., Reavis, S., et al. (2007). A parent-report instrument for identifying 1-year-olds at risk for an eventual diagnosis of autism: The first year inventory. *Journal of Autism and Developmental Disorders*, 37, 1691–1710.
- Robins, D. L., Casagrande, K., Barton, M., Chen, C.-M. A., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the modified checklist for autism in toddlers, revised with follow-up (M-CHAT-R/F). *Pediatrics*, 133(1), 37–45.
- Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The modified checklist for autism in toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 31(2), 131–144.
- Rutter, M., Le Couteur, A., Lord, C., & Faggioli, R. (2005). *ADI-R: Autism diagnostic interview–revised: Manual. OS, Organizzazioni speciali*. Torrance: Western Psychological Services
- Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: 4-year outcome and predictors. *American Journal on Mental Retardation*, 110(6), 417–438.
- 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.
- Schopler, E., Short, A., & Mesibov, G. (1989). Relation of behavioral treatment to “normal functioning”: Comment on Lovaas. *Journal of Consulting and Clinical Psychology*, 57, 162–164.
- Sevin, J. A., Matson, J. L., Williams, D., & Kirkpatrick-Sanchez, S. (1995). Reliability of emotional problems with the diagnostic assessment for the severely handicapped (DASH). *British Journal of Clinical Psychology*, 34, 93–94.
- Shonkoff, J. P., Hauser-Cram, P., Krauss, M. W., & Upshur, C. C. (1992). Development of infants with disabilities and their families. *Monographs of the Society for Research in Child Development*, 57(6), 230.
- 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*, 64(1), 1–114.
- Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *The Vineland adaptive behavior scales–interview edition*. Circle Pines, MN: American Guidance Service.
- Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland adaptive behavior scales* (2nd edn.). Circle Pines, MN: American Guidance Services.
- 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(2), 219–226.
- Stone, W. L., McMahon, C. R., & Henderson, L. M. (2008). Use of the screening tool for autism in 2-Year-olds (STAT) for children under 24 months: An exploratory study. *Autism: The International Journal of Research and Practice*, 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.
- Swinkels, S. H., Dietz, C., Van Daalen, E., et al. (2006). Screening for autistic spectrum in children aged 14 to 15 months: The development of the early screening of autistic traits questionnaire (ESAT). *Journal of Autism and Developmental Disorders*, 36(6), 723–732.
- Wetherby, A. M., Brosnan-Maddox, S., Peace, V., et al. (2008). Validation of the infant-toddler checklist as a broadband screener for autism spectrum disorders from 9 to 24 months of age. *Autism: The International Journal of Research and Practice*, 12(5), 487–511.
- Wolery, M., & Dyk, L. (1985). The evaluation of two levels of a center based early intervention project. *Topics in Early Childhood Special Education*, 5(2), 66–77.