ORIGINAL PAPER

Support for a Dimensional View of Autism Spectrum Disorders in Toddlers

**Lisa D. Wiggins · Diana L. Robins ·
Lauren B. Adamson · Roger Bakeman ·
Christopher C. Henrich**

Abstract We examined whether clinically distinct subgroups can be derived from a sample of toddlers ($n = 186$) who failed the Modified Checklist for Autism in Toddlers, received a comprehensive clinical evaluation, and were diagnosed with an autism spectrum disorder (ASD). Three subgroups emerged from cluster analysis distinguished by (a) social, communication, and intellectual skills and (b) the rate and intensity of repetitive behaviors and abnormal sensory response. Preoccupations, compulsions, and rituals did not distinguish resultant subgroups. These results support a dimensional diagnostic view of ASDs in toddlers since subgroup differences were based on symptom severity rather than different symptom profiles. Results also identify specific types and levels of behavioral deficit relevant to toddler populations. Implications for early diagnosis are discussed.

Keywords Autism spectrum disorders (ASDs) · Early identification · Early diagnosis · Cluster analysis

L. D. Wiggins · D. L. Robins · L. B. Adamson · R. Bakeman · C. C. Henrich
Georgia State University, Atlanta, GA, USA

Present Address:
L. D. Wiggins (✉)
NCBDDD/CDC, 1600 Clifton Road MS E-86, Atlanta, GA 30333, USA
e-mail: lwiggins@cdc.gov

Introduction

Children with autism spectrum disorders (ASDs) comprise a heterogeneous group that shows diverse levels of social, communication, behavioral, and intellectual development. Consequently, attempts to discover common features of ASDs essential for categorical classification have faced many challenges. Current diagnostic schemes typically recognize three distinct diagnoses within the class of ASDs: (a) Autistic disorder, (b) Asperger's disorder, and (c) Pervasive developmental disorder-not otherwise specified (PDD-NOS; American Psychiatric Association 1994); although a more dimensional classification system is being considered in revised diagnostic manuals (American Psychiatric Association 2009). Yet it is unclear whether a categorical or dimensional view of ASDs is more appropriate in toddler populations and, if clinically distinct categories can be derived; whether these categories can delineate etiology, trajectory, and treatment options for the young child (Fein et al. 1999). If clinically distinct categories of ASDs do exist, identification of symptoms that differentiate categories in the first few years of life can inform diagnostic practices and enhance knowledge of early manifestations of the disorders and factors that influence developmental course.

Past research suggests that as many as four categories of ASDs can be empirically derived, but that level of symptom severity is primarily responsible for distinguishing resultant subgroups (see Table 1 for select cluster analytic studies). Specifically, the degree of impairment in social, communication, and intellectual abilities and the presence of stereotyped interests and behaviors (SIB) have been found to be important factors that define ASD subgroups (Eaves et al. 1994; Fein et al. 1999; Sevin et al. 1995; Siegel et al. 1986; Stevens et al. 2000). Cluster analysis,

Table 1 Published cluster analyses on persons with autism spectrum disorders

Primary author	Year published	Sample size and age	Sample diagnoses	Primary independent variables	Clusters identified	Cluster descriptions
Siegel	1986	$n = 46$ 4–20 years, $M = 10.7$ years	AD	Developmental history questionnaire, parent report of current behavior, behaviors coded from a videotaped diagnostic evaluation	4	1. Classic autism 2. Severe intellectual disability 3. Schizotypal personality traits 4. Anxious/negativistic behaviors
Eaves	1994	$n = 166$ 3–12 years, $M = 7.1$ years	AD, ASD, ASP, AT, HFA	Behaviors coded retrospectively from hospital charts	4	1. Typically autistic 2. Low-functioning autism 3. High-functioning autism 4. Hard-to-diagnose autism
Sevin	1995	$n = 34$ 2–22 years, $M = 7.7$ years	AD, ASD	Childhood autism rating scale, autism behavior checklist, ritvo-freeman real life rating scale, Vineland adaptive behavior scales	4	1. Severe autism 2. Moderate autism 3. Mild autism 4. High-functioning autism
Prior	1998	$n = 135$ 3–21 years, $M = 10.22$ years	AD, ASD, HFA (only high functioning children included)	Autistic disorders checklist, peabody picture vocabulary test	3	1. Autism 2. ASD 3. Mild ASD
Fein	1999	$n = 194$ 3–7 years, $M = 4.9$ for ASD	ASD, Developmental Language Disorder, low-IQ	Autistic disorders checklist, peabody picture vocabulary test, Vineland adaptive behavior scales	2	1. Low-functioning autism 2. High-functioning autism
Stevens	2000	$n = 138$ 7–9 years, $M = 8.6$ years	ASD	Autistic disorders checklist, peabody picture vocabulary test, Vineland adaptive behavior scales	2	1. Low-functioning autism 2. High-functioning autism

AD autistic disorder, including infantile autism and residual state autism, *ASD* autism spectrum disorder or pervasive developmental disorder, *ASP* Asperger's syndrome, *AT* autistic tendencies, *HFA* high functioning autism

which relies on the partitioning of data into homogeneous groups, has been used to identify subgroups of ASDs in older populations. One of the earliest cluster analytic studies of ASDs found four clinically distinct subgroups defined by classic autism, severe intellectual disability, schizotypal personality traits, and anxious/negativistic behaviors (Siegel et al. 1986). A 4-cluster solution was also found by Eaves et al. (1994) and Sevin et al. (1995); both studies found subgroups of children who had low-functioning autism, high-functioning autism, moderate or typical autism, and mild or hard-to-diagnose autism. But only two ASD subgroups of low-functioning autism and high-functioning autism were found when Fein et al. (1999) limited their cluster analysis to 633 preschool children with delayed or deviant communication. The significance of this latter study is that fewer subgroups of children with ASDs were found when the focus was narrowed to a preschool population. This finding suggests that the course of ASDs may become more heterogeneous as children age and that distinct variables may predict level of severity of ASDs in young children.

Yet all of the aforementioned studies focused on older children and adults who present with distinctly different symptom sets than toddlers with ASDs. For instance, some researchers propose a 2-factor model of SIB that consists of "lower-order" sensorimotor behaviors and "higher-order" cognitive rigidity; lower-order sensorimotor behaviors may occur more often in younger samples than older samples and higher-order cognitive rigidity may occur more often in older samples than younger samples (Szatmari et al. 2006; Richler et al. 2007, 2010). However, there are only a few published studies on these factors of SIB in very young children with ASDs. In one report, Moore and Goodson (2003) found that parents of toddlers with ASDs reported more impairment in lower-order SIB than higher-order SIB. These results were further supported by Richler et al. (2007, 2010) who found that parents of toddlers with ASDs reported more lower-order behaviors than parents of children with other delays or typical development; higher-order SIB did not distinguish study groups. A decreased frequency of higher-order SIB in toddlers is not surprising given they are positively correlated with nonverbal abilities

(Bishop et al. 2006; Lord et al. 2006) and most children with ASDs identified in the first few years of life have below average nonverbal skills.

It is not clear from previous research whether lower-order SIB occur at the rate or intensity in toddlers to meet diagnostic classification that requires clinically significant impairment in social, communication, *and* behavioral domains. Moreover, there are other limitations of previous analyses that deserve consideration. First, the youngest mean age of children studied in past subgroup reports was almost 5 years (Fein et al. 1999) and there are no published studies on categories of ASDs in children younger than 5 years of age. Second, many studies used unstandardized or unpublished measures designed solely for the purpose of cluster analysis or measures that are not routinely used in clinical or research practice (Eaves et al. 1994; Fein et al. 1999; Prior et al. 1998; Siegel et al. 1986; Stevens et al. 2000); using items on "gold standard" measures as cluster and validation variables may lend more credence to resultant subgroups. Third, there are no published studies that examine subgroup differences in lower-order versus higher-order SIB to determine whether some groups of toddlers with ASDs would be missed by classification systems that require impairment in all three diagnostic domains. Finally, there are no published studies that examine characteristics of early ASD subgroups that can predict later ASD diagnosis.

Given the aforementioned limitations, the primary purpose of our study was to examine whether empirically derived subgroups could be derived from a sample of toddlers with ASDs and whether resultant subgroups would be based on level of ASD severity or different symptom profiles. Based on past research, it was hypothesized that we would find 2–3 ASD subgroups distinguished by level of impairment in social, verbal, and nonverbal abilities and the presence of SIB. We thought lower-order SIB would distinguish ASD subgroups in our sample of toddlers whereas higher-order SIB would not, suggesting that a 2-factor model of SIB may not emerge until after the toddler years. We also thought many of the toddlers with ASD in our sample would show clinically significant social and communication deficit but not clinically significant SIB and that this group of children would be more likely to lose their ASD diagnoses 2 years later.

Methods

Participants

Participants were retrospectively identified from two early screening studies at the University of Connecticut (UConn) and Georgia State University (GSU) that prospectively identified young children with ASDs through screening in general pediatric practices and early intervention programs. Specifically, families of participants who provided written informed consent were administered the Modified Checklist for Autism in Toddlers (M-CHAT; Robins et al. 1999a) during a routine 18- or 24-month well child visit or a visit to a state-wide early intervention program that serves children from birth to 36 months of age. The M-CHAT is a short parent-report checklist designed to detect risk for ASDs in very young children. A child screens positive on the M-CHAT when any three of 23 items are failed, or any two of six critical items are failed. Critical items were identified by empirical methods using discriminant function analysis. The most current estimate of M-CHAT sensitivity suggests an upper bound of .91; which corroborates the original validation study (Kleinman et al. 2008; Robins et al. 2001).

If M-CHAT results indicated risk for an ASD (i.e., screen positive), a member of the study team called the family to administer the M-CHAT Follow-up Interview (Robins et al. 1999b) to clarify responses and elicit examples of target behaviors. If risk for ASDs was still indicated after the M-CHAT Follow-up Interview, the family was invited for a free, comprehensive clinical evaluation. Three hundred children screened positive on the M-CHAT and Interview and received a comprehensive clinical evaluation. Our sample consisted of the subsample of 186 toddlers who were diagnosed with an ASD after the clinical evaluation. Mean age at evaluation was 26 months (range = 13–37 months; SD = 5 months). The racial make-up of the sample was 88% White, 4% Black, 4% Hispanic (including Puerto Rican), 2% Asian, and 2% "other" ($n = 112$). The sample was 80% male and 20% female. The average cognitive standard score yielded at the clinical evaluation was 61 ($n = 173$; range = 49–127; SD = 16), reflecting mild intellectual disability.

After the clinical evaluation, 113 children were diagnosed with Autistic Disorder, 72 were diagnosed with PDD-NOS, and one was diagnosed with Asperger's Disorder. One hundred thirty six of the 186 children in the sample were evaluated again around 4 years of age (mean = 53 months; range = 41–79 months; SD = 7 months). The same measures used in toddler evaluations were repeated at 4-years evaluations. After the clinical evaluation around 4 years of age, 79 children were diagnosed with Autistic Disorder, 34 were diagnosed with PDD-NOS, and 23 were not diagnosed with an ASD. See Table 2 for a crosstab of 2- and 4-years diagnoses.

Measures

The *Autism Diagnostic Interview-Revised* (Lord et al. 1994) is a semi-structured, parent interview used to classify children with a mental age of ≥24 months as autism or no

Table 2 Crosstabs between 4-year diagnoses and 2-year diagnoses and 2-year cluster membership

	4-Year diagnoses		
	NonASD	PDD-NOS	Autistic disorder
2-Year diagnoses			
Asperger's disorder	0	1	0
PDD-NOS	14	18	13
Autistic disorder	9	15	66
Totals (*n*)	23	34	79
2-Year cluster membership			
Cluster 1	10	16	14
Cluster 2	6	11	15
Cluster 3	7	7	50
Totals (*n*)	23	34	79

autism; the ADI-R does not classify children with other ASDs. The ADI-R gathers comprehensive information about the child from a parent in three domains of development: social, communication, and SIB. Individual items are scored as 0, 1, or 2 on the diagnostic algorithm. It is important to note that the ADI-R is often used in clinical and research practice with very young children because of lack of other appropriate measures. In response to this dilemma, the authors of the ADI-R have created a toddler version that is currently being field tested and was used in a portion of the current sample. The diagnostic algorithm for the toddler version is an exact replica of the diagnostic algorithm of the ADI-R (although different items are included in the broader interview). Furthermore, criteria for scoring and determining autism classification are the same. Therefore, both versions of the instrument will be called the ADI-R throughout this report.

The *Autism Diagnostic Observation Schedule* (Lord et al. 1999) is a standardized observation of a child that tries to elicit social interaction and communication using structured play activities. The examiner implements the module that best corresponds to the child's expressive language level in order to prevent language aptitude from impeding accurate classification. Most children in this study were administered Module 1, designed for children who are not regularly using phrase speech. ASD classification, subsequently referred to as the ADOS total score, is determined by scores on a subset of items from the social and communication domains. The algorithm page also includes SIB and play items, although they are not considered for ASD classification. Individual items are scored as 0, 1, or 2 on the algorithm page.

New ADOS algorithms have been proposed and are currently being validated in different samples of children (Gotham et al. 2007; Oosterling et al. 2010). The new ADOS algorithms combine items from the former social

and communication domains to derive a social affect total score and combine items from the former social and behavioral domains to derive a restricted and repetitive behavior total score. ASD classification is based on resultant scores from both the social affect and restricted and repetitive behavior domains. Results from both the former ADOS algorithms and revised ADOS algorithms will be presented in this analysis.

The *Childhood Autism Rating Scale* (Schopler et al. 1988) is a standardized observation instrument used to help diagnose ASDs in young children; parent report can also be considered during scoring. The CARS rates children suspected of having an ASD on 15 items that include social and communication skills and SIB. Individual items are scored on a 7-point Likert scale rated from one to four in half-point increments. The final diagnostic algorithm represents a sum of item scores and classifies the child as having severe autism, mild-moderate autism, or no autism indicated; a cut-off score of 30 is needed to be classified as having an ASD. Previous analyses on a subsample of children included in this study found that inter-rater reliability for the CARS total score was .94 (Chlebowski et al. 2010).

The *Mullen Scales of Early Learning* (MSEL, Mullen 1995) is a standardized measure of cognition appropriate for children from birth to 68 months of age. The examiner presents a series of tasks created to measure gross motor, fine motor, expressive language, receptive language, and visual reception skills. Raw scores can be converted to t-scores, percentile ranks, and age equivalents. An early learning composite, created from all domains except gross motor, is also provided.

The *Vineland Adaptive Behavior Scales* (VABS, Sparrow et al. 1984, 2005) is a semi-structured parent interview that assesses personal and social sufficiency in individuals from birth to 18 years. The VABS assesses four domains of adaptive behavior: communication, daily living skills, socialization, and motor abilities. Raw scores can be converted to standard scores, percentile ranks, and age equivalents. An adaptive behavior composite, created from all domains, is also provided. It is important to note that 4% of the sample received the VABS-II (revised edition; Sparrow et al. 2005). The VABS-II is similar to the VABS but offers updated norms, an expanded age range, updated item content, and revised interview format. Correlations between the VABS and VABS-II range from .65 to .91 for children 0–2 years of age.

Procedures

Families of children who screened positive on the M-CHAT and subsequent M-CHAT Follow-up Interview and agreed to participate in the study were scheduled for a

clinical evaluation. The clinical evaluation took place at the UConn Psychology Clinic, the GSU Psychology Clinic, the child's home, or the early intervention provider's site. Evaluations consisted of the ADI-R, ADOS, CARS, MSEL, and VABS. All clinicians had prior experience with the diagnostic measures before study administration and clinicians who administered the ADI-R and ADOS had established research reliability. After the evaluation was complete, clinicians immediately scored the instruments, discussed evaluation results, and provided feedback to the family. A licensed clinical psychologist or developmental pediatrician provided ASD diagnoses after careful review of all available data and completion of a DSM-IV checklist that supported an ASD diagnosis. Scores on each of the autism diagnostic instruments informed clinical diagnosis although ASD cut-off criteria on the ADI-R, ADOS, and CARS were not required for a clinical diagnosis (so four children with sub-threshold scores on the ADOS, ADI-R, and CARS were diagnosed with ASD and included in the sample; review of participant data confirms these children met criteria for PDD-NOS at the time of evaluation and two of these three children who were re-evaluated around 4 years of age still met criteria for PDD-NOS). A comprehensive evaluation report was mailed to the families within 6 weeks of the clinical evaluation. All families were invited to receive another comprehensive evaluation using the same measures around the child's fourth birthday.

Data Analyses

Ward's cluster analysis was used to identify empirically derived subgroups of toddlers with an ASD. We chose cluster analysis as our analytic method since we wanted to generate empirically derived and homogeneous groups of toddlers with ASD, and cluster analysis identifies children with similar behavioral profiles given performance on clinical evaluation measures. The standardized instrument chosen for cluster analysis was the CARS. Individual items from the CARS were chosen as cluster variables because the CARS was associated with the highest agreement with clinical judgment when used in a sub-set of toddlers from this sample (Ventola et al. 2006; Wiggins and Robins 2008). Further, the CARS has a broad range of items that may be important in defining subgroups of toddlers with ASDs and CARS items are rated on a 7-point scale, which provides a broader range of scores than other diagnostic instruments (such as the typical 3-point range found on the ADOS and ADI-R). Items from other diagnostic, cognitive, or adaptive measures were not used as cluster variables since these items were used to validate the cluster solution and were used as dependent variables in subsequent analyses.

Discriminant function analysis (DFA) of CARS items was used to identify CARS functions that best defined resultant subgroups and the amount of variance in cluster membership accounted for by each of these functions. Clusters were then validated by assessing mean differences between subgroups on MSEL domain scores, VABS domain scores, and ADI-R domain scores; the Bonferroni correction was applied to adjust for multiple comparisons. Validation analyses were conducted to assess whether resultant subgroups differed in terms of symptom profiles or level of severity on measures that were not used to generate the cluster solution but still had relevance to the presentation of ASDs in toddlers. Validation analyses will primarily focus on MSEL, VABS, and ADI-R comparisons. Validation results from the former and revised ADOS domains will only be mentioned briefly in text (and excluded from tables) since the ADOS was based on the same behavioral sample as the CARS. MSEL age equivalents were analyzed instead of t-scores because of common floor effects produced on this measure.

Separate ANOVAs were also performed to determine subgroup differences on individual SIB item scores included on the ADI-R and ADOS diagnostic algorithms in order to determine if resultant subgroups differed on additional SIB than those included as cluster variables (i.e., CARS items); the Bonferroni correction was again applied to adjust for multiple comparisons. Multinomial logistic regression was conducted to examine how ASD subtype membership around 2 years predicted ASD diagnosis around 4 years. Specifically, clinical diagnosis around 4 years of age was coded into three categories: nonASD, PDD-NOS, and Autistic Disorder. These categories were entered as the dependent variable and ASD cluster membership was entered as the independent variable. The reference category for 4 years diagnoses was Autistic Disorder. (B) coefficients were interpreted as odds-ratios, which are differences in the odds likelihood of membership in various diagnostic groups.

Results

Subgroups of Toddlers with ASD

Ward's cluster analysis revealed three clusters of toddlers with ASDs: Cluster 1 consisted of 47 children, Cluster 2 consisted of 44 children, and Cluster 3 consisted of 95 children. Of the 47 children in Cluster 1, 35 (74%) were diagnosed with PDD-NOS or Asperger's Disorder (ASP) and 12 (26%) were diagnosed with Autistic Disorder; of the 44 children in Cluster 2, 22 (50%) were diagnosed with PDD-NOS and 22 (50%) were diagnosed with Autistic Disorder; and of the 95 children in Cluster 3, 16 (17%)

were diagnosed with PDD-NOS and 79 (83%) were diagnosed with Autistic Disorder. There were no significant age or sex differences between cluster subgroups. There was a significant difference between cluster subgroups in total MSEL standard scores in that Cluster 1 ($M = 71$) performed better than Cluster 2 ($M = 61$) or Cluster 3 ($M = 56$), F (2, 171) = 17.20, $p < .01$.

A DFA of CARS items was performed to identify functions that best defined cluster subgroups and the amount of variance in cluster membership accounted for by each function. The DFA found that two discriminant functions were significant in distinguishing subgroups, Wilks' lambda = .15, χ^2 (30, $n = 186$) = 331.91, $p < .00$ for the first function and Wilks' lambda = .55, χ^2 (14, $n = 186$) = 105.06, $p < .00$ for the second function. The first function accounted for 76% of the variance and the second function accounted for 24% of the variance. The first function was labeled by the authors as "social and communication skills" and the second function was labeled by the authors as "SIB." CARS items included in the first function were verbal communication, emotional response, imitation, nonverbal communication, and relating to people; CARS items included in the second function were object use, body use, and sensory response (see Table 3 for a list of all CARS items as they pertain to each function).

Differences Between Subgroups of Toddlers with ASD

The three subgroups derived from cluster analysis were next compared on MSEL, VABS, and ADI-R domain

Table 3 Structure coefficient (SC) and discriminant function coefficients (DFC) for cluster variables

	Function 1 Social communication SC (DFC)	Function 2 SIB SC (DFC)
Imitation	.62 (.36)	
General impressions	.52 (.55)	
Relating to people	.45 (.30)	
Nonverbal communication	.41 (.35)	
Verbal communication	.41 (.53)	
Listening response	.35 (.07)	
Intellectual response	.29 (.26)	
Emotional response	.29 (.40)	
Visual response	.26 (.14)	
Activity level		.50 (.51)
Object use		.41 (.35)
Fear		.40 (.52)
Resistance to change		.37 (.08)
Body use		.26 (.13)
Sensory response		.21 (.09)

scores to determine how these subgroups differed in cognitive, adaptive, and autism-specific domains. ASD subgroups differed on all MSEL, VABS, and ADI-R domains, except the MSEL motor domain, VABS motor domain, and ADI-R SIB domain (see Table 4). Table 4 shows the first cluster had more communication abilities than the second or third cluster; the first cluster performed significantly better than both the second and third cluster on the VABS and ADI-R communication domains. The first cluster also had more social abilities than the second or third cluster, although group differences only reached statistical significance between the first and third clusters on the VABS and ADI-R social domains. ADOS analyses supports these results in that the first and second cluster showed significantly less impairment on the former social, F (2, 137) = 34.03, $p < .01$; communication, F (2, 137) = 25.38, $p < .01$; and SIB domains, F (2, 137) = 12.69, $p < .01$. The first and second cluster also showed significantly less impairment on the revised social affect, F (2, 137) = 14.10, $p < .01$ and restricted and repetitive behavior domains, F (2, 171) = 17.20, $p < .01$, indicating consistent findings across the old and new ADOS algorithms. Therefore, given results of the DFA and ANOVA analyses, cluster subgroups were distinguished by level of social, communication, and intellectual abilities and the rate and intensity of SIB. Consequently, the first cluster subgroup, which was characterized by relatively few social and communication deficits, few SIB, and low-average intellectual abilities, was labeled "ASD, mild impairment." The second cluster subgroup, which was characterized by many social and communication deficits, few SIB, and mild intellectual disability, was labeled "ASD, moderate impairment." The third cluster subgroup, which was characterized by many social and communication deficits, many SIB, and mild-moderate intellectual disability, was labeled "ASD, severe impairment."

Cluster subgroups were next compared on ADOS and ADI-R SIB algorithm items appropriate for toddlers (i.e., all algorithm items except ADI-R "circumscribed interests," which is only appropriate for children 36 months and older) to further classify subgroups and offer additional validation of cluster labels. There were no significant group differences in unusual preoccupations, verbal rituals, compulsions and rituals, and repetitive interests; the majority of these higher-order SIB were found in the ADI-R behavioral domain and few children exhibited such behaviors (Table 5). There were also no significant group differences in hand and finger or other complex body mannerisms on either diagnostic instrument (Table 5). There were, however, significant group differences in repetitive behaviors and abnormal sensory response on both the ADI-R and the ADOS (Table 5). In these analyses, the subgroup labeled "ASD, severe impairment" had

Table 4 Toddler autism spectrum disorder subgroup differences in general developmental and autism-specific domains

	Cluster 1 ($n = 47$) M (SD)	Cluster 2 ($n = 44$) M (SD)	Cluster 3 ($n = 95$) M (SD)	F	p	η^2
CARS total score	27.3 (3.64)$_a$	30.4 (1.73)$_b$	36.8 (3.43)$_c$	160	<.01	.64
MSEL expressive language age equivalent	19.9 (9.25)$_a$	15.1 (9.16)$_{a,b}$	13.1 (10.1)$_b$	6.43	<.01	.08
MSEL receptive language age equivalent	18.1 (10.2)$_a$	14.2 (9.72)$_{a,b}$	12.7 (11.8)$_b$	3.37	.04	.04
MSEL visual reception age equivalent	21.8 (8.89)$_a$	19.3 (9.76)$_a$	14.7 (4.80)$_b$	11.6	<.01	.14
MSEL fine motor age equivalent	23.0 (10.2)$_a$	21.4 (8.81)$_a$	19.9 (11.7)$_a$	1.18	.31	.02
VABS communication standard score	75.6 (14.3)$_a$	66.5 (5.28)$_b$	63.1 (6.59)$_b$	24.8	<.01	.22
VABS socialization standard score	73.7 (9.35)$_a$	69.8 (6.36)$_{a,b}$	66.6 (8.94)$_b$	10.7	<.01	.12
VABS daily living standard score	74.1 (12.3)$_a$	69.3 (6.86)$_{a,b}$	68.6 (10.5)$_b$	4.51	.01	.05
VABS motor standard score	83.7 (11.0)$_a$	84.7 (10.3)$_a$	81.3 (13.1)$_a$	1.39	.25	.02
ADI-R socialization total score	12.1 (4.19)$_a$	14.0 (3.31)$_{a,b}$	15.7 (3.63)$_b$	11.3	<.01	.15
ADI-R communication total score	8.41 (3.55)$_a$	10.3 (2.42)$_b$	11.2 (2.88)$_b$	10.5	<.01	.14
ADI-R SIB total score	1.81 (1.98)$_a$	2.30 (1.92)$_a$	2.81 (2.16)$_a$	2.39	.09	.04

Group differences are indicated by subscripts and reflect a Bonferroni adjustment for multiple comparisons

Table 5 Toddler autism spectrum disorder subgroup differences on stereotyped interests and behavior (SIB) items found on diagnostic instruments

	Cluster 1 ($n = 47$) M (SD)	Cluster 2 ($n = 44$) M (SD)	Cluster 3 ($n = 95$) M (SD)	F	p	η^2
SIB items on the ADI-R						
Unusual preoccupations	.64 (.87)$_a$	.95 (.95)$_a$	.98 (1.52)$_a$	.69	.50	.01
Verbal rituals	.46 (.98)$_a$	.13 (.50)$_a$	.08 (.50)$_a$	2.30	.11	.06
Compulsions/rituals	.54 (.88)$_a$	.09 (.29)$_a$	.49 (1.38)$_a$	1.27	.29	.03
Hand and finger mannerisms	.71 (.85)$_a$	.96 (1.02)$_a$	1.24 (.93)$_a$	3.00	.06	.06
Other complex mannerisms	1.16 (1.28)$_a$	.83 (.94)$_a$	1.07 (1.24)$_a$	.57	.57	.01
Repetitive use of objects	.65 (.89)$_a$	.96 (.88)$_a$	1.72 (1.94)$_b$	4.74	.01	.09
Unusual sensory interests	.62 (.70)$_a$	.96 (.71)$_a$	1.23 (.70)$_b$	6.60	<.01	.13
SIB items on the ADOS						
Unusual sensory interests	.55 (.69)$_a$	.72 (.82)$_a$	1.24 (.80)$_b$	10.8	<.01	.14
Hand, finger, or other complex mannerisms	.66 (.82)$_a$	.58 (.69)$_a$	.89 (.87)$_a$	1.89	.16	.03
Repetitive interests or stereotyped behaviors	.66 (.67)$_a$	.67 (.83)$_a$	1.37 (1.01)$_b$	10.9	<.01	.14

Group differences are indicated by subscripts and reflect a Bonferroni adjustment for multiple comparisons

significantly more repetitive behaviors and abnormal sensory response than other cluster subgroups.

Diagnostic Prediction

The percent of children in ASD clusters diagnosed with nonASD, PDD-NOS, and Autistic Disorder at their re-evaluation are summarized in Table 2 for clarity. One hundred thirty six of the 186 children diagnosed with an ASD as a toddler were assessed again around 4 years of age. There were no differences in 2-years ADI-R, ADOS, MSEL, or VABS scores between the 136 children re-evaluated around 4 years old and the 50 children not re-evaluated around age four. Twenty three of these 136 children no longer met

criteria for an ASD when evaluated around 4 years of age; instead these children were defined as having intellectual disability ($n = 7$), language delay ($n = 3$), motor delay ($n = 1$), and typical development ($n = 12$). Results found that children in the "ASD, mild impairment" subgroup were five times as likely than children in the "ASD, severe impairment" subgroup to receive a nonASD diagnosis as compared to a diagnosis of Autistic Disorder, Wald = 7.94, $p = .01$. There were no significant differences between children in the "ASD, moderate impairment" and "ASD, severe impairment" subgroups in terms of likelihood of receiving a diagnosis of nonASD compared to a diagnosis of Autistic Disorder. Furthermore, children in the "ASD, mild impairment" subgroup were eight times as likely than

children in the "ASD, severe impairment" subgroup to receive a diagnosis of PDD-NOS as compared to a diagnosis of Autistic Disorder, Wald = 14.85, $p = .00$, and children in the "ASD, moderate impairment" subgroup were five times as likely than children in the "ASD, severe impairment" subgroup to receive a diagnosis of PDD-NOS as compared to a diagnosis of Autistic Disorder, Wald = 8.56, $p = .00$. It is also important to note that 12 out of 14 children (86%) who did not have SIB noted on the ADOS at 2-years and were re-evaluated at years retained an ASD diagnosis and 11 out of 13 children (85%) who did not have SIB noted on the ADI-R at 2-years and were re-evaluated at 4-years retained an ASD diagnosis.

Discussion

Cluster analysis identified three subgroups of toddlers with ASDs distinguished by level of social, communication, and intellectual abilities and the rate and intensity of repetitive behaviors and abnormal sensory response. These results support past research in that 76% of the variance in distinguishing ASD cluster subgroups was accounted for by social and communication skills, suggesting that social and communication impairments are particularly relevant for the definition and classification of young children with ASDs. It is important to note that, on average, the subgroup with a clear social and communication advantage still performed below average in these domains and still met ASD criteria on standardized diagnostic instruments. Therefore, even the subgroup with more social and communicative abilities showed clinically significant social and communication impairments. These results are not surprising given that social and communication impairments are defining feature of ASDs, which are a heterogeneous group of disorders (American Psychiatric Association 1994).

These findings bring into question whether a dimensional or categorical view of ASDs is more appropriate for toddler populations. Currently, diagnostic classification systems adopt a categorical perception of ASDs and clinically distinct subtypes are thought to delineate symptom profiles and possibly influence developmental course. The diagnosis of Autistic Disorder is reserved for individuals who show social, communication, and behavioral deficits and the diagnosis of PDD-NOS is reserved for individuals who have symptoms of Autistic Disorder but do not meet full diagnostic criteria or have an atypical symptom presentation. Thus, the diagnostic category of PDD-NOS includes a broad range of symptoms and may not represent a clinically distinct subtype of toddlers. A dimensional view of ASDs in toddlers would support a single spectrum of behaviors, rather than distinct diagnoses of Autistic Disorder, PDD-NOS, and Asperger's Disorder, with a

range of symptoms and associated pathology (American Psychiatric Association 2009). This dimensional view could represent two domains of deficit (i.e., social affect and restricted and repetitive behaviors) instead of three domains of deficit (i.e., social, communication, and behavioral; American Psychiatric Association 2009).

The results of our study support a dimensional view of ASDs in toddlers since subgroups were distinguished by level of social, communication, and intellectual abilities and the rate and intensity of SIB rather than distinct symptom profiles. The most important implication of these findings is that toddlers in one subgroup showed more impairments in repetitive behaviors and abnormal sensory response than toddlers in the other two subgroups and many toddlers in the other two subgroups had few (or subclinical) behavioral deficits. Therefore, dimensional classification systems that require clinically significant deficits in SIB using the current DSM definitions may miss many young children with ASDs who show social and communication deficits only (with clinically significant SIB that may not develop until after the toddler years). Dimensional classification systems will thus need to consider the types of SIB appropriate for ASD classification in toddlers (i.e., repetitive behaviors and abnormal sensory response) as well as the rate and intensity at which they occur.

As just mentioned, it was not the *presence* of repetitive behaviors and abnormal sensory response that distinguished toddler ASD subgroups, but the *rate* and *intensity* at which these SIB occurred. For instance, even though repetitive body use was more frequently observed in children labeled "ASD, severe impairment," children with an ASD placed in other subgroups still showed "mildly abnormal body use" associated with "minor peculiarities" (Schopler et al. 1988). These findings support the hypothesis that SIB represent a continuum of behaviors that may or may not reach clinical significance in toddlers (Richler et al. 2007, 2010). Again, these findings caution that a diagnostic requirement of clinically significant impairments in multiple SIB may exclude many toddlers who retain their diagnosis into the pre-school years and delay early intervention referrals.

The subgroup that consistently had higher rates of repetitive behaviors and abnormal sensory response also had mild-moderate intellectual disability, which begs the question of how this subgroup differed from the other subgroup with mild intellectual disability. Results found that the "ASD, severe impairment" subgroup differed from the "ASD, moderate impairment" subgroup in that the former had lower visual reception scores and more autistic deficit than the latter, despite similar expressive and receptive language skills. It may be, then, that developmental level is responsible for the initial appearance of certain SIB and higher rates of these SIB further disrupts

social development which leads to more impaired functioning (Bishop et al. 2006). This hypothesis was partially supported in that toddlers with severe ASD and many lower-order SIB were much more likely than toddlers with mild and moderate ASD and few SIB to receive a diagnosis of Autistic Disorder, compared to a diagnosis of nonASD or PDD-NOS, around 4 years of age.

Higher-order SIB, such as unusual preoccupations and compulsions and rituals, did not distinguish ASD subgroups in this analysis. This lack of difference can be explained by low frequency of higher-order SIB for all ASD subgroups and suggests that higher-order SIB are not particularly relevant to younger cohorts and are not useful in classifying and diagnosing toddlers with ASDs. This perspective is shared among others who also failed to find significant group differences based on higher-order SIB in younger cohorts (e.g., group differences between toddlers with various forms of ASDs as well as toddlers with ASD and DD; Moore and Goodson 2003; Richler et al. 2007, 2010). Yet higher-order SIB are consistently found in older cohorts and do distinguish older children and adults with ASDs. Therefore, higher-order SIB may not develop until after the toddler years or may be related to skills not typically found in toddler populations (e.g., typical or advanced mental age).

One limitation of our study was that ADOS and CARS scores were based on an overlapping behavioral sample and ADOS analyses were used to offer additional support for the cluster solution. However, validation of the cluster solution primarily involved MSEL, VABS, and ADI-R analyses and validation of the cluster solution using former and revised ADOS algorithms were only offered as additional support for these results. In addition, although clinical diagnosis was partially based on CARS ratings, previous analyses on some of the participants in this sample suggest that inter-rater reliability for the CARS total score was .94 (Chlebowski et al. 2010), reflecting standardized scoring for cluster variables. It is also important to note that 27% of the sample was not re-evaluated around 4 years of age due to refusal or migration. Thus, some participant characteristics could have influenced study results. However, additional analyses showed no significant differences in 2-year ADI-R, ADOS, MSEL, or VABS scores between the 136 children re-evaluated around 4 years old and the 50 children not re-evaluated around 4 years old. Therefore, we believe the limitations of the study do not negate the importance of our results.

In conclusion, our study is the first to explore empirically derived subgroups of toddlers with ASDs using a standardized instrument that represents behaviors commonly found in the first few years of life. These types of analyses are useful in generating hypotheses on the development and course of ASDs in childhood and to inform diagnostic practices. We found three subgroups of toddlers with ASDs primarily distinguished by social, communication, and intellectual skills and the rate and intensity of repetitive behaviors and abnormal sensory response, which supports a dimensional diagnostic view of ASDs in toddlers focused on these specific developmental domains. We encourage replication of these analyses with different cluster variables and more diverse samples of children (e.g., identified by other methods and at different ages), in order to support the external validity of our findings. We also encourage diagnostic systems to consider the type and level of behavioral deficit needed for ASD classification in toddlers so all children with ASDs can be identified as soon as possible and referred to appropriate interventions.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2009). Report of the DSM-V Neurodevelopmental Disorders Workgroup. Retrieved 11 July 2009 from http://www.psych.org/MainMenu/Research/DSMIV/DSMV/DSMRevisionActivities/DSM-V-Work-Group-Reports/Neurodevelopmental-Disorders-Work-Group-Report.aspx.

Bishop, S., Richler, J., & Lord, C. (2006). Association between restricted and repetitive behaviors and nonverbal IQ in children with autism. *Child Neuropsychology, 12*, 247–267.

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,* Available Online First 01/07/2010.

DiGiacomo, A., & Fombonne, E. (1998). Parental recognition of developmental abnormalities in autism. *European Child and Adolescent Psychiatry, 7*, 131–136.

Eaves, L., Ho, H., & Eaves, D. (1994). Subtypes of autism by cluster analysis. *Journal of Autism and Developmental Disorders, 24*, 3–22.

Fein, D., Stevens, M., Dunn, M., Waterhouse, L., Allen, D., Rapin, I., et al. (1999). Subtypes of pervasive developmental disorder: Clinical characteristics. *Child Neuropsychology, 5*, 1–23.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The autism diagnostic observation schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders, 34*, 613–627.

Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., et al. (2008). The modified checklist for autism in toddlers: A follow-up study investigating the early detection of autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*, 827–839.

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

Lord, C., Rutter, M., DiLavore, P., & Risi, S. (1999). *Autism diagnostic observation schedule–WPS edition*. Los Angeles, CA: Western Psychological Services.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659–685.

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*, 47–63.

Mullen, E. (1995). Mullen scales of early learning. American Guidance Service.

Oosterling, I., Roos, S., DeBilt, A., Rommeise, N., De Jonge, M., Visser, J., et al. (2010). Improved diagnostic validity of the revised ADOS algorithms: A replication study in an independent sample. *Journal of Autism and Developmental Disorders, 40*, 689–703.

Prior, M., Leekam, S., Ong, B., Eisenmajer, R., Wing, L., Gould, J., et al. (1998). Are there subgroups within the autism spectrum? A cluster analysis of a group of children with autistic spectrum disorders. *Journal of Child Psychology and Psychiatry, 39*, 893–902.

Richler, J., Bishop, S. L., Klienke, J. R., & Lord, C. (2007). Restricted and repetitive behaviors in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*, 73–85.

Richler, J., Huerta, M., Bishop, S. L., & Lord, C. (2010). Developmental trajectories of restricted and repetitive behaviors and interests in children with autism spectrum disorders. *Developmental Psychopathology, 22*, 55–69.

Robins, D. L., Fein, D., & Barton, M. (1999a). *The modified checklist for autism in toddlers (M-CHAT)*. Self-published.

Robins, D. L., Fein, D., & Barton, M. (1999b). *Follow-up interview for the modified checklist for autism in toddlers (M-CHAT FUI)*. Self-published.

Robins, D. L., Fein, D., Barton, M., & Green, J. (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*, 131–144.

Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *The childhood autism rating scale*. Los Angeles, CA: Western Psychological Services.

Sevin, J., Matson, J., Coe, D., Love, S., Matese, M., & Benavidez, D. (1995). Empirically derived subtypes of pervasive developmental disorders: A cluster analytic study. *Journal of Autism and Developmental Disorders, 25*, 561–578.

Siegel, B., Anders, T., Ciaranello, R., Bienenstock, B., & Kramer, H. (1986). Empirically derived sub-classification of the autistic syndrome. *Journal of Autism and Developmental Disorders, 16*, 275–294.

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

Sparrow, S., Balla, D., & Cicchetti, D. V. (2005). *Vineland adaptive behavior scales*. Circle Pines, MN: American Guidance Services.

Stevens, M., Fein, D., Dunn, D., Allen, D., Waterhouse, L., Feinstein, C., et al. (2000). Subgroups of children with autism by cluster analysis: A longitudinal examination. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 346–352.

Szatmari, P., Georgiades, S., Bryson, S., Zwaigenbaum, L., Roberts, W., Mahoney, W., et al. (2006). Investigating the structure of the restricted, repetitive behaviors and interests domain of autism. *Journal of Child Psychology and Psychiatry, 47*, 582–590.

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

Wiggins, L. D., & Robins, D. L. (2008). Brief report: Excluding the ADI-R behavioral domain improves diagnostic agreement in toddlers. *Journal of Autism and Developmental Disorders, 38*, 972–976.