Evidence for Latent Classes of IQ in Young Children With Autism Spectrum Disorder

Jeffrey Munson, Geraldine Dawson, Lindsey Sterling, Theodore Beauchaine, Andrew Zhou, and Elizabeth Koehler
University of Washington

Catherine Lord
University of Michigan

Sally Rogers
University of California Davis Medical Center

Marian Sigman
University of California Los Angeles

Annette Estes and Robert Abbott
University of Washington

Abstract

Autism is currently viewed as a spectrum condition that includes strikingly different severity levels; IQ is consistently described as one of the primary aspects of the heterogeneity in autism. To investigate the possibility of more than one distinct subtype of autism based on IQ, both latent class analysis and taxometrics methods were used to classify Mullen IQs in a sample of 456 children with autism spectrum disorder. We found evidence for multiple IQ-based subgroups using both methods. Groups differed in level of intellectual functioning and patterns of verbal versus nonverbal ability. Results support the notion of distinct subtypes of autism that differ in severity of intellectual ability, patterns of cognitive strengths and weaknesses, and severity of autism symptoms.

DOI: 10.1352/2008.113:439–452

Autism is characterized by impairments in social interaction and communication as well as a restricted repertoire of activities and interests. Children and adults with autism have specific deficits in social and emotional information-processing (Davies, Bishop, Manstead, & Tanta, 1994; Dawson, Meltzoff, Sterling, & Renaldo, 1998) that are considered to be common features of individuals with the disorder. Yet autism is also characterized by a wide variability in more specific impairments, range of symptoms, levels of adaptive and intellectual functioning, and prognosis. These differences in presentation make conceptualization of the disorder difficult, especially given that diagnosis is based on behavioral observations and standardized parental interviews administered by clinicians. Thus, in almost all cases, behavioral characteristics, rather than laboratory or medical tests, determine diagnostic assignment.

Because of the significant symptom heterogeneity found in autism, it is often conceptualized as autism spectrum disorder. Variability in IQ is one of the most salient dimensions of this heterogeneity. Both the *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, 1994) and *International Classification of Dis-*

eases (World Health Organization, 1992) have definitions of Asperger syndrome that include cognitive developmental level as one of the key features distinguishing it from autism. It is estimated that 70% of individuals with autism have IQs in the mentally retarded range (Fombonne, 2003), yet some individuals have above average intellectual ability (Miller & Ozonoff, 2000). Moreover, the IQ profiles of individuals with and those without mental retardation tend to differ, with individuals who have higher IQ typically showing a higher Verbal IQ, on average (Ghaziuddin & Mountain-Kimchi, 2004; Gilchrist et al., 2001). However, there is extreme individual variability in IQ, making it unlikely that a specific cognitive profile can be used for differential diagnostic purposes (Filipek et al., 1999; Siegel, Minshew, & Goldstein, 1996). Nevertheless, researchers have used various strategies to subtype individuals with autism. Some have focused on medical conditions or known biological etiologies contributing to the disorder. Miles et al. (2005) defined subtypes of autism based on whether individuals had features that were stable from birth, suggesting an organic factor, including all of the syndromes that are currently acknowledged as causes of autism (e.g., fragile X syndrome). These individuals, comprising what the authors called the "complex" autism group, also tend to have more seizures, dysmorphic physical features, microcephaly, lower IQs, and a tendency toward poorer outcome than the "essential" autism group. The essential autism group is characterized by higher incidence of sibling recurrence and a family history of autism, higher male to female ratio, higher likelihood of regression and macrocephaly, and overall higher IQs. According to Miles et al., assigning individuals to the complex and essential groups allows for the first stage of characterizing the etiologic heterogeneity of those with autism spectrum disorders. This separation might be especially useful for genetic analyses because it provides a more homogeneous group of individuals (essentials). However, until the etiological substrates of autism are identified, it is impossible to know how truly homogeneous this group is.

Other researchers have focused on defining autism spectrum subgroups according to behavioral patterns of social interaction. Wing and Gould (1979) first characterized autism according to three subtypes: aloof, passive, or active-but-odd. The aloof subtype, which includes children who tend to reject contact and avoid gaze, is typically the most impaired and severely autistic (e.g., Castelloe & Dawson, 1993; Sevin et al., 1995). Levels of IQ tend to correspond to social typology, with the aloof group having the lowest IQ, followed by the passive and then active-but-odd groups (Borden & Ollendick, 1994). The aloof group also tends to have the lowest levels of adaptive behavior, worse language and communication skills, and higher ratings of stereotyped behavior/restricted interests. Intellectual functioning likely accounts for a large proportion of the variance in predicting language and communication skills, the presence of stereotyped behaviors, and other prototypically autistic behaviors, which may partly contribute to group assignment. Indeed, Volkmer, Cohen, Bregman, Hooks, and Stevenson (1998) found that IQ is often a predictor of social subtype assignment; however, it may not fully account for it. Because level of intellectual functioning may be among the strongest indicators of subtype, investigators have often attempted to divide the autism spectrum disorder group by choosing an a priori IQ cutoff in order to designate and then characterize the resulting high- and low-functioning groups (Allen et al., 2001; Bartek & Rutter, 1976). The lower functioning cognitive subgroup, defined as having an IQ below 70 or 80, tends to exhibit more self-injury, stereotypies, and prototypical autism behaviors. Yet such cutoff points are somewhat arbitrary, making it likely that there is diagnostic overlap between the cognitive subgroups generated.

Furthermore, distinctions between verbal and nonverbal information-processing abilities are often not explored, but may be important in identifying subtypes in autism. Tager-Flusberg and Joseph (2003) investigated discrepancies between verbal and nonverbal IQ in children with autism and found that children with discrepantly high nonverbal skills relative to verbal skills had greater social impairment, independent of absolute level of verbal ability and overall ability.

Attempts have been made to investigate subgroups within a dimensional construct on the basis of non-unimodal distributions. Meehl (1995) noted that although bimodality and marked skewness may be suggestive of latent groups, the presence of bimodality is neither a necessary nor sufficient condition for the existence of latent subgroups. For example, when two latent distributions have a mean difference of 2 *SD*s and equal variances, bimodality may not even be apparent. On the other had, Grayson (1987) noted that even when bimodality is observed in measured variables, the underlying structure may still be a continuous dimension.

Statistical strategies may provide a more empirical basis for characterizing individuals within possible autism spectrum disorder subgroups. A review of the literature indicates that most cluster analytic studies yield two, three, or four subgroups based on degree of impairment. Sevin et al. (1995), for example, used cluster analysis to classify 34 children with autism or pervasive development disorder-not otherwise specified (PDD-NOS) into four groups, described as ranging from high-functioning to low-functioning (severe) autism, with IQ decreasing with severity, and differing significantly between groups. Similarly, Eaves, Ho, and Eaves (1994) used a standard clustering algorithm and principal components analysis of variables to assign 166 children into four groups, ranging from typically autistic and lower functioning to a higher functioning group that more closely resembled Asperger syndrome. Again, severity of autism was related to intellectual impairment in that the most impaired subtype had the lowest average IQ. In a longitudinal examination of 138 school-age children with autism, Stevens et al. (2000) employed hierarchical agglomerative cluster analysis to validate a two group solution, in which cognitive level was the largest separating variable. Children who were lower functioning as defined by nonverbal IQ at preschool age tended to show poorer outcome at school age, suggesting that nonverbal IQ is an extremely potent predictor of membership among school-age children.

Often, cluster analytic techniques have been used to determine which behavioral features of autism tend to correlate or account for the majority of variance—or which factors cluster together. Once a cluster solution is determined and individuals are assigned to groups, the subtypes are characterized using various descriptors, including level of intellectual functioning. In using IQ as a descriptor only *after* the groups have been defined, however, makes it difficult to determine the true role of intellectual capacity in the formation of subgroups, and the actual distribution of IQ in the samples. Few investigators have focused exclusively on cognitive functioning as the empirical indicator of subgroup classification. Those who have specifically investigated the role of intellectual capacity in differentiating autism spectrum disorder subtypes have often found that IQ is the most significant contributor in discriminating between groups and the basis of differences between subtypes (e.g., Miller & Ozonoff, 2000).

Although the goal of cluster analysis is to determine the categories underlying autism spectrum disorders, these methods often yield groups with considerable diagnostic overlap. Unfortunately, under such conditions, cluster analysis often (a) fails to identify the correct number of clusters in datasets where group membership is known and (b) performs poorly in sorting individuals into subgroups (e.g., Krieger & Green, 1999; Tonidandel & Overall, 2004). Furthermore, statisticians have long recognized that clustering algorithms partition datasets into subgroups, even if the distributions are known to be continuous (see Beauchaine, 2003). Thus, results derived solely from cluster analysis do not provide strong evidence for subgroups of autism and do not eliminate the possibility of a spectrum of autistic-like disorders (Prior et al., 1998). In fact, data from eight cluster analytic studies suggest that children with PDD-NOS may fit into one of two overlapping groups and that the subtypes resemble each other, existing along a continuum, and differing only by degree of impairment (Myhr, 1998).

In a review of subtyping studies of autism, Beglinger and Smith (2001) posited their best guess that symptom heterogeneity can be represented by three continua (developmental delay, social impairment, and repetitive behaviors), and rough divisions can be drawn along these continua yielding four subgroups. The authors also noted the weaknesses associated with cluster analytic techniques, including the dependence on the investigators' choice of variables and characteristics of the sample. This conclusion of the presence of a "continuum containing subgroups" highlights the continued difficulty researchers in this area have in determining whether true differences between subgroups in autism can be reliably distinguished.

In part as a result of the limitations of cluster analysis, additional classification techniques, including latent class analysis and taxometrics, have been developed. Although rarely used to evaluate whether subgroups of autism exist, these techniques offer several advantages over clustering algorithms (Beauchaine & Marsh, 2006). For example, latent class analysis provides objective measures of fit for comparing alternative subgroupings, and taxometric analyses are far less prone to identify spurious subgroups within continuous distributions. The lone example of taxometric analysis (based on an adaptation of the regression-mixture model, Golden & Mayer, 1995) in autism is the Autism and Language Disorders Nosology project (Rapin, 1996), in which investigators found evidence for two discrete subgroups, or taxa, in a sample of children with PDD (Fein, et al., 1999) with a nonverbal IQ of about 65, optimally dividing the groups. In the present paper, we used both latent class analysis and maximum covariance (MAXCOV), the most widely studied taxometric algorithm, to address the question of whether subgroups of autism spectrum disorder can be identified from the verbal and nonverbal IQs of probands.

To summarize, although it is unclear whether distinct subtypes of autism exist, a recurring pattern emerges in which IQ strongly predicts social functioning, adaptive behavior, severity of symptoms, and prognosis (Bolte & Poustka, 2002; Carpentieri & Morgan, 1996; Coplan & Jawad, 2005; Howlin, Goode, Hutton, & Rutter, 2004; Liss et al., 2001). We used both MAXCOV and latent class analysis to analyze verbal and nonverbal IQs obtained from a large sample of preschool-age children diagnosed with autism spectrum disorder, who were evaluated through the National Institute for Child Health and Human Development (NICHD) Collaborative Program of Excellence in Autism. Although cluster analysis offers no proven means of choosing among models with different numbers of classes and tends to overextract classes when defining subtypes, latent class analysis and MAXCOV offer an alternative and more conservative approach in determining whether there is a bimodal or multimodal distribution of intellectual functioning among individuals with autism. By using young children in this analysis, we hoped to minimize individual difference related to experience and treatment.

Method

Participants

Participants were 456 children (370 boys [81%], 86 girls [19%]) with autism spectrum disorder who were between the ages of 24 and 66 months (M = 43.4, SD = 8.7); they were participating in studies affiliated with the NICHD Collaborative Program of Excellence in Autism. Exclusionary criteria included the presence of a neurological disorder of known etiology, significant sensory or motor impairment, major physical abnormalities, and history of serious head injury and/or neurological disease. Diagnosis of autism spectrum disorder was based on administration of the Autism Diagnostic Observation Schedule-Generic (ADOS-G) and Autism Diagnostic Interview-Revised (ADI-R). All of the children met criteria for autism (n = 357, 78%) or autism spectrum disorder (n = 99, 22%) on the ADOS-G. Nearly all of the children met criteria for a diagnosis of autism on the ADI-R (n = 431, 95%), with the remaining 25 children within 2 points of a diagnosis of autism on the ADI-R. Informed consent was appropriately obtained from each child's parent/guardian prior to their participation in this study.

Measures

Autism Diagnostic Interview–Revised. The ADI-R (Lord, Rutter, & Le Couteur, 1994) is a structured, standardized parent interview developed to assess the presence and severity of symptoms of autism in early childhood across all three main symptom areas (social relatedness; communication; and repetitive, restrictive behaviors). This interview has been psychometrically validated across a wide range of ages and severity levels in autism. Each site contained one experimenter who was trained to reliability by one of the authors (C. Lord) on the ADI-R; that person then trained other raters in her lab to a reliability of 85% or better.

Autism Diagnostic Observation Schedule–Generic (Lord et al., 2000). The ADOS-G is a semi-structured standardized interview using developmentally appropriate social and toy-based interactions in a 30- to 45-min interview to elicit symptoms of autism in four areas: social interaction, communication, play, and repetitive behaviors. The ADOS-G consists of four modules, each directed at a particular level of language ability. In the present study, all participants received Module 1, developed for preverbal children or those just beginning to speak. The ADOS-G has been psychometrically validated across a wide range of ages and severity levels in autism (Lord et al., 2000). Lord trained an experimenter at each site to reliability on the ADOS-G at the University of Chicago; that person then trained other raters in the lab to a reliability of 85% or better.

Mullen Scales of Early Learning (1997). This instrument is a standardized developmental test for children ages 3 months to 60 months consisting of five subscales: Gross Motor, Fine Motor, Visual Reception, Expressive Language, and Receptive

Language. The last four are combined to yield an overall composite score of intellectual functioning. The Mullen Scales of Early Learning demonstrates strong concurrent validity with other well-known developmental tests of motor, language, and cognitive development. We administered this test to all participants according to standard instructions by raters who were trained in assessing young children with autism and other developmental disorders. Reinforcers for all participants in all groups were used at times to reward cooperation and attention. A set of four developmental quotients for each participant was constructed dividing the age equivalence score for each Mullen subscale by the child's chronological age (CA) and then multiplying by 100. We calculated an overall verbal score by averaging the receptive and expressive language scores and a nonverbal IQ by averaging the visual reception and fine motor scores. As has been done in other samples with young children who have autism (e.g., Lord et al., 2006), we used ratio based scores throughout this paper because the Mullen subscale T-scores commonly yielded a floor score in this sample (% with floor T score of 20: Visual Reception, 59%; Fine Motor, 72%; Receptive Language, 80%; Expressive Language, 76%).

Vineland Scales of Adaptive Behavior, Interview Edition. The Vineland (Sparrow, Balla, & Cicchetti, 1984) is a standardized parent interview designed to assess adaptive behavior across four domains: Social, Communication, Daily Living, and Motor Skills. Standard scores of these four domains were used because floor effects were uncommon given the wide range of normative data available for these domain scores.

Results

Maximum Covariance Analysis

Taxometric analyses were conducted using MAXCOV, which is particularly well suited for identifying overlapping dichotomous distributions embedded within a range of observed scores. (A second taxometric algorithm, mean above minus below a cut was also applied to the data. Because the results of these analyses were fully consistent with the results obtained from MAXCOV, the more commonly used taxometric algorithm [Haslam & Kim, 2002], only the latter are reported.) In contrast to common clustering algorithms that tend to impose structure upon a dataset regardless of whether it is continuous or discrete in nature, taxometric procedures begin with the null hypothesis that the measured traits represent continua and seek disconfirming evidence of this assumption (Beauchaine, 2003). In MAXCOV, variables are taken three at a time and the covariance of two is calculated within adjacent intervals of the third. A smoothed regression function is then fitted through the resulting covariance values. Peaked regression functions are suggestive of discrete latent classes, whereas flat regression functions characterize continua (see Beauchaine & Marsh, 2006; Waller & Meehl, 1998). In the discrete case, the location of the MAXCOV peak identifies the most efficient cutoff point for dividing the sample, which in turn allows for estimation of the base rates for both groups. With four variables for analysis (Mullen IQs), 12 nonredundant MAXCOV plots can be generated. When peaked functions consistently emerge within the same interval, indicating similar baserates regardless of the variable combinations used, more confidence can be placed in the identified classes as being truly discrete. Given such consistency, Bayesian-estimated group membership probabilities are calculated by combining information from each MAXCOV run. Six of the 12 conditional covariance plots indicate a sharp peak in the function indicating a low base rate taxon at the high end of the Mullen IQ distributions. Baserate estimates based on the posterior probability of taxon membership resulted in a subgroup of high Mullen scores that comprised 17.8% of the sample.

Latent Class Analysis

After finding evidence for discontinuity in the IQ distributions in the sample using the MAXCOV procedure, we then examined the data using latent class analysis (Lazarsfeld & Henry, 1968), which is a maximum likelihood-based method that provides model-based parameter estimates in contrast to model-free methods such as cluster analysis. The software program M-plus (Muthen & Muthen, 2004) was used to assess whether there is evidence of multiple, unobserved latent classes in this sample based on the four subscales of the Mullen. Models specifying 2, 3, and 4 latent classes were run. Variances for each class were assumed to be equal in order to minimize the number of parameters being estimated.

Results indicated that a two-group model fit the data significantly better than a single group model (see Table 1). Similarly, three groups fit

Table 1. Latent Class Analysis Fit Indices

No. of classes	Bayesian information criteria	Entropy	Loglikli-hood	Lo-Mendel-Rubin value
One			897.0	
Two	−1882.3	0.840	1017.7	237.8**
Three	−1886.4	0.690	1053.4	70.4*
Four	−1892.5	0.773	1090.1	68.8*
Five	−1850.9	0.748	1103.0	73.9

* $p < .05$. ** $p < .01$.

better than two, and four groups fit better than three. No improvement in fit was seen with a five-group model. Means on the Mullen IQs for the four groups identified in the four-groups latent class analysis solution are illustrated in Figure 1. Rather than simply reflecting overall differences in intellectual functioning, these groups illustrate striking differences in both the absolute level of functioning as well as the relative abilities of the verbal and nonverbal areas. The lowest group (59% of the sample) was characterized by extremely low verbal scores as well as very low nonverbal scores, with an average difference between verbal and nonverbal scores of 22 points. The second group (12.5%) was similar; however, the discrepancy between verbal and nonverbal scores in this group was even more extreme, with nonverbal scores averaging 42 points higher than verbal scores. The third group (21.7%) showed moderate to mild level of impairment, with scores in the 60 to 70 range, with verbal abilities commensurate with nonverbal abilities. The final group (7.0%) reflected a subgroup of children functioning in the average range, again with verbal and nonverbal areas at roughly comparable levels. This fourth group was the only one in which the pair of Verbal or Nonverbal subtests showed a widely different pattern. In this group the Visual Reception scale was much noticeable higher than the Fine Motor subscales.

Children in Group 2, which showed the very large discrepancy between verbal and nonverbal scores, were, on average, nearly one year younger than the children in the other groups, $F(3, 452) = 26.6$, $p < .001$. Means and *SD*s for the age in months for each group were Group 1, 45.3 (8.5); Group 2, 34.8 (6.8); Group 3, 43.5 (7.7), Group 4, 42.5 (7.0). Boys and girls were equally likely to be present in each of the four groups, $\chi^2 = 0.85$, $p = .839$. The number and percentage of boys and girls present in each group are as follows: boys, Groups 1 to 4, respectively, 217 (58.6%), 45 (12.2%), 83 (22.4%), and 25 (6.8%); girls, Groups 1 to 4, respectively, 51 (59.3%), 12 (14.0%), 16 (18.6%), and 7 (8.1%).

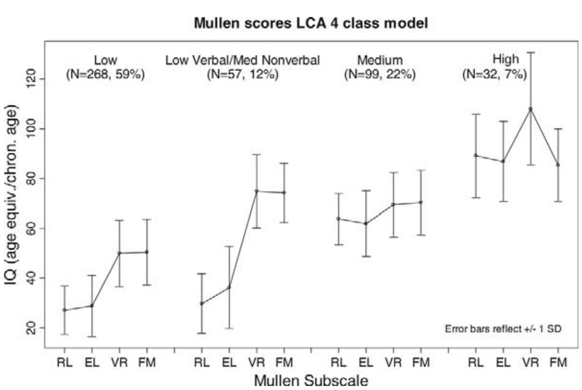

Figure 1. Mean Mullen IQs of latent groups. LCA = latent class analysis. Rec = Receptive Language, Exp = Expressive Language, VR = Visual Reception, and FM = Fine Motor.

Comparison Between Latent Class Analysis and MAXCOV Procedures

Although the four-group model was found to provide the best fit of the data in the latent class analyses, we wanted to compare the classification results of the two class latent class analysis model with the MAXCOV results to directly compare the two-group solution of these different methods. Both methods identified a relatively small group of higher functioning children, with the two-group latent class analysis model placing 19.1% of the sample in this high group; and MAXCOV, 15.1% of the sample in this group. Overall, there was high agreement between the latent class analysis constrained variance classification results and the MAXCOV procedure because 85.6% of individuals were classified similarly by the two approaches. Thus, 14.4% was classified differently by these two methods. When examining the combination of the latent class analysis, we found that the two-group results and the MAXCOV results with the latent class analysis four-group results described above, had a striking similarity. Children placed in the low group by the latent class analysis two-group model but in the high group by MAXCOV show the same extreme difference between their verbal and non-

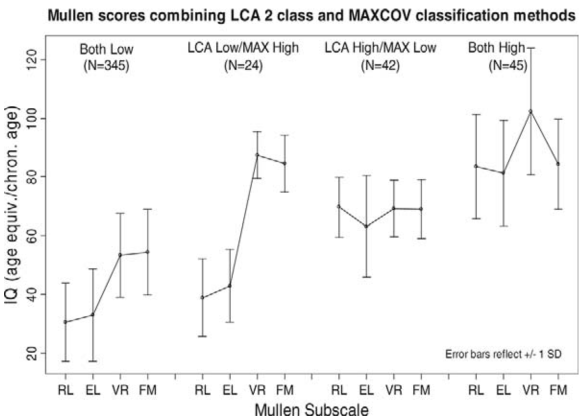

Figure 2. Mean Mullen scores of latent groups by combining classification methods. LCA = latent class analysis, RL. = Receptive Language, EL = Expressive Language, VR = Visual Reception, and FM = Fine Motor.

verbal scores (Figure 2). Children in the opposite group (latent class analysis high/MAXCOV low) show relatively equal verbal and nonverbal scores in the 60 to 70 range.

Figure 3 illustrates the relationship between the Mullen Verbal (*M* of Receptive and Expressive Language subscales) and Nonverbal (*M* of Visual Reception and Fine Motor subscales) IQs using a convex hull plot (Vida & Polar, 2005). Given the substantial overlap of verbal and nonverbal scores between the groups, the convex hull for each subgroup, rather than each individual point, is displayed. Relative density plots for each variable are shown in each margin. This plot reveals the complexities involved in categorizing young children with autism spectrum disorder in terms of their intellectual functioning. Clearly the low and low/medium groups showed strengths in nonverbal relative to verbal performance. However, the two higher groups show more commensurate verbal and nonverbal performance. The relative density plots nicely reveal the overlapping verbal and nonverbal distributions; however, information about the relative sample size is lost. The line graphs depict the overall verbal and nonverbal IQ distributions broken down by latent class analysis group membership.

Notably, the groups depicted in these plots have been assigned descriptive labels of low, low verbal/medium nonverbal, medium, and high. Although helpful as a shorthand method to referring to these latent classes, one should not place undo emphasis on the meaning of these labels. For ex-

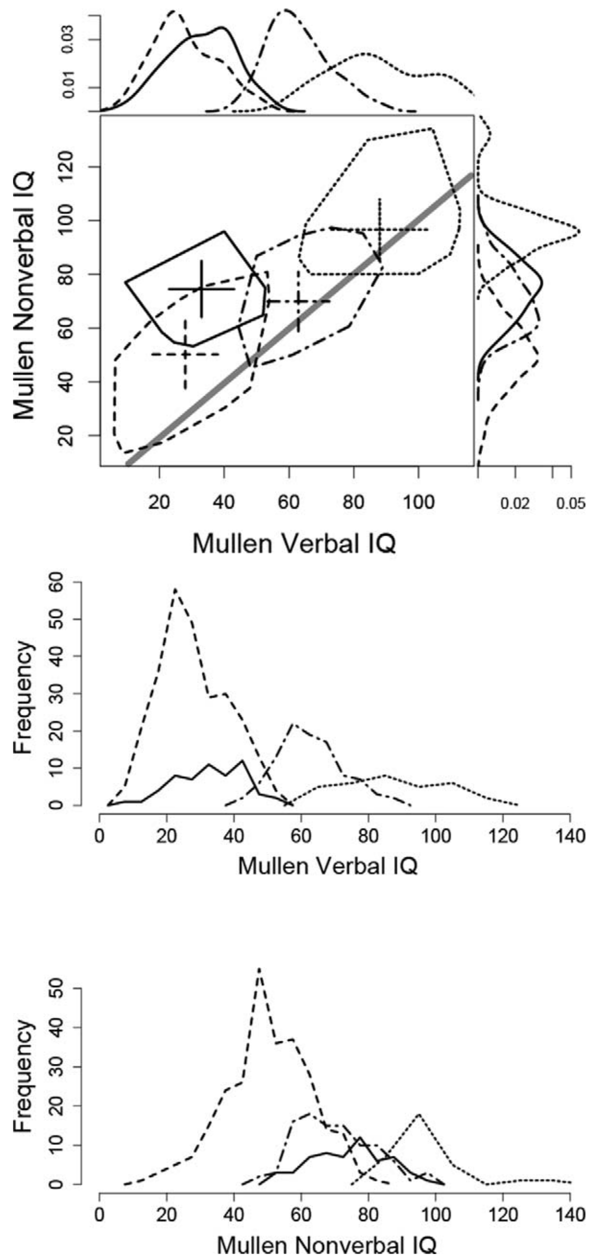

Figure 3. Convex hull plots and frequency distributions of Mullen Verbal and Nonverbal IQ by latent class analysis (LCA) groups. For the convex hull plot, marginal distributions reflect relative density plots for each group. Each cross reflects the bivariate group mean ± 1 *SD*. The diagonal line depicts equivalent verbal and nonverbal IQ. Group 1, Low (dashed line), Group 2, Low/Medium (solid line); Group 3, Medium (dash-dotted line); Group 4 (dotted line).

Table 2. Vineland, Autism Diagnostic Interview-Rev. (ADI), and Autism Diagnostic Observation Schedule-Generic (ADOS) Scores by Latent Class Analysis Group

	Low (1)		Low verbal/nonverbal (2)		Medium (3)		High (4)	
Measure	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Vineland								
Socialization	57.6	6.4	60.1	9.0	65.2	8.4	66.8	10.0
Com.[b]	54.0	6.8	58.6	5.3	67.3	8.2	77.9	12.8
Daily Living Skills	57.2	7.8	65.3	6.8	64.0	8.3	67.4	8.4
Motor Skills	63.2	14.2	77.8	16.9	72.1	12.9	81.7	16.7
ADOS								
Social	11.4	2.1	11.5	2.1	9.2	2.5	8.5	2.7
Com.	6.5	1.4	6.5	1.5	5.7	1.6	5.8	1.6
ADI								
Social	20.1	4.7	18.7	3.9	17.1	4.9	15.7	5.0
Com.	11.9	2.3	11.2	2.2	12.5	4.1	13.3	3.7
Repetitive	5.1	1.8	4.3	1.7	5.4	2.2	5.9	2.8

[a]Cell values indicate significant pairwise group differences at $p < .05$. [b]Communication.

ample, children placed in the high group in the latent class analysis will not necessarily be described clinically as having high-functioning autism. Some children in both the medium and high groups have verbal IQs above 80, and some children in all but the low group have nonverbal IQs above 80. One can see that the medium group in Figure 3 overlaps with every other group. Children whose scores fall into these areas of overlap have notably lower posterior probabilities of group membership. Though the mean posterior probability for group membership was .88, 24% of the sample had probabilities of less than .80, and 8%, less than .60.

Group Classification and Relationship to ADI, ADOS, and Vineland

Finally, we compared the groups identified via the latent class analysis four-group model on the Vineland, ADI, and ADOS measures (Table 2) with ANOVA and Bonferroni post-hoc comparisons. Vineland Socialization and Communication domains both showed a gradation in scores consistent with the findings for the Mullen. Vineland Communication scores were increasingly higher across Groups 1 to 4 and Socialization scores were increasingly higher across Groups 1 to 3, with no difference between Groups 3 and 4. On the Daily Living Skills and Motor Skills domains, only Group 1 had significantly lower scores compared with the other three groups.

The ADOS social and ADI Social scores showed similar patterns in which Groups 1 and 2 tended to show greater impairment (i.e., higher scores) than did Groups 3 and 4. Groups 1 and 2 also showed more impairment than did Group 3 on the ADOS Communication score. In contrast, on the ADI Communication scores, Groups 3 and 4 had higher scores than did Group 2. This likely reflects the greater number of items for which verbal children may be rated compared to nonverbal children. Finally, Groups 3 and 4 had significantly higher ADI Repetitive and Stereotyped Behaviors scores than did Group 2.

Clearly, these groups differed in their adaptive behavior and symptom presentation; however, the question remains as to whether these differences simply reflect the overall relationship of verbal and nonverbal abilities to adaptive behavior and autism symptoms. To examine this question, we regressed verbal (*M* of Receptive and Expressive Language subscales) and nonverbal (*M* of Visual Reception and Fine Motor subscales) IQs on each Vineland, ADOS, and ADI measure. Then, we added latent class analysis group membership as a second step in the regression, using three dichotomous indicators (Variable 1 coded 1 for Group 2, Variable 2 coded 1 for Group 3, and Variable 3 coded 1 for Group 4; Group 1 was coded zero on all three variables). Table 3 shows these results in which latent class analysis group membership added significantly to the prediction

Table 2. Extended

	Post-hoc Bonferroni comparisons[a]					
F	1 vs. 2	1 vs. 3	1 vs. 4	2 vs. 3	2 vs. 4	3 vs. 4
37.72	1 < 2	1 < 3	1 < 4	2 < 3	2 < 4	
131.06	1 < 2	1 < 3	1 < 4	2 < 3	2 < 4	3 < 4
35.01	1 < 2	1 < 3	1 < 4			
27.41	1 < 2	1 < 3	1 < 4			
35.80		1 > 3	1 > 4	2 > 3	2 > 4	
8.57		1 > 3		2 > 3		
16.00		1 > 3	1 > 4		2 > 4	
4.68				2 < 3	2 < 4	
5.40				2 < 3	2 < 4	

of Vineland Socialization, Communication, Daily Living Skills, and ADI Social scores above and beyond that of verbal and nonverbal IQ. This provides another line of evidence for latent classes of IQ as opposed to a single or even bivariate (namely, verbal and nonverbal) continuous dimension of IQ in young children with autism spectrum disorders.

Discussion

Heterogeneity in autism has long been noted. Variability in IQ is routinely one of the largest contributors to this heterogeneity. In the current study we examined the degree to which variability in information-processing, both verbal and nonverbal, may reflect unobserved or latent groups within a large sample of preschool children with autism spectrum disorder. Thus, rather than using a single wide-ranging distribution of IQ, we sought to determine whether there is evidence that this wide range of functioning may reflect the presence of discrete latent groups. By utilizing two different techniques, latent class analysis and taxometric analysis (Beauchaine, 2003), we avoided some of the difficulties present in earlier work in which researchers investigated subgroups in autism based on IQ. First, we employed a taxometric analysis to assess whether evidence for discrete distributions of IQ could be found. This approach did suggest discontinuity in the IQ distribution at the higher end of the Mullen scale.

With this evidence, we used latent class analyses to explore this question and to assess whether two or more latent groups based on IQ could be identified. The latent class analysis revealed a four-group solution as providing the best fit for these data. These four groups can be described as a mixture of the overall level of functioning in the group, coupled with the presence or absence of a large discrepancy between verbal and nonverbal functioning. The largest group, comprised of children with low nonverbal and very low verbal scores, reflected those with the most severe cognitive impairments. A second group also showed very low verbal abilities but had nonverbal scores over 40 points higher, on average. The third and fourth groups reflected commensurate verbal and nonverbal abilities, with the third group showing mild to moderate impairments in their cognitive functioning; and the fourth, highest functioning group scoring in the low average range.

Groups with identical patterns of Mullen scores were observed when contrasting the two group latent class analysis results with those from the MAXCOV method, which yields, by definition, two groups when evidence of taxonicity is present. Children who were classified in a similar manner with these two methods matched the lowest and highest functioning groups identified in the latent class analysis four-class model. Children classified differently across the two methods comprised two distinct IQ-based subgroups that were different from both a simple low- or high-func-

Table 3. Linear Regression Analyses Predicting Vineland, Autism Diagnostic Interview-Rev. (ADI), and Autism Diagnostic Observation Schedule-Generic (ADOS) From Mullen IQ and Latent Class Analysis Group

	Step 1		
Measure	R_2	Verbal IQ β (*t*)	Nonverbal IQ β (*t*)
Vineland			
Socialization	.69***	0.75 (19.49***)	0.11 (2.94**)
Com.[a]	.31***	0.46 (8.29***)	0.12 (2.18*)
Daily Living Skills	.38***	0.04 (0.67)	0.59 (10.73***)
Motor Skills	.319***	−0.04 (−0.76)	0.59 (10.40***)
ADOS			
Social	.249***	−0.59 (−10.13***)	0.14 (2.39*)
Com.	.061***	−0.30 (−4.56***)	0.07 (1.15)
ADI			
Social	.166***	−0.40 (−6.61***)	0.00 (−0.08)
Com.	.044***	0.30 (4.55***)	−0.23 (−3.56***)
Repetitive	.032***	0.25 (3.82***)	−0.20 (−3.06**)

[a]Communication.
*$p < .05$. **$p < .01$. ***$p < .001$.

tioning group as well as being different from each other. The verbal versus nonverbal dimension of intellectual functioning in addition to the absolute level of overall functioning is, therefore, important to evaluate when considering subgroups of children with autism.

Though no gender differences were present in the likelihood of group classification, there was an intriguing mean difference in age among the groups. Children in Group 2, who showed an extreme discrepancy between verbal and nonverbal scores, were, on average, nearly one year younger than the children in the other groups. In addition to the low verbal scores and large verbal–nonverbal discrepancy, this group showed the lowest level of repetitive and stereotyped behaviors on the ADI. One interpretation of these findings is that Group 2 may reflect a developmental or maturational effect rather than a uniquely different IQ subgroup. As children in this group age, those whose language and communication ability improved would look like those in Group 3 (with more commensurate verbal and nonverbal abilities), whereas those who made slow progress would tend to resemble members of Group 1 (who had more profound verbal impairments). Perhaps, as well, levels of repetitive behaviors may increase to levels similar to the other groups because these behaviors often follow the social and communication impairments in young children with autism. Alternatively, this group may signify a specific language impairment in these children, as described by Kjelgaard and Tager-Flusberg (2001). Whether language deficits relative to nonverbal IQ in children with autism compared with children who have specific language impairments represent the same or different underlying mechanisms is not known. However, clarifying areas of similarity and difference among children with autism spectrum disorder, children with other developmental disabilities, and typically developing children continues to be an important line of inquiry for understanding the genetics and underlying neurobiologic and neurocognitive mechanisms involved in autism.

This age difference in Group 2 also raises an important measurement issue when working with children who have moderate to severe cognitive impairments. Ratio-based IQs were used in this study to avoid the problem of a floor effect on low-end scores; differences in mean age among the groups (which serves as the denominator in the scores) may come into play. For example, there may be specific items on the Mullen that tend to be particularly difficult (e.g., those requiring imitation?) or easy (e.g., those clearly requiring no spoken instruction) for children with autism in comparison to children in the normative sample functioning at roughly the same developmental level. Simply examining age-equivalence scores

Table 3. Extended

Step 2					
Δ R_2	Verbal IQ β (*t*)	Nonverbal IQ β (*t*)	Group 2 β (*t*)	Group 3 β (*t*)	Group 4 β (*t*)
.008*	0.91 (14.19***)	0.07 (1.54)	0.05 (1.38)	−0.1 (−2.23*)	−0.11 (−2.22*)
.035***	0.71 (7.86***)	0.18 (2.65**)	−0.03 (−0.70)	−0.17 (−2.55*)	−0.34 (−4.71***)
.027***	0.23 (2.59**)	0.59 (8.89***)	0.04 (0.75)	−0.08 (−1.25)	−0.25 (−3.53***)
.013+	0.13 (1.36)	0.57 (8.06***)	0.04 (0.84)	−0.1 (−1.36)	−0.17 (−2.24*)
.003	−0.56 (−5.85***)	0.12 (1.66)	0.01 (0.11)	−0.04 (−0.63)	0.02 (0.33)
.006	−0.28 (−2.56*)	0.06 (0.81)	−0.02 (−0.34)	−0.07 (−0.86)	0.04 (0.44)
.019*	−0.65 (−6.45***)	0.04 (0.50)	−0.06 (−1.13)	0.16 (2.11*)	0.20 (2.60**)
.008	0.19 (1.72)	−0.26 (−3.21**)	0.01 (0.20)	0.07 (0.88)	0.15 (1.84)
.013	0.08 (0.70)	−0.17 (−2.13*)	−0.05 (−0.94)	0.09 (1.17)	0.16 (1.92)

rather than the ratio-based scores does not address this issue because functioning at a 20-month-old level will mean vastly different things for a 2- versus a 4-year-old child.

The highest functioning group identified by the latent class analysis (Group 4) had much higher mean Mullen scores (verbal IQ = 88, nonverbal IQ = 97) compared to Group 3 (verbal IQ = 63, nonverbal IQ = 70). Despite this large difference in intellectual functioning, these groups did not differ on the Vineland Socialization, Daily Living Skills, or Motor Skills domains. The average standard score across the four Vineland subdomains was 73 for Group 4 and 67 for Group 3, showing a much bigger relative weakness in adaptive functioning for Group 4. In addition, Groups 3 and 4 did not differ on any ADOS or ADI scores. Despite the striking difference in levels of cognitive functioning between Groups 3 and 4, this did not translate into systematic differences in adaptive functioning or level of autism symptoms as measured by the ADOS and ADI.

We also found that latent class analysis group membership accounted for a significant proportion of the variability of Vineland Socialization, Communication, Daily Living Skills, and ADI social scores beyond that accounted for by the Mullen Verbal and Nonverbal IQs. This provides additional evidence of the importance of considering IQ in young children with autism spectrum disorder more than simply a single dimension. Indeed, the results presented here suggest that there is more than a simple linear relationship between intellectual functioning and adaptive behavior and autism symptoms, even when independently measuring both verbal and nonverbal intellectual abilities.

This study comprises a very large sample of young children who have been carefully diagnosed as being on the autism spectrum, and, thus, we believe the findings here should generalize to the broader population of children with autism spectrum disorder in this age range. However, this sample was not collected as a population-based epidemiologic sample. Replication is clearly needed. One should be particularly cautious in making assumptions regarding the relative size of each of the IQ groups described. Of the 1,824 Mullen subscales (Four Subscales × 456 Children), only 11% represent scores in the "average" range (T scores of 40 or greater, 16th percentile or greater). In many studies of older children with autism spectrum disorder, researchers have reported a much higher proportion in the average range of intellectual functioning. There are several issues that may bear on the representativeness of the present sample. First, this sample may truly contain fewer high-functioning children with autism spectrum disorder than are present in the population because families who seek to participate in university studies with their preschooler who has autism spectrum disorder may not be representative of all families with preschoolers who have autism spectrum disorder. Second, these children were tested primarily between late 1998 and 2001, which may represent a cohort effect compared to more recent samples. Third, by assessing preschoolers and using a developmental instrument designed for use from infancy through age 5, we may have produced a sample that included a greater proportion of more impaired children than is often found in samples of older children. It is often a fine line between getting a valid estimate of a nonverbal child's level of intellectual functioning and simply being unable to collect valid, interpretable, quantitative data. Some of these children will continue to fall further and further behind their same-age peers in terms of their cognitive development. Many well-developed instruments for older children are limited in their use with the low end of intellectual functioning because normative data are simply not available. Thus, there may be a portion of children in the present sample, who, by virtue of the difficulties in obtaining meaningful quantitative data, may simply never take part in studies of older children with autism spectrum disorder.

Whether the methods used here or similar statistical methods will provide evidence for latent classes based on IQ in older children remains an open question that necessities further study. However, careful attention needs to be paid to how the use of a given measurement instrument will impact the scores it yields (e.g., How much variability on the low end of the scale is there?) and even the resulting sample the study produces. This is particularly true when verbal information-processing and language development in children with autism are studied (Tager-Flusberg, 2000). The likelihood that a sample of preschoolers with autism, when assessed 5 or 10 years later, will present an even wider range of functioning than they did at the outset makes the identification and measurement of the underlying neurocognitive mechanisms at work in autism that much more difficult. In this pursuit, we must be aware of how our measurement tools, presumably impartial and objective, may also become construction tools that actively shape and influence the phenomenon we are studying. Any seasoned child assessor knows, a la Heisenberg, that observing a child at a table with an open briefcase full of booklets and blocks can have a profound influence on what we end up measuring.

In summary, by examining a large, well-characterized sample of preschool children with autism spectrum disorder, we have provided an important evidence of the presence of multiple IQ-based subgroups within autism. Four latent classes were identified that represent very different levels of intellectual functioning as well as different patterns of relative verbal versus nonverbal abilities. We found that group membership was related to adaptive functioning and social impairment, above and beyond the direct relationship of verbal and nonverbal IQ. Cross-sectional samples such as this must be complemented with longitudinal data because variability in course represents yet another area of heterogeneity in autism where much remains to be learned.

References

Allen, D. A., Steinberg, M., Dunn, M., Fein, D., Feinstein, C., Waterhouse, L., & Rapin, I. (2001). Autistic disorder versus other pervasive developmental disorders in young children: Same or different? *European Child and Adolescent Psychiatry, 10,* 67–78.

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

Bartek, L., & Rutter, M. (1976). Differences between mentally retarded and normally intelligent autistic children. *Journal of Autism and Childhood Schizophrenia, 6,* 109–120.

Beauchaine, T. P. (2003). Taxometrics and developmental psychopathology. *Development and Psychopathology, 15,* 501–527.

Beauchaine, T. P., & Marsh, P. (2006). Taxometric methods: Enhancing early detection, diagnosis, and prevention of psychopathology by identifying latent vulnerability traits. In D. Cicchetti & D. Cohen (Eds.), *Developmental psychopathology* (2nd ed., pp. 931–967). Hoboken, NJ: Wiley.

Beglinger, L. J., & Smith, T. H. (2001). A review of subtyping in autism and proposed dimensional classification model. *Journal of Autism and Developmental Disorders, 31,* 411–422.

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

Borden, M. C., & Ollendick, T. H. (1994). An examination of the validity of social subtypes in autism. *Journal of Autism and Developmental Disorders, 24,* 23–37.

Carpentieri, S., & Morgan, S. B. (1996). Adaptive and intellectual functioning in autistic and non–autistic retarded children. *Journal of Autism and Developmental Disorders, 26,* 611–620.

Castelloe, P., & Dawson, G. (1993). Subclassification of children with autism and pervasive developmental disorder: A questionnaire based on Wing's subgrouping scheme. *Journal of Autism and Developmental Disorders, 23,* 229–241.

Coplan, J., & Jawad, A. F. (2005). Modeling clinical outcome of children with autistic spectrum disorders. *Pediatrics, 116,* 117–122.

Davies, S., Bishop, D., Manstead, A. S., & Tantam, D. (1994). Face perception in children with autism and Asperger's syndrome. *Journal of Child Psychology and Psychiatry, 35,* 1033–1057.

Dawson, G., Meltzoff, A. N., Osterling, J., & Rinaldi, J. (1998). Neuropsychological correlates of early symptoms of autism. *Journal of Autism and Developmental Disorders, 69,* 1276–1285.

Eaves, L. C., Ho, H. H., & Eaves, D. M. (1994). Subtypes of autism by cluster analysis. *Journal of Autism and Developmental Disorders, 24,* 3–21.

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

Filipek, P. A., Accardo, P. J., Baranek, G. T., Cook, E. H., Dawson, G., Gordon, B., Gravel, J. S., Johnson, C. P., Kallen, R. J., Levy, S. E. et al. (1999). The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 29,* 439–484.

Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: An update. *Journal of Autism and Developmental Disorders, 33,* 365–382.

Ghaziuddin, M., & Mountain-Kimchi, K. (2004). Defining the intellectual profile of Asperger syndrome: Comparison with high-functioning autism. *Journal of Autism and Developmental Disorders, 34,* 279–284.

Gilchrist, A., Green, J., Cox, A., Burton, D., Rutter, M., & Couteur, A. (2001). Development and current functioning in adolescents with Asperger syndrome: A comparative study. *Journal of Child Psychology and Psychiatry, 42,* 227–240.

Golden, R. R., & Mayer, M. J. (1995). Peaked indicators: A source of pseudotaxonicity of a latent trait. In D. Lubinski & R. Dawis (Eds.), *Assessing individual differences in human behavior: New concepts, methods, and findings* (pp. 93–115). Palo Alto, CA: Davies-Black.

Grayson, D. A. (1987). Can categorical and dimensional views of psychiatric illness be distinguished? *British Journal of Psychiatry, 151,* 355–361.

Haslam, N., & Kim, H. C. (2002). Categories and continua: A review of taxometric research. *Genetic, Social, and General Psychology Monographs, 128,* 271–320.

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

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes, 16,* 287–308.

Krieger, A. M., & Green, P. E. (1999). A cautionary note on using internal cross validation to select the number of clusters. *Psychometrika, 64,* 341–353.

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis.* Boston: Houghton Mifflin.

Liss, M., Harel, B., Fein, D., Allen, D., Dunn, M., Feinstein, C., Morris, R., Waterhouse, L., & Rapin, I. (2001). Predictors and correlates of adaptive functioning in children with developmental disorders. *Journal of Autism and Developmental Disorders, 31,* 219–230.

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

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (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,* 205–223.

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

Meehl, P. E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist, 50,* 266–275.

Miles, J. H., Takahashi, T. N., Bagby, S., Sahota, P. K., Vaslow, D. F., Wang, C. H., Hillman, R. E., & Farmer, J. E. (2005). Essential versus complex autism: Definition of fundamental prognostic subtypes. *American Journal of Medical Genetics, 135A,* 171–180.

Miller, J. N., & Ozonoff, S. (2000). The external validity of Asperger disorder: Lack of evidence from the domain of neuropsychology. *Journal of Abnormal Psychology, 109,* 227–238.

Mullen, E. M. (1997). *Mullen Scales of Early Learning.* Los Angeles: Western Psychological Services.

Muthen, L. K., & Muthen, B. O. (2004). *Mplus user's guide* (3rd ed.). Los Angeles: Author.

Myhr, G. (1998). Autism and other pervasive developmental disorders: Exploring the dimensional view. *Journal of Abnormal Psychology, 43,* 589–595.

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

Rapin, I. (1996). Preschool children with inadequate communication: Developmental language disorders, autism, low IQ. *Clinics in Developmental Medicine* (No. 139). London: MacKeith Press.

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

Siegel, D. J., Minshew, N. J., & Goldstein, G. (1996). Wechsler IQ profiles in diagnosis of high–functioning autism. *Journal of Autism and Developmental Disorders, 26,* 389–406.

Sparrow, S. S., Balla, S., & Cicchetti, D. V. (1984). *The Vineland Adaptive Behavior Scales* (Survey Form). Circle Pines, MN: American Guidance Service.

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

Tager-Flusberg, H. (2000). The challenge of studying language development in children with autism. In L. Menn & N. Bernstein Ratner (Eds.), *Methods for studying language production* (pp. 313–332). Mahwah, NJ: Erlbaum.

Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society of London, B, 358,* 303–314.

Tonidandel, S., & Overall, J. E. (2004). Determining the number of clusters by sampling with replacement. *Psychological Methods, 9,* 238–249.

Vidmar, G., & Pohar, M. (2005). Augmented convex hull plots: Rationale, implementation in R and biomedical applications. *Computer Methods and Programs in Biomedicine, 78,* 69–74.

Volkmer, F. R., Cohen, D. J., Bregman, J. D., Hooks, M. Y., & Stevenson, J. M. (1998). An examination of social typologies in autism. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 82–86.

Waller, N. G., & Meehl, P. E. (1998). *Multivariate taxometric procedures: Distinguishing types from continua.* Newbury Park, CA: Sage.

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Childhood Schizophrenia, 9,* 11–29.

World Health Organization. (1992). *ICD-10 classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines.* Geneva: Author.

Received 2/4/07, accepted 4/12/08.
Editor-in-charge: Frances Conners

Correspondence regarding this article should be sent to Jeffrey Munson, University of Washington, Psychiatry and Behavioral Sciences, Seattle, WA 98195. E-mail: jeffmun@u.washington.edu

