



Published in final edited form as:

J Autism Dev Disord. 2015 May ; 45(5): 1329–1344. doi:10.1007/s10803-014-2295-3.

Detecting Subgroups in Children Diagnosed with Pervasive Developmental Disorder – Not Otherwise Specified

Laura Brennan, M.A., Marianne Barton, Ph.D., Chi-Ming Chen, Ph.D., James Green, Ph.D., and Deborah Fein, Ph.D.

University of Connecticut, Department of Psychology, 406 Babbidge Road, Unit 1020, Storrs, CT 06269-1020

Abstract

Hierarchical cluster analyses were used to detect three subgroups in a sample of children with Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS) evaluated at ages 2 and 4. At age 2, Cluster 1 demonstrated few autism symptoms and high cognitive scores; 60% no longer met criteria for PDD at 4. Cluster 2 exhibited more autism symptoms and lower cognitive scores at 2; 89.5% met criteria for ASD at 4. Cluster 3 had the lowest cognitive scores and most impaired social/communication skills at 2, but no repetitive behaviors; 60% diagnosed with Autistic Disorder at 4. Results shed light on outcomes for different PDD-NOS types and raise questions regarding the increased importance of repetitive behaviors in DSM-5.

Keywords

PDD-NOS; cluster analysis; DSM 5; ASD; subgroups

Previous editions of the Diagnostic and Statistical Manual of Mental Disorders (DSM) included several specific subcategories (Autistic Disorder, Asperger's Syndrome, and Pervasive Developmental Disorder-Not Otherwise Specified [PDD-NOS]) under the broader rubric of Pervasive Developmental Disorders. However, the most recent version of the DSM, the DSM-5, eliminated these subcategories and created the single diagnosis of Autism Spectrum Disorder (ASD; Frazier et al. 2012). The DSM-IV-TR, the DSM-5's predecessor, characterized individuals diagnosed with an ASD based on behavioral features across three domains: social reciprocity, communication, and restricted or stereotyped behaviors or interests (APA 2000; APA 2013). In contrast, the DSM-5 requires impairments from two

Correspondence concerning this article should be addressed to: Laura Brennan, M.A., Department of Psychology, University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, CT 06269-1020, (860) 486-5767, laura.brennan@uconn.edu.
Author Contact Information: Laura Brennan, laura.brennan@uconn.edu, Phone: (860) 486-5767, Fax: (860) 486-2760
Laura Brennan, M.A. was a graduate student in the Clinical Psychology Ph.D. Program at the University of Connecticut, U.S.A. at the time this study was completed. Marianne Barton, Ph.D. was a faculty member in the Clinical Psychology Ph.D. Program at the University of Connecticut, U.S.A. at the time this study was completed. Chi-Ming Chen, Ph.D. was a faculty member in the Clinical Psychology Ph.D. Program at the University of Connecticut, U.S.A. at the time this study was completed. James Green, Ph.D. was a faculty member in the Developmental Psychology Ph.D. Program at the University of Connecticut, U.S.A. at the time this study was completed. Deborah Fein, Ph.D. was a faculty member in the Clinical Psychology Ph.D. Program at the University of Connecticut, U.S.A. at the time this study was completed. Authors' affiliations have remained the same since the time of this study.

This paper was prepared from a paper presentation at the 2012 International Meeting for Autism Research and the Master's thesis of Laura Brennan at the University of Connecticut.

domains: social reciprocity and restricted or stereotyped behaviors or interests (Mandy, Charman, & Skuse 2012; APA 2013). The symptoms of ASDs are sometimes described as falling along a continuum, with more severe symptoms at one end of this spectrum and milder symptoms at the other (Buitelaar, Van der Gaag, Klin, & Volkmar 1999; Walker et al. 2004). The DSM-5 was designed with this assumption of a severity continuum; however, research prior to its publication suggested that one group of individuals, those diagnosed with PDD-NOS, might differ qualitatively from those diagnosed with other ASDs (Paul et al. 2004; Walker et al. 2004).

The PDD-NOS Diagnosis

The DSM-IV-TR outlined specific diagnostic criteria for two of the three ASDs, including Autistic Disorder and Asperger's Syndrome. However, the diagnostic criteria for PDD-NOS did not explicitly define the behaviors necessary for diagnosis. Rather, a diagnosis of PDD-NOS was given when a child demonstrated a combination of symptoms, to include impairments in social interaction skills and *either* communication difficulties, *or* the presence of repetitive or stereotyped behaviors (APA 2000).

As a result of its poorly defined criteria, PDD-NOS was described as a potentially problematic, "catchall" diagnosis that lacked explicit operational definitions and had poor inter-rater reliability (Mandy, Charman, Gilmour, & Skuse 2011; Prior et al. 1998; Walker et al. 2004). Despite these critiques, PDD-NOS remained a highly prevalent diagnosis; it was assigned at a rate 1.7 times that of Autistic Disorder (Chakrabarti & Fombonne 2005).

The absence of more specific criteria for a diagnosis of PDD-NOS suggested a likely degree of heterogeneity within this population. There was, however, little research attempting to further "specify" the PDD-NOS diagnosis (Buitelaar et al. 1999). Instead, research primarily described PDD-NOS in relation to other ASDs in order to examine whether each disorder presented unique and varied profiles or whether each diagnosis differed only by their position along a spectrum of symptom severity (Buitelaar et al. 1999; Paul et al. 2004; Walker et al. 2004).

Subgroups in the Literature

In the past, validating subgroups under the broader umbrella of Pervasive Developmental Disorders was expected to clarify some of the heterogeneity in the presentations of each disorder and thereby clarify the etiology and trajectory of each disorder, as well as to help develop effective treatment plans for children with ASDs (Stevens et al. 2000; Roux, Garreau, Barthelemy, & Hameury 1994). Clarifying the characteristics of children with distinct presentations allowed for a more comprehensive understanding of the unique profiles and needs of children in each diagnostic category.

Research comparing PDD-NOS to other ASDs defined by the DSM-IV-TR is extensive and reveals differing perspectives. Several studies proposed that the ASD diagnoses represented a spectrum of symptom severity (Buitelaar et al. 1999; Fein et al. 1999; Prior et al. 1998; Stevens et al. 2000). This conceptualization of ASD, now embraced by the DSM-5, argued that each disorder varied only by the severity of a child's autism related symptoms. Thus,

under this interpretation, PDD-NOS did not differ qualitatively from other ASDs. A second perspective, however, suggested that the PDD-NOS profile differed distinctly from other ASDs, indicating that PDD-NOS may not have fit neatly along the proposed continuum of symptom severity (Paul et al. 2004; Walker et al. 2004). One study suggested that children with PDD-NOS often demonstrated stronger cognitive and adaptive functioning than children with Autistic Disorder, had histories of language delays uncommon in Asperger's Syndrome, and exhibited repetitive and stereotyped behaviors less frequently than children with either Autistic Disorder or Asperger's Syndrome (Walker et al. 2004).

In contrast to the number of studies examining the boundaries between ASD disorders as a whole, only two studies looked within a sample of children diagnosed with PDD-NOS in order to detect subgroups and further define the characteristics of these children. Walker and colleagues (2004) conducted a qualitative assessment of a small sample of children diagnosed with PDD-NOS (Mean age=86.3 months, SD=38 months) and identified three groups that emerged from their data. The first group (n=11) demonstrated few repetitive and stereotyped behaviors, were described as cognitively "high functioning," and had a "transient or persistent language delay." The second group (n=5) exhibited numerous repetitive and stereotyped behaviors, yet had "good" current language skills. The authors hypothesized that this group might have met criteria for Asperger's Disorder, except for a mild language delay earlier in development. Finally, the third group (n=5) was characterized as being potentially "too young or too delayed" to effectively assess for repetitive or stereotyped behaviors. The authors also posited that these children might have presented with a late age of onset for Autistic Disorder (Walker et al. 2004).

A study conducted by Mandy and colleagues (2011) looked at a sample of children diagnosed with PDD-NOS (N=256, Mean age 9.1 years) and grouped them according to DSM-IV-TR symptomatology. Their results indicated that 97% of children with PDD-NOS in their sample presented with a combination of social interaction and communication impairments, while only 3% presented with the combination of social interaction deficits and repetitive or stereotyped behaviors (Mandy et al. 2011). These findings suggested that a majority of children with PDD-NOS presented with difficulties in communication skills in addition to deficits in reciprocal social interaction, but very few of these children demonstrated repetitive and stereotyped behaviors. Both of these studies suggest that repetitive or stereotyped behaviors may appear later in development or might not be consistently observed. This finding has important implications for the identification of ASD in young children with less severe autism symptomatology.

Outcomes of Children with PDD-NOS

Several studies have found that children with PDD-NOS were more likely than those with other ASD diagnoses to achieve "optimal outcomes" as they grew older. An "optimal outcome" refers to children who were diagnosed with an ASD at an early age but no longer demonstrate ASD symptoms when reevaluated later in development. A study by Sutera and colleagues (2007) found that 39% of their sample of 11 children diagnosed with PDD-NOS at approximately 2 years of age no longer met criteria for an ASD by the time they were 4-years-old. This rate of achieving "optimal outcomes" by children with PDD-NOS was

significantly greater than that of children diagnosed with Autistic Disorder at 11% (Sutera et al. 2007). This pattern of children with PDD-NOS no longer meeting criteria for an ASD at age 4 has been demonstrated in several other studies (Berry 2009; Helt et al. 2008; Lord et al. 2006). Berry and colleagues (2009) looked at diagnostic outcomes for children diagnosed with PDD-NOS at age 2 when re-evaluated at age 4 and found several factors to be predictive of “optimal outcomes” at age 4 (Berry 2009). These included better motor abilities early in development as reported by the parents, low symptom severity at initial diagnosis, presence of few repetitive behaviors, higher adaptive skills as measured by parent-report, and higher expressive language abilities on a developmental assessment measure (Berry 2009). These findings suggest that there may be patterns of characteristics within PDD-NOS that might provide information about potential future outcomes.

Specific Aims

Identifying subgroups of children within the DSM-IV-TR’s PDD-NOS population may enhance our ability to identify, understand, and provide services for these children. It may also provide valuable information about a group of children who may be affected by the modifications to ASD diagnoses under the new DSM-5 criteria.

The current study sought to examine the characteristics of a sample of children diagnosed with PDD-NOS at approximately 2 years of age. The study’s specific aim was to identify more homogeneous and clinically meaningful subgroups within a sample of children diagnosed with PDD-NOS in the hope that those subgroups would have predictive validity for future diagnosis. This aim was addressed through: (a) utilization of a hierarchical cluster analysis to detect clusters in the current sample, (b) description of the characteristics within the subgroups detected by the cluster analysis, (c) determination of the predictive validity of subgroups by demonstrating differential outcomes based on the diagnosis received when the children were reevaluated at age 4, and (d) external validation of the subgroups using variables not included in the cluster analysis.

The following hypotheses were made concerning the outcomes of this analysis. First, given the common suggestion that PDD-NOS was a “catchall” diagnosis, we predicted that the characteristics of the subgroups detected by the cluster analyses at age two would show a varied, or atypical profile, meaning that the children in each subgroup would present with a profile that differed across multiple domains, rather than along a spectrum of symptom severity. Second, emphasis was placed on the predictive value of the clusters. Patterns of behaviors in children with PDD-NOS at age 2 might provide important information about their potential developmental course. It was therefore hypothesized that subgroup membership at age 2 would be related to diagnostic outcome at age 4.

Methods

Participants

Participants were selected from a larger sample of children taking part in an ongoing study examining the effectiveness of screening questionnaires designed to detect ASD symptoms in young children, which included the Modified Checklist for Autism in Toddlers (M-CHAT;

Robins, Fein, Barton, & Green 2001) and a more recent, amended version, the M-CHAT-Revised with Follow Up (M-CHAT-R/F; Robins, Casagrande, Barton, Chen, Dumont-Mathieu, & Fein 2014). This study was conducted under the supervision of the University of Connecticut's Institutional Review Board. Pediatrician's staff provided parents the opportunity to participate in our study during their child's 18 and 24 month well-child visits. Participants were given an information sheet detailing a waiver of signed consent at the fourth to fifth grade reading level. Parents were free to decline participation at any time.

Within this larger sample of children, 123 diagnosed with PDD-NOS between the ages of 18–34 months (Mean age = 25.5 months, SD=4.39) were included in the current study. Of the 123 children with PDD-NOS, 20 children were excluded from the analyses due to missing data. One child was excluded because he was the only child who received the Module 2 version of the Autism Diagnostic Observation Schedule (ADOS). The sample of 102 children consisted of predominantly male (n=78, 76%) toddlers. Of these children, 78% were reported as being White (n=80, 78%), with 7% Hispanic (n=7), 5% Black or African American (n=5), 5% Asian or Pacific Islander (n=5), 2% Biracial (n=2), and 1% described as “other” (n=2; See Table 1). Two children were missing Race or Ethnicity data.

Of the 102 children diagnosed with PDD-NOS, 71 (70%) received a re-evaluation between the ages of 48–64 months of age, as part of the larger study protocol. Thirteen (18%) of these 71 children were excluded from analyses because of missing data. As a result, 58 of the 71 children (Mean age = 51.1, SD = 6.98) were included in our second series of analyses aimed at determining the predictive value of the clusters by looking at diagnostic outcome at age 4. The demographic characteristics of this second sample (n=58) were similar to the initial sample, with 76% of these children being male (n=44). In terms of Race/Ethnicity, 83% of the children were described as White (n=48), 7% Hispanic (n=4), 5% Black or African American, 3% Asian or Pacific Islander (n=2), and one child was missing this data (See Table 1).

Procedures

Completed M-CHAT or M-CHAT-R/F screeners were sent to the University of Connecticut (UConn) Early Detection laboratory for scoring. If the child failed (screened positive on) the M-CHAT or M-CHAT-R/F, caregivers were called to confirm items missed. Children who continued to fail the screener after the follow-up phone interview were invited for a free developmental and diagnostic evaluation at UConn conducted by a graduate student and a licensed clinical psychologist or developmental pediatrician. The study then invited participants to return for a re-evaluation at approximately 4 years of age. The diagnosis of an ASD was assigned based upon the clinical judgment of experienced clinicians, using scores from the ADOS, Autism Diagnostic Interview (ADI), Childhood Autism Rating Scale (CARS), and developmental and adaptive behavior measures and according to DSM-IV-TR criteria for an ASD or PDD-NOS diagnosis.

Measures

All measures have been used extensively in clinical practice and research in order to detect and diagnose ASD in young children, and are considered to have strong psychometric

properties (Kleinman et al. 2007; Lord et al. 2000; Mullen 1994; Sparrow, Cicchetti, & Balla 2005).

M-CHAT and M-CHAT-R/F—The M-CHAT is a 23-item parent-report measure with yes/no questions designed to detect ASD symptoms in young children (Robins, Fein, Barton & Green 2001). This screening measure was adapted from the Checklist for Autism in Toddlers (CHAT; Baron-Cohen, Allen, & Gillberg 1992), in order to tailor the questionnaire to be appropriate for a parent-report format (Kleinman et al. 2007). Internal consistency was found to be sufficient for the complete screener and for six critical items (Cronbach's alpha values = .85 and .84, respectively) in a recent replication study (Kleinman et al. 2007). The M-CHAT-R/F is an updated and recently validated version of the M-CHAT yielding robust psychometric properties (Robins et al., 2014). Sensitivity and specificity were estimated to be 0.854 and 0.003 respectively, with positive predictive value (PPV) for receiving an ASD diagnosis of 0.475 (Robins et al., 2014). Both the M-CHAT and M-CHAT-R have a demonstrated history of effectively identifying children at risk for autism spectrum disorders, they are therefore considered comparable means of participant recruitment for the current sample. Further, screen positive rates and positive predictive power are similar for both measures, thus they are comparable for present purposes (Kleinman et al. 2007, Robins et al., 2014).

ADOS—The ADOS is a semi-structured, play-based interview that has been standardized for the purpose of diagnosing individuals with ASD (Lord et al. 2000). The ADOS assesses individual performance within four domains: Communication, Reciprocal Social Interaction, Play, and Repetitive Behaviors. The algorithm for scoring the ADOS follows this domain structure and cut-off scores for an ASD diagnosis have been established in the Communication and Reciprocal Social Interaction domains (Lord et al. 2000). Interrater reliability (mean weighted kappas, $M\kappa_W$) was high for both Modules 1 and 2 ($M\kappa_W = .78$ and $M\kappa_W = .70$, respectively; Lord et al. 2000).

ADI/ADI-Revised (ADI-R)—Both the original version of the ADI and a modified version, the ADI-R, were used to aid in the diagnosis of ASD in the current sample. The ADI and ADI-R are semi-structured interviews for parents of children with ASD that assess autism symptomatology based on The International Classification of Diseases (ICD-10) and DSM-IV (Lord, Rutter, & Le Couteur, 1994). Both measures are for use with children who have a mental age over 2 and both have sound psychometric properties (Lord et al. 1994, Le Couteur et al. 1989). Interrater reliability for the ADI and ADI-R communication and social domains was high (κ_W ranging from .64–.97 and κ_W ranging from .62–.89, respectively; Le Couteur et al., 1989; Lord et al., 1994), as were the interrater reliability results for the restricted and repetitive behaviors and interests of both versions ($M\kappa_W = .70$ and κ_W ranging from .55 to .87, respectively; Le Couteur et al., 1989; Lord et al., 1994). Multiple versions of the ADI were utilized over the course of this study to ensure the most updated version of the measure was used for diagnostic purposes. Data from the ADI were not used in the current analyses except for diagnostic purposes.

Childhood Autism Rating Scale (CARS, Schopler, Reichler, & Renner 1980 CARS, Schopler, Reichler, & Renner 1988) a 15 item behavioral rating scale that measures the

presence and severity of ASD symptoms. Clinicians scored the CARS following the developmental evaluation based on direct observation from the evaluation, parent-report, and test results. This measure has been found to have strong interrater reliability ($M\kappa_W = .71$) and internal consistency (Cronbach's alpha values = .94; Schopler et al., 1988).

Mullen Scales of Early Learning (Mullen)—The Mullen is a cognitive assessment standardized for use with children from birth to 68 months, which consists of five subdomains: Gross Motor, Visual Perception, and Fine Motor, as well as Receptive and Expressive Language (Mullen, 1994). Each subdomain score is assigned a t-score, as well as age equivalents and percentile rank for ease of interpretation. Internal consistency for the measure is reported as being very satisfactory (.75 to .83) and the test re-test reliability was high for both younger and older children (.84 and .76 respectively; Mullen, 1994). Children in the current sample completed this measure at both time points.

Vineland Adaptive Behavioral Scales – Interview Edition (Vineland)—The Vineland is a parent-report measure designed to assess adaptive skills of children across four domains: Communication, Daily Living Skills, Socialization, and Motor Skills (Sparrow et al., 2005). Chronbach's alphas were computed for the domain scores and for the Adaptive Behavior Composite (ABC) score based on the internal-consistency reliabilities of the subdomains. All were found to be above .80 for the age ranges included in the Early Detection sample (Sparrow et al. 2005). Both the Vineland I and Vineland II were used over the course of this study to ensure that the most updated version at time of evaluation was used. The current study obtained scores on this measure for children at both time points.

Data Analytic Plan

Data analyses for the current study occurred in two phases, both of which utilized exploratory hierarchical cluster analyses to detect potential subgroups within a sample of 102 children diagnosed with PDD-NOS at age 2. Hierarchical methods are ideal for samples of this size in order to keep calculations feasible (Aldenderfer & Blashfield 1984; Hair & Black 2000). This procedure uses stepwise clustering methods to combine observations into subgroups using, in this case, agglomerative methods to assign observations to clusters (Hair & Black 2000). Agglomerative methods place each observation into individual clusters initially and, through a stepwise process, merge the most similar clusters together to create a new cluster; this process continues until all clusters form a single group (Hair & Black 2000). The agglomerative methods used here are described in the appendix.

Decisions about cluster numbers are dependent upon statistical considerations and upon theoretical and practical considerations (Hair & Black 2000). Given that current research suggests the PDD-NOS population is heterogeneous and ill-defined, and that the intent of this study was to better understand the characteristics of subtypes in order to guide diagnostic and treatment development, it was important to restrict the number of clusters to ensure the subgroups identified would be clinically relevant. It was decided that more than four clusters in a sample of 102 children would likely yield clusters with few observations in each and could potentially be less representative of the population as a whole.

The predictive value of the clusters produced by the hierarchical cluster analysis was also a key factor in determining the optimal cluster procedures. In order to establish ‘outcome,’ diagnostic data from 58 children included in the initial analyses, who received a reevaluation at age 4, were examined. These participants were grouped according to whether they retained a diagnosis of PDD-NOS, received a diagnosis of Autistic Disorder, or no longer met criteria for an ASD or Developmental Delay (DD) diagnosis at age 4. Children who no longer met criteria for an ASD were considered by the experimenters to have achieved ‘better’ outcomes and children who received a diagnosis of Autistic Disorder were considered to have had ‘poorer’ outcomes. The examiners then calculated the specificity and sensitivity, as well as the positive and negative predictive values, of the clusters to determine whether cluster assignment at age 2 provided information about having ‘better’ or ‘poorer’ outcomes at age 4.

Phase I—In the initial phase of the study, a hierarchical cluster analysis was conducted using Ward’s method to detect clusters in the current sample (Ward 1963). Variables included each individual item from the ADOS, Module 1, as well as each subdomain score from the Mullen. Because the scales differed across assessment tools, scores were standardized to allow for comparison between measures (Hair & Black 2000).

A three-cluster structure best fit the data (see Fig. 1 top and bottom panels for dendrogram and scree plot, respectively). However, this cluster structure was found to have insufficient predictive performance due to poor specificity (0.68) and negative predictive value (0.42). Therefore, the examiners reevaluated the variables included in the analyses, as it became evident from these results that the quality of the variables included was more important for predicting outcome than the quantity of the variables included. Evidence supporting this conclusion can be found in the literature surrounding cluster analyses. Researchers suggest that selection of variables for cluster analyses must have theoretical and practical foundations (Hair & Black 2000). More importantly, the literature suggests that only variables describing the observations to be clustered and that directly pertain to the particular aims of the analyses should be included. Including variables that are irrelevant to either of these premises can mask the underlying cluster structure that exists, making identification of these clusters extremely difficult (Milligan & Hirtle 2003; Hair & Black 2000).

Phase II—In light of these findings, a second hierarchical cluster analysis was conducted, again using Ward’s method, in which only variables that provided predictive information about participants at age 4 were included (Ward 1963). The level of predictive performance demonstrated by each variable was determined by plotting each item used in the original analyses against the reevaluation diagnosis of ‘ASD,’ a group that included children who retained their PDD-NOS diagnosis or received an Autistic Disorder diagnosis at age 4, or ‘No ASD/No DD’ at age 4. The examiners then selected the item from each subdomain from the ADOS Module 1 and the subdomain score from the Mullen that best differentiated between groups at age 2 based on diagnostic classification at age 4. The Visual Reception subdomain score from the Mullen and four items from the ADOS were selected, to include Item A6: Use of Other’s Body to Communicate, B7: Requesting, C1: Functional Play with Objects, and D4: Unusually Repetitive Interests or Stereotyped Behaviors. This provided a

list of variables collected during the children's first evaluation that offered the most information about their future diagnostic outcome at age 4.

Evaluation of the clusters—Typical statistical procedures, such as using an analysis of variance (ANOVA) to demonstrate that significant differences exist between clusters by using the variables included in the cluster analysis, are not valid means of evaluating the cluster structures. Instead, what is referred to as 'external' validation procedures are suggested (Milligan & Hirtle 2003; Hair & Black 2000): External validity can be established by conducting ANOVAs that utilize variables not included in the hierarchical cluster analysis. In the current study, a one-way multivariate analysis of variance (MANOVA) was used to establish the external validity of the selected clusters by comparing the groups' standardized scores on each item of the CARS, an autism symptom severity measure, and subdomain scores from the Vineland. Due to missing CARS and Vineland data, three persons were excluded from the external validity analyses; a total of 99 participants diagnosed with PDD-NOS at age 2 were included in the one-way MANOVA. Follow-up ANOVAs and appropriate post hoc tests were performed (e.g., if equal variances were not assumed, post hoc Games-Howell test was used; otherwise, post hoc Fisher's least significant difference [LSD] test was run). The alpha value was set at 0.05 for all statistical tests.

Results

Results from the Phase II hierarchical cluster analysis indicated that a three cluster structure best fits the data (see Fig. 3 for the dendrogram and scree plot produced by the analysis and corresponding cluster labels, and Table 2 for demographic information on each cluster). Cluster 1 (n=26) scored the highest of the three clusters on the Visual Reception subdomain (M = 38.65, SD = 11.5) from the Mullen. Children in this cluster also exhibited the least social and communicative impairments and the fewest repetitive behaviors, as indicated by their scores on the item from each subdomain of the ADOS used in the analysis. Cluster 2 (n = 68) demonstrated lower scores than Cluster 1 on the Visual Reception subdomain (M = 30.95, SD = 9), and presented with more social impairments and communication difficulties than Cluster 1 on the included ADOS items. Children in Cluster 2 also had the most repetitive and stereotyped behaviors and interests of all three clusters. Children in the third cluster (Cluster 3, n = 8) were considered 'atypical' and the most difficult to characterize. Children in Cluster 3 received the lowest scores of all three clusters on the Mullen Visual Reception subdomain (M = 27.25, SD = 10.14) and were the most impaired in areas of communication and social interaction on the ADOS. Despite exhibiting greater impairment within these domains, children in Cluster 3 demonstrated fewer repetitive behaviors than those in Cluster 2.

Mullen

The clusters demonstrated a consistent pattern across all subdomains of the Mullen (see Table 3). In addition to the Visual Reception subdomain, which was the only score from this measure used in Phase II of the cluster analysis, Cluster 1 continued to have the highest scores of each cluster in the remaining subdomains (Fine Motor and Receptive and

Expressive Language; $M = 35.23$, $SD = 12.7$, $M = 31.85$, $SD = 10.45$, and $M = 31.62$, $SD = 8.26$, respectively). Cluster 2's scores were between Clusters 1 and 3 for each of these clusters, though scores were more similar to Cluster 1 on the Fine Motor subdomain ($M = 32.46$, $SD = 9.31$) and closer to the lower scores found in Cluster 3 for the Receptive and Expressive Language subdomains ($M = 22.72$, $SD = 6.38$, $M = 28.03$, $SD = 8.18$, respectively), suggesting a more significant impairment in communication abilities in this cluster when compared to Cluster 1. Cluster 3 also had the lowest scores across all remaining Mullen subdomains ($M = 27.13$, $SD = 10.27$, $M = 20.75$, $SD = 2.12$, $M = 24.38$, $SD = 4.96$), which indicated the greatest cognitive impairment of all three clusters.

ADOS

ADOS A1, use of other's body to communicate—For the ADOS item examining a child's use of another person's body to communicate, results (see Fig. 3) indicated that Cluster 1 was the least likely to demonstrate this behavior, with 81% of the children in this cluster receiving a score of zero for “no use of another's body to communicate.” 57% of children in Cluster 2 received scores indicating the children used another person's hand to lead them to or reach for an item they desired to a mild or moderate degree, as indicated by a score of one or two. Most striking about the current results for this ADOS item, was that 100% of Cluster 3 children received a score of three, which is indicative of “little or no spontaneous communication,” suggesting severe communicative impairments.

ADOS B7, requesting—Results for the ADOS B7 item show that 65% of children in Cluster 1 were able to successfully request items using joint attention (see Fig. 4). The remaining 35% of the children in Cluster 1 received a mild score of 1, meaning these children used pointing to draw another's attention to an object, but their use of coordinated eye contact was not yet fluent enough for a score of zero. 84% of children in Cluster 2 demonstrated mild to moderate impairments in their ability to use joint attention to request objects and 63% of children in Cluster 3 demonstrated impairments in this skill area.

ADOS C1, functional play with objects—Cluster results for this item (see Fig. 5) indicated that the majority (92%) of children in Cluster 1 demonstrated unimpaired play skills, with 8% of the children in this cluster demonstrating mild impairment, as indicated by a score of one. Ninety-six percent of children in Cluster 2 demonstrated mild to moderate deficits this area, with only 4% of the children in this cluster being found to have no impairment in their play abilities. 88% of children in Cluster 3 received scores of one, two, or three, indicating mild to moderate impairment in this play domain.

ADOS D4, unusually repetitive interests or stereotyped behaviors—85% of children in Cluster 1 did not demonstrate any repetitive or stereotyped behaviors during the administration of the ADOS (see Fig. 6). In contrast, 53% of the children in Cluster 2 received a score indicating that these behaviors were present to either a mild or moderate degree. Children in cluster 3 were more similar to children in Cluster 1 than on any previous item included in the cluster analysis, as 63% of children in Cluster 3 did not demonstrate any repetitive or stereotyped behaviors during the ADOS. When these behaviors were present in a child from Cluster 3, they received a milder score of one (37%).

External Validity

In order to explore whether the cluster structure (three clusters) detected by the hierarchical cluster analysis remained consistent when compared using variables not included in the original cluster analyses, a one-way MANOVA was conducted on all 15 CARS items, the CARS Total Score, and four Vineland subdomain scores. A trend toward differences was found among the three clusters on the dependent measures, Wilks's $\Lambda = .59$, $F(38,156) = 1.25$, $p = .17$, $\eta_p^2 = .23$. One-way ANOVAs on all dependent variables were conducted as follow-up tests to the MANOVA, and post hoc tests of the significant ANOVAs were performed. ANOVAs and post hoc analyses revealed significant differences between clusters for seven items and for the CARS Total Score. Clusters 1 and 2 differed significantly on the following items: I. Relating to People, II. Imitation, V. Object Use, VIII. Listening Response, XI. Verbal Communication, XII. Nonverbal Communication, and XV. General Impressions, as well as CARS total score. For each item, Cluster 1's CARS scores indicated the least severe autism symptom presentation, as they were significantly lower than Cluster 2's scores, which were suggestive of the most severe autism symptomatology. Cluster 3's scores, though not significantly different from either those of Clusters 1 or 2, fell consistently between Clusters 1 and 2's scores (see Table 4 for a summary of the external validity results).

No significant differences were found between clusters on the Vineland.

Diagnosis at Age 4 By Cluster

Due to insufficient sample size for Pearson's chi-square test (three of the cells [50%] had expected frequencies that were less than five), Fisher's exact test was performed to determine if the cluster membership at age 2 would be related to diagnostic outcome at age 4. Results suggested that, as would be expected given that the variables included in the cluster analysis were selected on the basis of their predictive value, the clusters detected by the Phase II hierarchical cluster analysis were found to provide important information about outcome for children who received a reevaluation at age 4 (two-sided exact $p = .001$, Table 6)

. Fifteen of the children in Cluster 1 received a reevaluation (see Tables 5 and 6 and Fig. 7). This cluster contained the greatest number of children who went on to no longer meet criteria for an ASD ($n = 9$, 60%), with the other six children (40%) remaining stable in their PDD-NOS diagnosis. Within the children who no longer met criteria for an ASD, five received no diagnosis through the study (33.3%), three were diagnosed with DD (20%), and one was given a Developmental Language Disorder (DLD) diagnosis (6.7%). In Cluster 2, 38 children were reevaluated at age 4. The majority of these children continued to meet criteria for PDD-NOS ($n = 15$, 39%) or went on to meet criteria for a diagnosis of Autistic Disorder at age 4 ($n = 19$, 50%). Only 11% of the children ($n = 4$) in Cluster 2 did not demonstrate ASD symptoms at age 4. Two of these children received a DD diagnosis (5.3%), one received a DLD diagnosis (2.6%), and one received "no diagnosis" (2.6%). Despite the limited repetitive behaviors at age 2 in Cluster 3, the majority of the five children in this cluster who received a reevaluation went on to meet criteria for Autistic Disorder at age 4 ($n = 3$, 60%). One child (20%) continued to meet criteria for PDD-NOS and one child

(20%) no longer met criteria for an ASD at age 4, and instead was given a DD diagnosis. Again, as expected given the variables used in the cluster analysis, the sensitivity and specificity values calculated for this three cluster structure were high. In Table 7, sensitivity for outcome (ASD vs. No ASD/No DD) at age 4 was 100%, while specificity was 83%. The positive and negative predictive values, as well as accuracy, were similarly high (95%, 100%, and 96%, respectively), indicating that our clusters demonstrated a highly accurate ability to predict age 4 diagnosis using scores at age 2.

Discussion

The current study used hierarchical clustering procedures to detect subgroups within a sample of children diagnosed with PDD-NOS in an attempt to clarify the characteristics of a diagnosis that was portrayed as ‘problematic’ in the literature under the DSM-IV-TR, as well as to investigate the potential impact the new DSM-5 criteria may have upon some or all of these children. The three clusters produced by these analyses are described in detail in the subsequent paragraphs.

Spectrum of Symptom Severity

In part, the results of the current study support the perspective that PDD-NOS is characterized by a spectrum of symptom severity (See Table 8). Clusters 1 and 2 appeared to differ along a continuum. Cluster 1 represented the higher functioning end of the spectrum, as this cluster consisted of children who received the highest scores on each subdomain of the Mullen, demonstrated the least impairment on social and communication skills, and exhibited the fewest repetitive behaviors and interests, as measured by the ADOS. These children also had the lowest total scores on the measure of autism symptom severity. Unsurprisingly, a majority of the children in Cluster 1 no longer met criteria for an ASD or remained stable in their PDD-NOS diagnosis when reevaluated at age 4. Cluster 2 represented the lower functioning end of the symptom severity spectrum, with these children receiving lower scores on the Mullen and demonstrating more impairment in social and communication skill areas than children in Cluster 1. Children in Cluster 2 also engaged in more restricted and repetitive behaviors or interests than children in either Clusters 1 or 3 at age 2. Consistent with this profile, children in Cluster 2 either continued to meet diagnostic criteria for PDD-NOS or received a diagnosis of Autistic Disorder when reevaluated at age 4.

Evidence in the literature supports the current findings that link specific skill profiles in children to future outcome. A 2007 study suggested that children with higher cognitive scores and fewer early social interaction impairments demonstrate a greater ability to develop skills, such as receptive and expressive language, as well as play skills, over time (Ben-Itzhak & Zachor 2007). Further, this study found that both cognitive levels and social-reciprocity skills were significantly correlated with outcome later in development (Ben-Itzhak & Zachor, 2007). Studies have also indicated that motor skills, symptom severity at age 2, number of repetitive behaviors present, adaptive functioning, and expressive language skills are characteristic of children who no longer meet criteria for an ASD when reevaluated at age 4 (Berry 2009).

In the current study, the children in Cluster 1 confirmed earlier findings that children with PDD-NOS who receive higher scores on nonverbal problem solving measures, demonstrate fewer social interaction impairments, and present with fewer repetitive behaviors and less severe autism symptomatology may be more likely to have 'better' outcomes later in development (e.g. no longer meet criteria for an ASD). Children in Cluster 2 demonstrated that lower cognitive scores combined with greater social impairment, more frequent repetitive behaviors, and more severe autism symptoms predict the retention of PDD-NOS diagnoses or the development of a more severe diagnosis of Autistic Disorder by age 4.

Clusters 1 and 2 may indicate that the PDD-NOS population is less heterogeneous than has been previously described in the literature. The implication of these findings may be that, regardless of the lack of explicitly defined criteria, there seemed to be a somewhat consistent pattern of symptoms in a proportion of children diagnosed with PDD-NOS at age 2, and that this pattern varied largely in terms of severity.

Finally, a potential concern from the data might suggest that children in Cluster 1 would fail to meet criteria for a diagnosis of ASD under the DSM-5 criteria, and might instead be diagnosed with Social Pragmatic Disorder. It remains unclear if these children can be expected to attain the positive results noted in this sample, if they receive less intensive early intervention. Although collection of treatment data was not a focus of this project, all children who received an ASD diagnosis at age 2 were referred for services at that time, thus each child had received some intervention in the time between initial diagnosis and reevaluation at age 4.

Atypical Profile

The current findings also provide support; however, for the perspective that PDD-NOS represents a more atypical profile and does not fit neatly within the spectrum of symptom severity, especially in young children. Instead, children in Cluster 3's presentations differed qualitatively from the children in Clusters 1 and 2 at age 2 (see Table 9). At the time of their first evaluation, these children exhibited the most severe cognitive, social, and communicative impairments when compared to Clusters 1 and 2, yet they demonstrated far fewer repetitive and stereotyped behaviors than children in Cluster 2. These findings were not expected given that these children demonstrated deficits in the social and communication items from the ADOS at this early age. Also surprising was the fact that the children in Cluster 3 received lower severity scores on the CARS than children in Cluster 2, indicating milder autism symptom severity presentation at age 2, despite more marked impairments in the cognitive, interpersonal, and communicative domains. Again, despite these more mild autism severity scores at age 2, it was found 60% of the children in Cluster 3 who received a reevaluation at age 4 went on to develop Autistic disorder, suggesting that Cluster 3 children are more likely to have 'poorer' outcomes later in development.

Cluster 3 also demonstrated the greatest impairment in functional play skills, when compared to Clusters 1 and 2. Play skills have been found in the literature to be highly correlated with communication, cognitive, and social development in young children (Bateson 1955; Piaget 1962; Vygotsky 1978; Bates 1979; Rapin 1996). Toy play in particular is thought to be related to the development of joint attention skills (Toth, Munson,

Meltzoff, & Dawson 2006). In both high functioning and low functioning children with autism, the frequency with which they engage in toy play and the developmental level of this play have been found to be significantly lower than their non-autistic peers (Rapin 1996). More important, evidence suggests that toy play in preschool aged children diagnosed with autism has been found to be predictive of communication development over the next several years of development (Toth et al. 2006). Given the findings on the correlation between play skills and other important developmental areas, Cluster 3's profile of low cognitive scores and severe social and communicative impairments may lend further support to the interrelatedness of these developmental domains.

Cluster 3 represents a group of children with PDD-NOS who are potentially difficult to characterize and diagnose accurately at age 2. As noted previously, prior research indicated that motor skills, severity scores, number of repetitive behaviors, and play skills, among others, are variables found to be predictive of developmental outcomes in children with ASD (Sutera et al. 2007; Berry 2009). Lower functioning children with autism have been found to be more easily diagnosed at age 2, especially when the children present with higher nonverbal than verbal scores (Rapin 1996). Children in Cluster 3 exhibited this pattern of higher nonverbal problem solving scores than receptive and expressive language scores, and exhibited severe social and communication deficits. However, they did not present with the repetitive behaviors required for an Autistic Disorder diagnosis, and their scores on the CARS were also less severe compared to children in Cluster 2, perhaps because they did not present some of these behaviors. Despite their milder autism symptom presentation at age 2; however, a greater proportion of children in Cluster 3 went on to have 'poorer' outcomes and met criteria for Autistic Disorder at age 4.

It is possible that children in Cluster 3 were not yet demonstrating the repetitive behaviors required for a diagnosis of Autistic Disorder when first diagnosed at age 2. Research in repetitive behaviors has shown the number of repetitive behaviors exhibited by children at age 4 is often higher than was present in those children at age 2 (Moore & Goodson 2003; Cox, Klein, Charman, Baird, Baron-Cohen, Swettenham, Drew, & Wheelwright 1999; Stone, Lee, Ashford, Brissie, Hepburn, Coonrod, & Weiss 1999). At age 2 ASD specific impairments in social and communication skills may be apparent on the ADOS and CARS, but symptoms in the restricted and repetitive behaviors and interests domain may not have developed yet. This may be especially true of the very delayed children in Cluster 3, whose repetitive behaviors may make their appearance at a later developmental time.

Given the reliance on repetitive behaviors and restricted interests in the DSM 5 diagnostic criteria for autism spectrum disorder, these data have important implications for diagnosis of ASDs. It may be that children in Cluster 3, who developed repetitive behaviors later in development, will no longer qualify for an ASD diagnosis as toddlers, despite their potentially negative trajectory.

External Validity

A MANOVA was utilized to determine whether differences between the clusters existed on variables not included in the original cluster analyses. Results indicated Clusters 1 and 2 differed significantly on seven items from the CARS, as well as the total score from this

measure. Cluster 3 scores were not found to be significantly different from either cluster and remained consistently between Cluster 1 and 2 on each item and on the total score. The total scores for each cluster were above what has been found to be the most accurate cut off score for PDD-NOS on the CARS, which is 25.5 for 2-year-olds (Chlebowski, Green, Barton, & Fein 2010).

DSM-5

The DSM-5 includes several significant changes to the criteria found in the DSM-IV-TR to diagnose ASDs. These changes include combining the specific sub categories of PDD into one diagnosis of ASD and reducing the three symptom domains (Social Interaction, Communication, and Restricted, Repetitive, and Stereotyped Behaviors) to two (A. Social Communication & B. Restricted, Repetitive Behaviors; RRB; Mandy, Charman, & Skuse 2012). In order to receive a diagnosis of ASD, a child must demonstrate symptoms from both symptom domains. Meeting criteria for the symptoms described in Criteria A requires a child demonstrate one symptom in all three of the symptom subdomains, which include A1 (Social-Emotional Reciprocity), A2 (Nonverbal Communication), and A3 (Relationships). Meeting criteria for the symptom cluster defined by Criteria B requires a child to have demonstrated one symptom from at least two symptom subdomains, including B1 (Stereotyped or Repetitive Speech, Motor Mannerisms, or Use of Objects), B2 (Excessive Adherence to Routines or Ritualized Speech), B3 (Restricted, Fixated Interests), or B4 (Hyper- or Hypo-reactivity to Sensory Input or Unusual Sensory Interests).

Children in two of the clusters detected in the current study, Clusters 1 and 3, did not demonstrate consistent repetitive and stereotyped behaviors at age 2. These findings are consistent with other studies investigating the characteristics of PDD-NOS samples. Walker et al.'s (2004) study found that 50% of their sample demonstrated only mild or transient repetitive behaviors or interests, while Mandy et al. (2011) found that 97% of their sample of toddlers (n=66) did not demonstrate these behaviors. Without the repetitive and stereotyped behaviors or interests, it is unclear whether young children would meet criteria for an ASD diagnosis under the DSM-5 criteria. It is therefore imperative that future researchers continue to understand the trajectories of children who do not present with consistent repetitive and stereotyped behaviors at 2 years of age in order to ensure that children with significant social and communicative delays can access the autism specific early intervention services needed. Limiting access to such services when autism symptoms are present to either a mild or moderate degree, as found in Clusters 1 and 3, would likely have significant impacts on children's outcome at age 4, though further research will be necessary to support this claim.

Limitations

Several limitations should be considered when interpreting the results of the current study. First, of the 102 children who received a diagnosis of PDD-NOS at age 2, only 58 were reevaluated at the approximate age of 4, as a result of caregivers being unavailable for contact, having relocated, or refusing the evaluation. Therefore, the results relating to outcome were based on a subset of children included in the original cluster analyses. However, the percentages of children from each cluster that received a reevaluation were

roughly equal, indicating that there was no differential attrition. This provided almost equal access to data on outcome for each cluster (See Table 2).

The current study was unable to examine the potential effects of interventions children received in the years between evaluations at ages 2 and 4. Services received by children during this critical time period in their development might have provided meaningful information about diagnostic stability; however, these data were unavailable.

Finally, the number of children in Cluster 3 was extremely small, thus limiting our ability to draw generalizable conclusions from this data. However, it should be noted that during the Phase II cluster analyses, when the number of clusters was expanded to include four total clusters or contracted to include only two clusters, Cluster 3 remained a distinct group while the configurations of Clusters 1 and 2 changed. It was therefore determined that Cluster 3 represented a discrete cluster with characteristics that differed notably from the other possible clusters.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are extremely grateful to all the children and families for their participation and to the pediatrician sites, as well as Birth-to-Three providers within Connecticut and Massachusetts, who assisted and continue to assist our study. We would like to extend sincere thanks to members of the Early Detection Advisory Board, especially Thyde Dumont-Mathieu, M.D., Ho-Wen Hsu, M.D., and Mark Greenstein, M.D., for their sage advice and support. We would also like to thank the clinicians, graduate students, undergraduate research assistants, and research assistants whose work has been invaluable to the study. This study is supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (RO1 HD039961).

References

- Aldenderfer, MS., Blashfield, RK. Cluster analysis. Beverly Hills; London: Sage; 1984.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-IV-TR. American Psychiatric Publishing, Inc; 2000.
- American Psychiatric Association. DSM 5. American Psychiatric Association; 2013.
- Baron-Cohen S, Allen J, Gillberg C. Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *The British Journal of Psychiatry*. 1992; 161(6):839–843. [PubMed: 1483172]
- Bateson G. A theory of play and fantasy. *Psychiatric research reports*. 1955; 2(39):39–51.
- Ben-Itzhak E, Zachor DA. The effects of intellectual functioning and autism severity on outcome of early behavioral intervention for children with autism. *Research in Developmental Disabilities*. 2007; 28(3):287–303. [PubMed: 16730944]
- Berry, LN. Unpublished doctoral dissertation. University of Connecticut; Storrs, Connecticut: 2009. Early treatments associated with optimal outcome in children with autism spectrum disorders.
- Buitelaar JK, Van der Gaag R, Klin A, Volkmar F. Exploring the Boundaries of Pervasive Developmental Disorder Not Otherwise Specified: Analyses of Data from the DSM-IV Autistic Disorder Field Trial. *Journal of Autism and Developmental Disorders*. 1999; 29(1):33–43. [PubMed: 10097993]
- Chakrabarti S, Fombonne E. Pervasive Developmental Disorders in Preschool Children: Confirmation of High Prevalence. *The American Journal of Psychiatry*. 2005; 162(6):1133–1141. [PubMed: 15930062]

- Cicchetti D, Sparrow S. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*. 1981; 86:127–137. [PubMed: 7315877]
- Chlebowski C, Green JA, Barton ML, Fein D. Using the Childhood Autism Rating Scale to Diagnose Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*. 2010; 40(7):787–799. [PubMed: 20054630]
- Cox A, Klein K, Charman T, Baird G, Baron-Cohen S, Swettenham J, Drew A, Wheelwright S. Autism spectrum disorders at 20 and 42 months of age: Stability of clinical and ADI-R diagnosis. *Journal of Child Psychology and Psychiatry*. 1999; 40(5):719–732. [PubMed: 10433406]
- Dougherty, G. *Pattern Recognition and Classification*. New York: Springer; 2013. (Agglomerative Hierarchical Clustering; p. 150-154.
- Fein D, Stevens M, Dunn M, Waterhouse L, Allen D, Rapin I, Feinstein C. Subtypes of Pervasive Developmental Disorder: Clinical Characteristics. *Child Neuropsychology*. 1999; 5(1):1–24.
- Frazier TW, Youngstrom EA, Speer L, Embacher R, Law P, Constantino J, Findling RL, Hardan AY, Eng C. Validation of Proposed DSM-5 Criteria for Autism Spectrum Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2012; 51(1):28–40. [PubMed: 22176937]
- Helt M, Kelley E, Kinsbourne M, Pandey J, Boorstein H, Herbert M, Fein D. Can Children with Autism Recover? If So, How. *Neuropsychology Review*. 2008; 18(4):339–366. DOI: 10.1007/s11065-008-9075-9 [PubMed: 19009353]
- Kleinman JM, Robins DL, Ventola PE, Pandey J, Boorstein HC, Esser EL, Wilson LB, et al. 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*. 2007; 38(5):827–839.
- Le Couteur A, Rutter M, Lord C, Rios P, Robertson S, Holdgrafer M, McLennan J. Autism Diagnostic Interview: A Standardized Investigator-Based Instrument. *Journal of Autism and Developmental Disorders*. 1989; 19(3):363–387. [PubMed: 2793783]
- Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Pickles A. Autism From 2 to 9 Years of Age. *Archives of General Psychiatry*. 2006; 63:694–701.
- Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, Pickles A, et al. 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*. 2000; 30(3)
- Lord C, Rutter M, Le Couteur A. 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*. 1994; 24(5):659–685. [PubMed: 7814313]
- Mandy WPL, Charman T, Skuse DH. Testing the Construct Validity of Proposed Criteria for DSM-5 Autism Spectrum Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2012; 51(1):41–50. [PubMed: 22176938]
- Mandy W, Charman T, Gilmour J, Skuse D. Toward specifying pervasive developmental disorder-not otherwise specified. *Autism Research*. 2011; 4(2):121–131. [PubMed: 21298812]
- Milligan GW. An Examination of the Effect of Six Types of Error Perturbation On Fifteen Clustering Algorithms. *Psychometrika*. 1980; 45(3):325–342.
- Milligan GW, Hirtle SC. Clustering and classification methods. *Handbook of psychology*. 2003
- Moore V, Goodson S. 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*. 2003; 7(1):47–63. [PubMed: 12638764]
- Mullen, EM. *The Mullen Scales of Early Development*. Circle Pines, MN: American Guidance Services; 1994.
- Paul R, Miles S, Cicchetti D, Sparrow S, Klin A, Volkmar F, Coflin M, Booker S. Adaptive Behavior in Autism and Pervasive Developmental Disorder-Not Otherwise Specified: Microanalysis of Scores on the Vineland Adaptive Behavior Scales. *Journal of Autism and Developmental Disorders*. 2004; 34(2):223–228. [PubMed: 15162940]
- Piaget, J. *Play, dreams and imitation*. Vol. 24. New York: Norton; 1962.

- Prior M, Eisenmajer R, Leekam S, Wing L, Gould J, Ong B, Dowe D. Are There Subgroups within the Autistic Spectrum? A Cluster Analysis of a Group of Children with Autistic Spectrum Disorder. *Journal of Child Psychology and Psychiatry*. 1998; 39(6):893–902. [PubMed: 9758197]
- Rapin, I., editor. *Preschool children with inadequate communication*. Vol. 139. Mac Keith Press; 1996.
- Robins DL, Casagrande K, Barton M, Chen CMA, Dumont-Mathieu T, Fein D. Validation of the Modified Checklist for Autism in Toddlers, Revised With Follow-up (M-CHAT-R/F). *Pediatrics*. 2014; 133(1):37–45. [PubMed: 24366990]
- Robins DL, Fein D, Barton ML, Green JA. 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*. 2001; 31(2):131–143. [PubMed: 11450812]
- Roux S, Garreau B, Barthélémy C, Hameury L. Implementation of a bioclinical database for research and treatment studies in childhood autism: Preliminary report on a concrete experience. *Developmental brain dysfunction*. 1994
- Schopler, E., Reichler, R., Rochen Renner, B. *The childhood autism rating scale*. Western Psychological Services; 1988.
- Stevens MC, Fein DA, Dunn M, Allen D, Waterhouse LH, Feinstein C, Rapin I. Subgroups of Children With Autism by Cluster Analysis: A Longitudinal Examination. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2000; 39(3):346–352. [PubMed: 10714055]
- Stone WL, Lee EB, Ashford L, Brissie J, Hepburn SL, Coonrod EE, Weiss BH. Can autism be diagnosed accurately in children under 3 years? *Journal of Child Psychology and Psychiatry*. 1999; 40(2):219–226. [PubMed: 10188704]
- Sutera S, Pandey J, Esser EL, Rosenthal MA, Wilson LB, Barton M, Green J, et al. Predictors of Optimal Outcome in Toddlers Diagnosed with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*. 2007; 37(1):98–107. [PubMed: 17206522]
- Sparrow, SS., Cicchetti, DV., Balla, DA. *The Vineland Adaptive Behavior Scales-II, Second Edition*. Circle Pines, MN: American Guidance Service, Inc; 2005. p. 109-117.
- Toth K, Munson J, Meltzoff NA, Dawson G. Early Predictors of Communication Development in Young Children with Autism Spectrum Disorder: Joint Attention, Imitation, and Toy Play. *Journal of Autism and Developmental Disorders*. 2006; 36(8):993–1005. DOI: 10.1007/s10803-006-0137-7 [PubMed: 16845578]
- United States Census Bureau. *US Census Bureau: State and County QuickFacts*. Connecticut: 2012. Retrieved from <http://quickfacts.census.gov/qfd/states/09000.html>
- Vygotsky, LLS. *Mind in society: The development of higher psychological processes*. Harvard university press; 1978.
- Walker DR, Thompson A, Zwaigenbaum L, Goldberg J, Bryson SE, Mahoney WJ, Strawbridge CP, et al. Specifying PDD-NOS: A Comparison of PDD-NOS, Asperger Syndrome, and Autism. *Journal of The American Academy of Child and Adolescent Psychiatry*. 2004; 43(2):172–180. [PubMed: 14726723]
- Ward JH. Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*. 1963; 58(301):236–244.

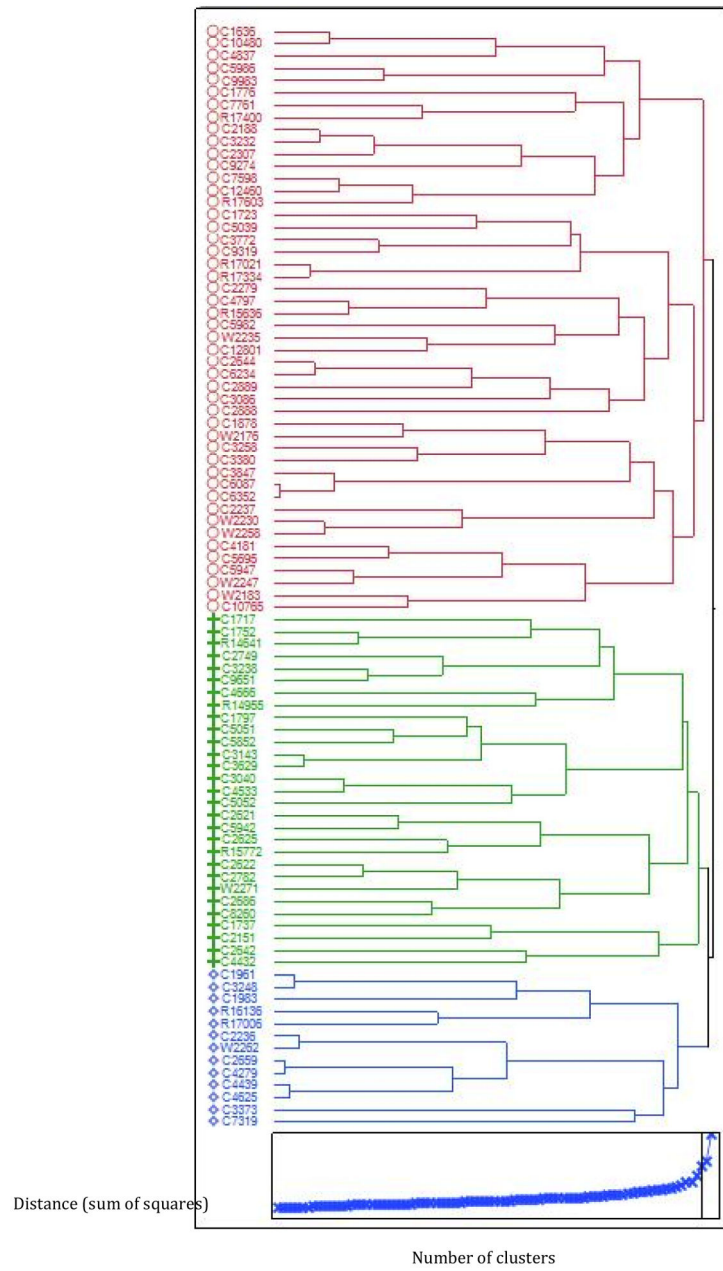


Figure 1.
Methods: Phase I cluster analysis dendrogram

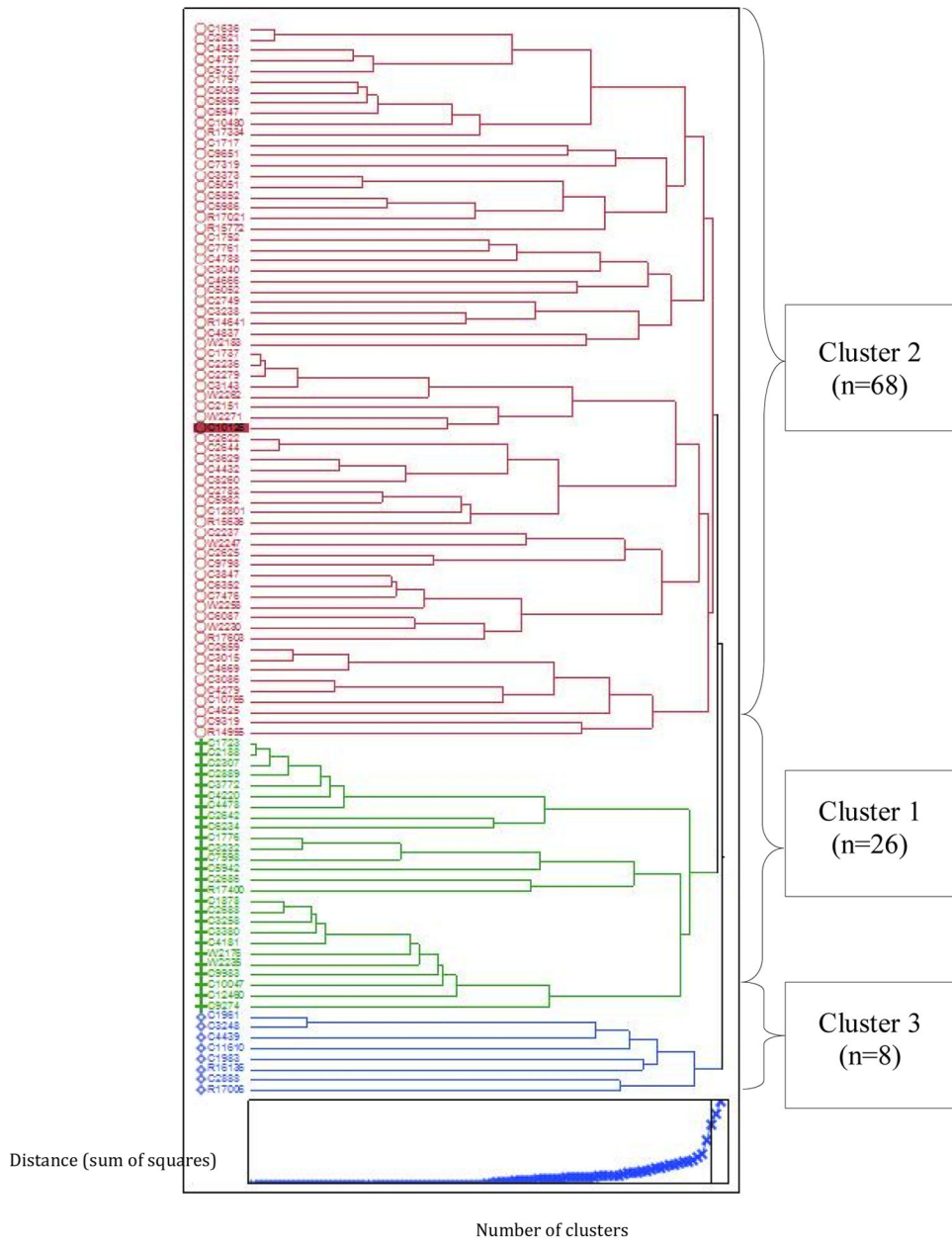


Figure 2.
Results: Phase II cluster analysis dendrogram and scree plot

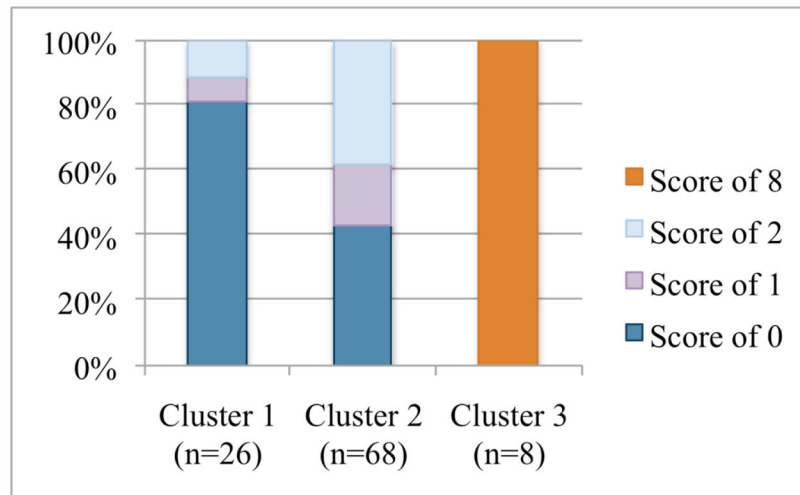


Figure 3.
ADOS item A6 “Use of Other’s Body to Communicate” scores by cluster.
0 = No use of another’s body to communicate
1 = Takes another person’s hand and leads him/her places without coordinated gaze
2 = Placement of another person’s hand or other body part on object
8 = Little or no spontaneous communication

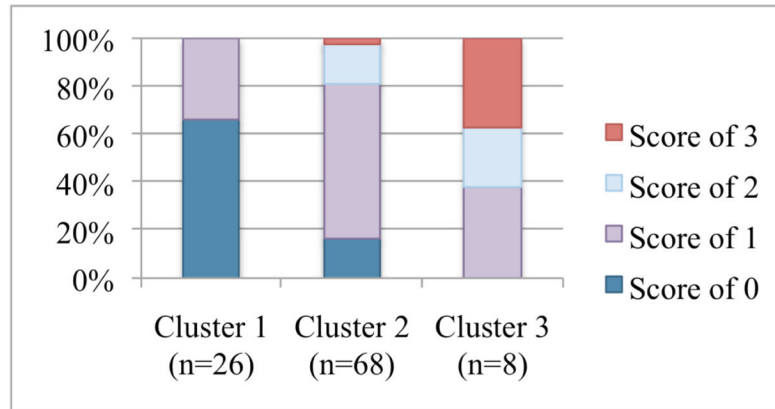


Figure 4.

ADOS Item B7 "Requesting" scores by cluster

0 = Points with index finger using coordinated gaze to object and person

1 = Using pointing to reference objects, without sufficient flexibility or frequency for '0'

2 = Points to objects when close to or touching object, no coordinated eye gaze or vocalization

3 = Does not point to objects in any way

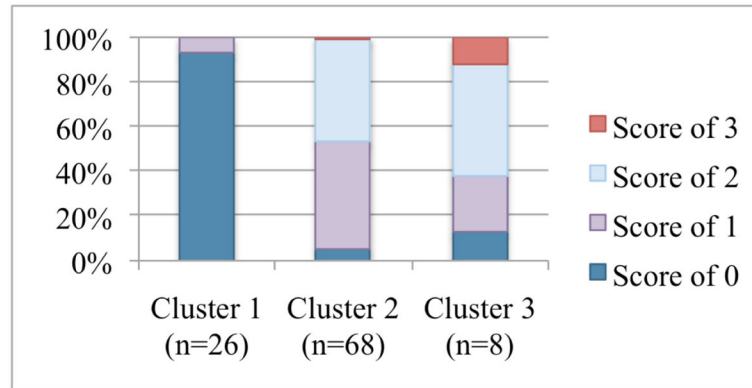


Figure 5.
 ADOS item C1 “Functional Play with Objects” scores by cluster
 0 = Spontaneously and appropriately plays with variety of toys
 1 = Some spontaneous functional play with cause-and-effect toys with at least 1 miniature
 2 = Plays appropriately with cause-and-effect toys only, and/or pushing car
 3 = No play with toys or only stereotyped play

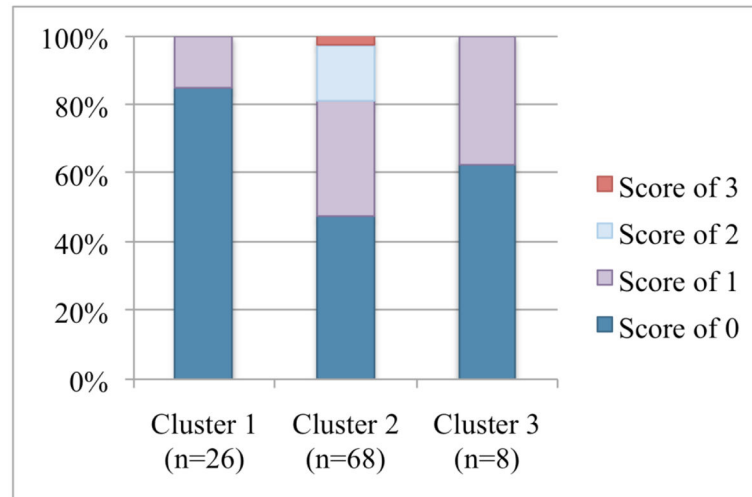


Figure 6.
 ADOS item D4 “Unusually Repetitive Interests or Stereotyped Behaviors” scores by cluster
 0 = No repetitive or stereotyped behaviors during the ADOS evaluation
 1 = An interest or behavior that is repetitive or stereotyped to an unusual degree
 2 = Repetitive or stereotyped interests and/or behaviors are minority of child’s interests or behaviors
 3 = Repetitive or stereotyped interests and/or behaviors form majority of child’s interests

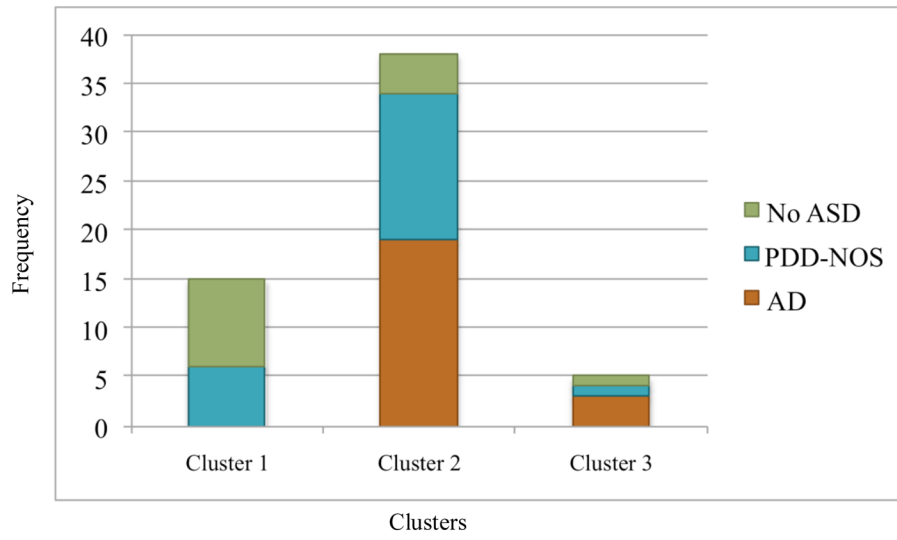


Figure 7.
Diagnosis at age four by cluster

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Sample Demographics

Sample	N	Mean Age in Months (SD)	Gender (Ratio)	Race/Ethnicity
<i>Age 2</i>	102	25.5 (4.39)	Males = 78 Females = 24 (3.25:1)	White (n=80) Hispanic/Latino (n=7) Black or African American (n=5) Asian or Pacific Islander (n=5) Biracial (n=2) Other (n=1) Missing (n=2)
<i>Age 4</i>	58	51.1 (6.98)	Males = 44 Females = 14 (3.14:1)	White (n=48) Hispanic/Latino (n=4) Black or African American (n=3) Asian or Pacific Islander (n=2) Missing (n=1)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Demographic Information of Clusters

Demographic Information	Cluster 1	Cluster 2	Cluster 3
n	26	68	8
Mean Age in Months (SD)	25.6 (4.71)	26.2 (4.44)	23.71 (3.31)
Gender (Ratio)	Males = 17 Females = 9 (1.89:1)	Males = 55 Females = 13 (4.23:1)	Males = 61 Females = 2 (3:1)
Race/Ethnicity	White (n=20) Hispanic/Latino (n=2) Asian or Pacific Islander (n=0) Black or African American (n=1) Biracial (n=1) Other (n=0) Missing (n=2)	White (n=54) Hispanic/Latino (n=4) Asian or Pacific Islander (n=5) Black or African American (n=4) Biracial (n=0) Other (n=1) Missing (n=0)	White (n=6) Hispanic/Latino (n=1) Asian or Pacific Islander (n=0) Black or African American (n=0) Biracial (n=1) Other (n=0) Missing (n=0)
Received Reevaluation at Age 4	15 (58%)	38 (55%)	5 (63%)

Table 3

Average Mullen Scales of Early Learning (Mullen) T Scores by Cluster

Cluster	<i>Mean Mullen Subdomain Scores (SD)</i>			
	Visual Processing*	Fine Motor	Receptive Language	Expressive Language
Cluster 1 (n=26)	38.65 (11.5)	35.23 (12.7)	31.85 (10.45)	31.62 (8.26)
Cluster 2 (n=68)	30.95 (9)	21.083 (3.53)	22.67 (6.59)	29.65 (8.38)
Cluster 3 (n=8)	27.25 (10.14)	27.13 (10.27)	20.75 (2.12)	24.38 (4.96)

* Indicates item included in Phase II of cluster analyses

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4
 External Validity Results: Post Hoc Analyses, Significant Childhood Autism Rating Scale (CARS) Items

CARS Item	Cluster 1 Mean (n=25)	Cluster 2 Mean (n=66)	Cluster 3 Mean (n=8)	Follow-up ANOVA and p value	Post hoc test and p value
I. Relating to People	1.940 ²	2.333 ¹	2.188	.015	LSD, .004
II. Imitation	1.740 ²	2.220 ¹	2.125	.007	LSD, .002
V. Object Use	1.580 ²	1.924 ¹	1.625	.005	LSD, .002
VIII. Listening Response	1.760 ²	2.114 ¹	2.000	.032	LSD, .009
XI. Verbal Communication	2.360 ²	2.674 ¹	2.313	.024	LSD, .017
XII. Nonverbal Communication	1.980 ²	2.326 ¹	2.188	.012	Games-Howell, .009
XV. General Impressions	1.820 ²	2.136 ¹	1.875	.011	LSD, .005
CARS Total Score	26.420 ²	29.455 ¹	27.438	.002	LSD, .001

Note: Superscripts indicate significant differences between clusters.

Table 5

Diagnoses at Age Four by Cluster

Cluster	<i>Diagnoses</i>				
	Autistic Disorder (AD)	PDD-NOS	No Diagnosis	Developmental Delay (DD)	Developmental Language Delay (DLD)
Cluster 1 (n=15)	0	6 (40%)	5 (33.3%)	3 (20%)	1 (6.7%)
Cluster 2 (n=38)	19 (50%)	15 (39.5%)	1 (2.6%)	2 (5.3%)	1 (2.6%)
Cluster 3 (n=5)	3 (60%)	1 (20%)	0	1 (20%)	0

Table 6

Diagnoses at Age Four by Cluster – ASD vs. No ASD

Cluster	<i>ASD vs. No ASD Diagnosis</i>	
	ASD	No ASD
Cluster 1 (n=15)	6 (40%)	9 (60%)
Cluster 2 (n=38)	34 (89.5%)	4 (10.5%)
Cluster 3 (n=5)	4 (80%)	1 (20%)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 7

Phase II Cluster Analysis Results: Sensitivity/Specificity Estimates for Diagnosis at Age Four

		<i>Diagnosis at Re-Evaluation (Gold Standard)</i>		
		Autistic Disorder	No ASD/No DD	Total
Cluster Analysis (Test)	Autistic Disorder (1)	21	1	22
	No ASD/No DD (2)	0	5	5
	Total	21	6	27
Sensitivity	1			
Specificity	0.8333			
Positive Predictive Value	0.9545			
Negative Predictive Value	1			
Accuracy	0.963			

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 8

Spectrum of Symptom Severity: Clusters 1 and 2

Dimension	Cluster 1 (n=26)	Cluster 2 (n=68)
Cognitive Functioning	High	Low
Social & Communication Deficits	Low	High
Repetitive Behaviors	Low	High
CARS Total Score	Low	High
Age 4 Diagnosis	Most Likely No-ASD/No-DD	PDD-NOS or AD

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 9

Atypical Profile: Cluster 3

Domain	Cluster 3 (n=8)
Cognitive Functioning	Low
Social & Communication Deficits	High
Repetitive Behaviors	Low
CARS Total Score	Mild
Age 4 Diagnosis	Most Likely AD

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript