

Behavioral and Emotional Problems in Young People with Pervasive Developmental Disorders: Relative Prevalence, Effects of Subject Characteristics, and Empirical Classification

Luc Lecavalier

Published online: 2 August 2006
© Springer Science+Business Media, Inc. 2006

Abstract Parents or teachers rated 487 non-clinically referred young people with Pervasive Developmental Disorders on the Nisonger Child Behavior Rating Form. The objectives of the study were to examine the relative prevalence of specific behavior problems, assess the impact of subject characteristics, and derive an empirical classification of behavioral and emotional problems for this population. Results indicated that the youngsters experienced high rates of behavior and emotional problems. Cluster analysis suggested that six- and eight-cluster solutions best fit the ratings provided by parents and teachers, respectively. Both parent and teacher cluster solutions contained groups of children characterized as problem free, well adapted, hyperactive, anxious, and with undifferentiated behavior disturbances. The empirically derived clusters were supported by data external to the analyses.

Keywords Autism · Pervasive developmental disorder · Behavior problems · Prevalence · Cluster analysis · Nisonger child behavior rating form

Introduction

Although mentioned as secondary features, behavior problems such as temper tantrums, overactivity, aggression, and self-injurious behaviors are part of the clinical

descriptions of autism in current psychiatric diagnostic systems (APA, 1994; WHO, 1992). There is ample evidence indicating that young people with pervasive developmental disorders (PDDs) present with a wide range of behavior and emotional problems, with symptoms of anxiety, depression, and attention deficit hyperactivity disorder (ADHD) being the most frequently reported (e.g., see Gillberg & Billstedt, 2000; Lainhart, 1999; Sverd, 2003).

An important barrier to the study of behavior and emotional problems is the heterogeneity of symptoms presented by individuals with PDDs. Such individuals often differ significantly in cognitive and adaptive functioning, and the nature and severity of autistic behaviors vary and change with development. It is not well understood how these individual differences impact the occurrence and presentation of behavior and emotional problems beyond the core symptoms that define the PDD population.

Comparisons across studies and generalization of findings have been hampered by the fact that many studies contained small or clinically referred samples or did not use standardized instruments. Noteworthy exceptions include the recent studies by Gadow and colleagues (Gadow, DeVincent, & Azizian, 2004; Gadow, DeVincent, Pomeroy, & Azizian, 2005) and Tonge and Einfeld (2003). Gadow et al. (2004) and Gadow et al. (2005) compared the severity and prevalence of DSM-IV symptoms in preschoolers and elementary school aged children to clinic controls and community-based samples. In both studies, children with PDDs were recruited from a developmental disabilities specialty clinic and were rated by parents and teachers on a rating scale containing all symptoms relevant to childhood disorders from the DSM-IV. Overall, preschoolers with PDDs ($n = 182$) presented with more severe DSM-IV psychiatric symptoms than their peers in regular and special education, and to some extent to non-PDD

L. Lecavalier (✉)
Nisonger Center, Ohio State University, 305 McCampbell Hall,
1581 Dodd Drive, Columbus, OH 43210-1257, USA
e-mail: lecavalier.1@osu.edu

psychiatric referrals. School-aged children with PDDs ($n = 301$) exhibited a pattern of psychiatric symptoms highly similar to non-PDD clinic referrals. The highest screening prevalence rates were for ADHD, oppositional defiant disorder, and generalized anxiety disorder. The results also suggested that children with the most severe PDD symptoms had fewer psychiatric symptoms.

In their epidemiological study of children with mental retardation (MR), Tonge and Einfeld (2003) had a subsample of 118 children with autism (mean age of 8.5 years). All participating children were rated on the *Developmental Behaviour Checklist* at three different intervals across an eight-year span. Results indicated that 73.5% of children with autism were above the cutoff considered clinically significant for caseness. Scores were fairly stable across time and the researchers commented that young people with autism are at high risk of suffering ongoing serious behavioral and emotional disturbance, over and beyond those defining the disorder.

In the current study, parents and teachers completed the *Nisonger Child Behavior Rating Form* (NCBRF; Aman, Tassé, Rojahn, & Hammer, 1996; Tassé, Aman, Hammer, & Rojahn, 1996) on a large non-clinically referred sample of children and adolescents with PDDs. The NCBRF is an empirically derived standardized instrument designed to measure behavior and emotional problems in young people with developmental disabilities. In addition to measuring behavior problems commonly seen in young people with PDDs, the NCBRF also contains two prosocial behavior subscales. Normative data were based on a clinically referred sample of children with developmental disabilities. Recently, the tool was shown to have good construct validity in a PDD population (Lecavalier, Aman, Hammer, Stoica, & Matthews, 2004).

The current study had three interrelated objectives. First, it was designed to examine the relative prevalence of specific behavior problems. Despite the accepted fact that people with autism experience high rates of behavior problems, few studies, if any, have reported such information. Categorical classifications and subscale scores make it difficult to know the proportion of individuals who present with a variety of behavior problems. The second objective was to examine the effects of subject characteristics for both parent and teacher versions of the instrument. Both English and French versions of the instrument have normative data for a MR population, but not for individuals with PDDs (Tassé et al., 1996; Tassé, Girouard, & Morin, 1999). Examining the effects of subject characteristics on problem behavior is important for several reasons. A number of disorders are more prevalent in one gender (e.g., conduct disorder or anxiety disorders), others change over the lifespan (e.g., ADHD), and level of functioning impacts the presentation and

occurrence of psychopathology (e.g., Brown, Aman, & Haverkamp, 2002; Rojahn, Matson, Lott, Esbensen, & Smalls, 2001; Tassé et al., 1996). The final goal was to derive an empirical classification of behavioral and emotional problems. Cluster analysis was used as a multivariate technique to isolate groups of subjects with similar behavioral profiles. It has been used successfully to examine the dimensions of psychopathology in young people with MR as well as their typically developing peers (e.g., Brown, Aman, & Lecavalier, 2004; Kamphaus, Huberty, DiStehano, & Petoskey, 1997; Kamphaus, Petoskey, Cody, Rowe, & Huberty, 1999). Most cluster analysis studies of psychopathology have not included prosocial behaviors, and none have been done assessing behavior problems in young people with PDDs. Prosocial behaviors are important to include in studies of psychopathology because they have been identified as significant protective factors against the development of childhood disorders (e.g., Kamphaus et al., 1999). This is the first study to address these objectives in a large non-clinically referred sample of young people with PDDs.

Method

Measures

Nisonger Child Behavior Rating Form

The content of the parent and teacher versions of the NCBRF is identical, but the subscale scoring methods differ slightly to reflect small variations that occurred in factor structures when the scales were developed (Aman et al., 1996). Ten social competence items are distributed on two subscales: Compliant/Calm and Adaptive/Social. Items are rated on a four-point Likert scale ranging from *not true* (0) to *completely or always true* (3). Sixty-six problem behavior items are also rated on a four-point Likert scale ranging from *did not occur or was not a problem* (0) to *occurred a lot or was a serious problem* (3). Raters are instructed to consider both the rate of occurrence and degree to which the behavior was a problem. Sixty items from the parent version and 62 items from the teacher version are distributed on six subscales: Conduct Problem, Insecure/Anxious, Hyperactive, Self-Injury/Stereotypic, Self-Isolated/Ritualistic, and Overly Sensitive (parent version) or Irritable (teacher version). With the exception of the Overly Sensitive/Irritable subscales, both versions share very similar subscale content. As assessed in several studies, the NCBRF has good psychometric properties (see Aman et al., 1996; Lecavalier et al., 2004; Tassé, Morin, & Girouard, 2000).

Scales of Independent Behavior-Revised

The SIB-R (Bruininks, Woodcock, Weatherman, & Hill, 1996) is a comprehensive and standardized measure of adaptive behavior. It contains 14 subscales distributed into four areas: (a) Motor Skills, (b) Social and Communication Skills, (c) Personal Living Skills, and (d) Community Living Skills. The Broad Independence score is a measure of overall adaptive behavior or functional independence and is based on the average of the four different areas of adaptive behavior. The test provides norms from early infancy to adulthood and has good psychometric properties (Bruininks et al., 1996).

Procedure

Data were collected in 37 school districts across Ohio over a two-year period as part of larger state evaluation project (Hammer & Lecavalier, 2003). Elsewhere, we assessed the psychometric properties of the NCBRF and *Gilliam Autism Rating Scale* on about half of the participants included in the current study (Lecavalier, 2005; Lecavalier et al., 2004). Inclusion criteria for the current study were that students be aged between 3 and 21 years and receiving educational services for PDDs. Students were not chosen on the basis of any demographic variables such as level of functioning, behavior, or academic functioning. Project coordinators from each district were asked to select between 5 and 20 students from their rosters, depending on the size of the school district. In all, 611 children were identified as potential participants.

Data were collected via questionnaires from parents and teachers on classroom environments, home and school resources, and on several areas of student functioning. Investigators held eight regional meetings over the two-year period with representatives of school districts, parents, and teachers to explain the purpose of the project and the completion of the different instruments. The SIB-R was completed as a checklist by parents only. It is designed to be administered in an interview format, but checklist administration is considered acceptable under certain circumstances (Bruininks et al., 1996). In addition to the regional meetings and technical support offered throughout, a summary sheet with instructions and examples was provided for accurate completion of the SIB-R. A doctoral level graduate student in psychology verified every completed SIB-R for anomalies in responding (e.g., not giving credit for skills that were *obviously* mastered such as crawling for an individual who walks, runs, and rides a bicycle). Parents also provided information on psychotropic medicine use, and teachers provided information on educational variables.

In all, 487 students were rated on the NCBRF; 353 were rated by parents and 437 were rated by teachers. Three hundred three participants were rated by both informants. Thus, 50 participants were rated only by parents and 184 were rated only by teachers. There were no differences in chronological age, gender, adaptive behavior, or on any of the eight subscale scores between the participants rated only by one informant and those rated by both.

Participants

Three hundred eighty-nine students were males (82.6%), 82 were females (17.4%), and 16 had missing data (3.3%). Most of the students (89.3%) were Caucasian and the average age was 9.6 years (SD = 3.8). The disability identified by the Individualized Education Plan (IEP) team and reported by teachers were as follows: 326 (67%) autism; 53 (11%) Preschooler with a Disability; 67 (14%) had a variety of other disability categories such as Speech and Language Impairment, Mental Retardation, or Other Health Impaired; and 40 (8%) had this information missing.

Parental reports indicated that 46.7% of the sample had taken at least one type of psychotropic medicine in the last 12 months. Three hundred forty-five parents (97.7%) confirmed that their child had been diagnosed by a physician or psychologist with a PDD. Students without a parental confirmation of a PDD diagnosis were classified as Autistic, Preschooler with a Disability, Speech and Language Impaired, or Other Health Impaired by their school districts.

Three hundred twenty parents provided complete adaptive behavior ratings. Average SIB-R age equivalents in months for the sample were as follows: 65 (SD = 43) for Broad Independence, 76 for Motor Skills (SD = 53), 58 for Social and Communication Skills (SD = 48), 66 for Personal Living Skills (SD = 50), and 66 for Community Living Skills (SD = 47). Average SIB-R standard scores for the sample were as follows: 52 (SD = 33) for Broad Independence, 69 for Motor Skills (SD = 30), 54 for Social and Communication Skills (SD = 35), 60 for Personal Living Skills (SD = 30), and 54 for Community Living Skills (SD = 33). Sixty-six percent of students obtained functional independence standard scores in the range of MR (i.e., 70 or lower).

Raters

Most ratings were provided by mothers ($n = 295$; 83.6%), but 23 were provided by fathers (6.5%), 24 by both parents (6.8%), and 11 by guardians or grandparents (3.1%). The average age of parents was 40.5 years (SD = 7.0; Range 21–59), and 52.6% graduated from college.

Two hundred sixty teachers rated the 437 students. Female respondents completed 396 of the 437 ratings (90.6%). Average teacher age was 37.9 years (SD = 10.2; Range 22–62) and average teaching experience was 11.2 years (SD = 9.1; Range 0–36). Level of teacher education was as follows: three had an Associate's degree (0.7%), 229 had Baccalaureates (52.4%), 200 had Master's degrees (45.8%), and five did not report this information (1.1%). Most of the teacher ratings were provided by primary instructors ($n = 311$; 71.2%) or instructors ($n = 88$; 20.1%), but 39 (8.9%) were provided by other types of raters such as classroom assistants or "shadows." Three hundred ninety-three of the 437 ratings (89.9%) were completed by raters who knew the students they were rating for more than 6 months.

Data Analysis

Prevalence and Effects of Subject Characteristics

The prevalence of problem behaviors was determined by simple frequency counts of items endorsed as occurring often/being a moderate problem or occurring a lot/being a severe problem (i.e., scores of 2 or 3). Subject variables were analyzed with a three-way analysis of variance in which each subscale was analyzed as a function of gender, age, and level of adaptive behavior. Ages were roughly divided into quartiles and adaptive behavior composite scores were roughly divided into three groups (severe/profound MR range; mild/moderate MR range; no MR range). The least squares means option was used (type III analysis) to correct for different sample sizes. To correct for multiple comparisons, a Bonferroni correction was used, setting the alpha level at .006. Tukey's HSD were used as post hoc tests.

Cluster Analysis

Cluster analysis refers to the process of applying multivariate heuristics in order to partition a set of entities into relatively homogeneous subgroups based on similarities between them (i.e., the goal is to derive clusters with minimal within-group variance and maximal between-group variance). Unlike factor analysis, the subject (rather than the variable) is the basic unit of analysis.

In the current study, a two-step procedure was followed: Ward's hierarchical method was followed by a K-Means analysis. Ward's analysis is a non-overlapping and agglomerative method. It begins with each case considered as a separate cluster. At each successive step in the clustering, the two most similar clusters are merged so that variance within clusters is minimized. This method has good cluster recovery ability and has performed well with

behavioral data (Milligan, 1996; Milligan & Cooper, 1987). The pseudo F statistic, pseudo t^2 statistic, and cubic clustering criterion were used as statistical indexes to determine the optimal number of clusters. In addition, clinical meaningfulness of solutions was taken into account. A drawback of Ward's method is that once a child is assigned to a cluster, its membership cannot change. The intent behind the K-means analysis is to make possible some shifts in cluster membership. It is an iterative cluster partitioning method that starts with a set number of clusters (in this case, determined by the results of Ward's analysis). At each step of the clustering process, the cases are moved to the nearest centroid, which are re-calculated based on the new assignment. The procedure is repeated until no reassignments are made.

The reliability of cluster solutions was examined by randomly dividing the sample into two three separate times. The six half samples were cluster analyzed based on the centroids of the total sample. A measure of agreement (coefficient Kappa) was then calculated between cluster assignments obtained for the total and half samples. The average Kappa for the six half samples were .84 and .91 for parent and teacher solutions, respectively.

Findings were also replicated in the subsample of student with an autism classification in their IEP for both parent ($n = 230$) and teacher ($n = 314$) ratings. Cluster solutions were replicated by using the centroids obtained from the entire sample and by conducting separate cluster analyses (Ward's followed by K-Means). The second procedure is a strict test of the robustness of the solutions as cluster analysis is heavily dependent on the sample used. Cluster centroids and percentages of individuals belonging to different clusters were *very* similar to those obtained for the entire PDD sample with both procedures. Therefore, only data on the broader PDD phenotype is presented.

The validity of the cluster solutions was appraised with data external to the cluster analysis. Groups were contrasted on demographic variables, psychotropic medication use, and the presence of a behavior plan in the student's IEP.

Results

Relative Prevalence

Table 1 shows the percentages of items rated as occurring often/being a moderate problem and occurring a lot/being a severe problem, separately for parents and teachers. Overall, the rates reported by parents and teachers were very similar, but varied significantly across items. For instance, parents and teachers reported moderate or severe problems with *Stealing* for 3.4% and 3.6% of the sample, respectively. On the other hand, the item *Easily frustrated*

Table 1 Percentage of behavior problems endorsed as moderate or severe in severity according to parents and teachers

	Parents (<i>n</i> = 353)			Teachers (<i>n</i> = 437)		
	Severity of problem			Severity of problem		
	Moderate	Severe	Total	Moderate	Severe	Total
NCBRF item ^a and corresponding parent subscale ^b						
<i>Conduct problem</i>						
Argues	19.8	5.1	24.9	17.8	6.6	24.4
Cruelty/meanness	4.8	4.8	9.6	10.8	4.1	14.9
Defiant	20.7	5.9	26.6	14.4	11.0	25.4
Destroys property	8.2	3.1	11.3	6.9	4.6	11.5
Disobedient	13.6	3.4	17.0	16.0	5.9	21.9
No guilt	15.0	8.2	23.2	13.0	9.8	22.8
Explosive (I)	15.0	7.9	22.9	14.0	10.5	24.5
Physical fights	4.2	1.1	5.3	3.7	1.6	5.3
Attacks people	6.8	3.1	9.9	8.9	5.9	14.8
Runs away	9.1	4.8	13.9	9.4	6.4	15.8
Stubborn (–)	29.7	21.0	50.7	24.7	19.7	44.4
Talks back	12.5	4.8	17.3	10.8	6.6	17.4
Temper tantrums (I)	21.0	7.6	28.6	18.5	11.7	30.2
Threatens people	2.8	1.7	4.5	5.3	2.3	7.6
Violates Rules	9.6	3.7	13.3	10.5	4.6	15.1
Argues (I/A)	9.3	2.8	12.1	8.0	4.1	12.1
<i>Insecure/Anxious</i>						
Exaggerates	5.4	3.1	8.5	4.1	3.0	7.1
Others against him	6.5	4.8	11.3	5.5	3.4	8.9
Feels worthless	5.4	2.3	7.7	3.2	0.9	4.1
Lying/cheating (–)	4.5	1.7	6.2	2.5	2.1	4.6
Nervous/tense	16.7	4.5	21.2	12.8	5.3	18.1
Anxious to please	8.8	1.4	10.2	8.2	1.6	9.8
Says no one likes him	4.0	4.0	8.0	3.7	1.8	5.5
Secretive (S-I/R)	5.1	2.2	7.3	3.7	1.1	4.8
Self-conscious	5.9	1.7	7.6	3.9	1.6	5.5
Shifts topics	9.9	3.4	13.3	8.9	3.2	12.1
Steals (CP)	1.7	1.7	3.4	2.5	1.1	3.6
Sulks/silent/moody	4.8	1.7	6.5	8.9	2.3	11.2
Talks too much/loud	15.6	8.8	24.4	13.3	9.2	22.5
Too fearful/anxious	12.2	4.5	16.7	8.2	2.7	10.9
Worrying	10.2	4.2	14.4	9.8	3.7	13.5
<i>Hyperactive</i>						
Difficulty concentrating	36.3	13.0	49.3	31.6	18.5	50.1
Easily distracted (–)	38.5	21.2	59.7	30.4	29.1	59.5
Fails to finish	21.0	10.2	31.2	14.9	12.4	27.3
Fidgets/wiggles/squirms	25.2	16.7	41.9	24.5	19.0	43.5
Overactive	22.4	18.7	41.1	17.4	11.9	29.3
Excited/exuberant	19.8	7.9	27.7	16.5	3.9	20.4
Repeats over and over	22.9	18.1	41.1	13.5	11.4	24.9
High energy level	25.2	18.7	43.9	18.3	12.1	30.4
Short attention span	33.1	20.4	53.5	26.3	21.1	47.4
<i>Self-Isolated/Ritualistic</i>						
Apathetic/unmotivated	14.7	4.0	18.7	19.2	9.4	28.6
Rituals	15.9	11.9	27.8	15.6	10.5	26.1
Shy/bashful	14.7	3.7	18.4	12.6	5.7	18.3
Isolates self (–)	23.5	12.2	35.7	22.0	13.5	35.5
Refuses to talk	11.3	8.5	19.8	10.1	5.7	15.8
Shy/timid	8.8	3.7	12.5	8.2	4.6	12.8
Odd repetitive beh.	17.8	10.5	28.3	19.2	13.3	32.5
Withdrawn/uninvolved	14.2	11.6	25.8	19.5	9.4	28.9
<i>Self-Injury/Stereotypic</i>						
Hits or slaps self	10.2	5.7	15.9	9.6	6.2	15.8
Rocks repetitively	9.6	2.5	12.1	11.0	5.0	16.0
Scratches skin/pulls hair	5.1	3.4	8.5	6.4	3.9	10.3

Table 1 continued

	Parents (<i>n</i> = 353)			Teachers (<i>n</i> = 437)		
	Severity of problem			Severity of problem		
	Moderate	Severe	Total	Moderate	Severe	Total
Gouges/eats inedible	8.2	4.0	12.2	6.2	5.3	11.5
Bites self	2.8	3.1	5.9	1.8	3.0	4.8
Physically harms self	6.8	4.2	11.0	4.6	5.7	10.3
Threatens to harm self	2.0	1.4	3.4	0.9	0.4	1.3
<i>Overly Sensitive</i>						
Clings to adults (–)	13.3	3.4	16.7	13.3	3.9	17.2
Crying (I)	19.3	4.0	23.4	17.6	5.5	23.1
Easily frustrated (I)	43.6	18.4	62.0	33.2	20.4	53.6
Overly sensitive (I/A)	24.4	7.9	32.3	16.9	4.8	21.7
Feelings easily hurt (I/A)	18.1	6.5	24.6	11.9	3.4	15.3
<i>No Subscale</i>						
Irritable (I)	14.2	5.1	19.3	15.6	7.6	23.2
Repeatedly flaps (S-I/S)	15.9	15.3	31.2	12.1	13.5	25.6
Changes in mood (I)	19.5	8.5	28.0	21.5	11.0	32.5
Repetitive mvt (S-I/S)	17.6	10.5	28.1	13.7	10.1	23.8
Under-active/slow (S-I/R)	7.4	3.7	11.1	8.7	4.8	13.5
Unhappy or sad (S-I/R)	4.8	0.8	5.6	6.9	2.5	9.4

^a Summary of items only; ^b Letters between parentheses represent the teacher subscale on which the item belongs. CP = Conduct Problem; I/A = Insecure/Anxious; S-I/R = Self-Isolated/Ritualistic; S-I/S = Self-Injury/Stereotypic; I = Irritable; – = no subscale

was endorsed by parents and teachers as a moderate or severe problem for 62.0 and 53.6% of the sample, respectively.

The most frequently endorsed items were those relating to symptoms of ADHD. For instance, in percentages, the following items were reported to be moderate or severe problems (parents/teachers): *Difficulty concentrating* (49/50), *Easily distracted* (60/60), and *Fidgeting, wiggling, squirming* (42/44).

The item *Physically harms self* was endorsed as being a moderate or severe problem for 11% and 10% of parents and teachers, respectively. Rates for specific self-injurious behaviors rated as moderate or severe problems ranged from 5% (teachers ratings on *Bites self*) to 16% (parent ratings on *Hits/slaps self*). Along the same lines, *Destruction of property* was reported to be a moderate or severe problem for 11% (parents) and 12% (teachers) of the sample. Both types of informants reported moderate or severe problems with *Physical fights* for 5% of the sample.

In percentages, the following symptoms of anxiety were reported to be moderate or severe problems (parents/teachers): *Nervous/tense* (21/18), *Too fearful/anxious* (17/11), and *Worrying* (14/14). In percentages, the following symptoms of mood problems were reported to be moderate or severe problems (parents/teachers): *Crying* (23/23), *Irritable* (19/23), *Temper tantrums* (29/30), and *Unhappy/sad* (6/9).

Subject Characteristics

No gender or interaction effects were found. Therefore both sexes were combined for subsequent analyses.

Table 2 shows average subscale scores and standard deviations based on the four age groups and by type of informant. Parent ratings yielded a significant main effect for the Insecure/Anxious subscale [$F(3, 340) = 11.76$], with the youngest participants obtaining the lowest scores. Teacher ratings revealed significant main effects for age for the Insecure/Anxious [$F(3, 422) = 12.07$], Self-Injury/Stereotypic [$F(3, 422) = 4.44$], and Self-Isolated/Ritualistic subscales [$F(3, 422) = 4.99$]. As with parent ratings, younger children had significantly lower scores on the Insecure/Anxious subscale. The 10- to 12-year-old age group had the highest scores on the Self-Injury/Stereotypic and Self-Isolated/Ritualistic subscales.

Table 3 shows average subscale scores and standard deviations based on the three levels of adaptive behavior. Parent ratings indicated significant differences across groups for five of the eight subscales. Participants with lower adaptive skills obtained lower scores (worse) on the Adaptive/Social [$F(2, 317) = 30.09$] and Insecure/Anxious subscales [$F(2, 317) = 16.42$]. They also had higher scores (worse) on the Hyperactive [$F(2, 317) = 7.03$], Self-Injury/Stereotypic [$F(2, 317) = 21.63$], and Self-Isolated/Ritualistic [$F(2, 317) = 10.92$] subscales. Teacher ratings were quite similar to parent ratings, with seven of the eight subscales reaching statistical significance. Participants with lower adaptive skills had lower scores on the Compliant/Calm [$F(2, 281) = 19.71$], Adaptive/Social [$F(2, 281) = 26.48$], Insecure/Anxious [$F(2, 281) = 8.45$], and Irritable [$F(2, 281) = 13.51$] subscales. Conversely, they obtained higher scores on the Conduct Problem [$F(2, 281) = 8.93$], Hyperactive

Table 2 Summary of the main effect of age for parent and teacher versions of the NCBRF

	Age in years				<i>F</i> (3,336)	Source of significance
	3–6 (<i>n</i> = 102) Mean (SD)	7–9 (<i>n</i> = 71) Mean (SD)	10–12 (<i>n</i> = 85) Mean (SD)	13–21 (<i>n</i> = 86) Mean (SD)		
<i>Parent subscales</i>						
Compliant/Calm (Range 0–18)	8.7 (2.4)	8.8 (3.0)	9.6 (2.9)	8.8 (3.0)	1.82	
Adaptive Social (Range 0–12)	4.7 (2.1)	5.2 (2.1)	5.3 (2.3)	5.2 (2.4)	1.41	
Conduct Problem (Range 0–48)	10.9 (7.4)	12.4 (8.1)	11.4 (8.4)	11.1 (9.0)	.53	
Insecure/Anxious (Range 0–45)	3.4 (3.8)	7.0 (7.4)	7.2 (7.6)	8.7 (6.7)	11.76*	1:2; 1:3; 1:4
Hyperactive (Range 0–27)	12.4 (6.1)	13.4 (6.0)	13.1 (6.3)	10.8 (6.3)	2.76	
Self-Injury/Stereotypic (Range 0–21)	2.0 (2.5)	2.5 (3.1)	2.9 (3.8)	2.7 (3.3)	1.18	
Self-Isolated/Ritualistic (Range 0–24)	6.9 (4.6)	5.9 (4.5)	6.6 (4.0)	7.2 (4.7)	1.16	
Overly Sensitive (Range 0–15)	5.3 (2.5)	5.6 (3.0)	5.5 (3.3)	5.2 (2.8)	.47	
	3–6 (<i>n</i> = 119) Mean (SD)	7–9 (<i>n</i> = 94) Mean (SD)	10–12 (<i>n</i> = 108) Mean (SD)	13–21 (<i>n</i> = 105) Mean (SD)	<i>F</i> (3, 419)	Source of significance
<i>Teacher subscales</i>						
Compliant/Calm (Range 0–15)	7.7 (3.0)	8.0 (3.0)	7.8 (2.9)	7.9 (3.3)	.18	
Adaptive Social (Range 0–15)	5.9 (3.0)	6.8 (3.0)	6.6 (2.8)	6.6 (3.0)	1.74	
Conduct Problem (Range 0–39)	6.3 (6.8)	8.0 (6.7)	8.7 (9.0)	7.2 (7.4)	2.17	
Insecure/Anxious (Range 0–45)	4.1 (4.7)	7.0 (7.4)	8.5 (7.1)	9.0 (7.8)	12.07*	1:2; 1:3; 1:4
Hyperactive (Range 0–24)	8.9 (5.2)	9.2 (5.7)	9.7 (4.9)	8.7 (5.5)	.69	
Self-Injury/Stereotypic (Range 0–27)	3.1 (3.6)	3.6 (4.0)	5.1 (4.8)	4.7 (5.5)	4.44*	1:3
Self-Isolated/Ritualistic (Range 0–33)	7.1 (4.8)	7.6 (5.8)	9.6 (5.3)	8.8 (5.0)	4.99*	1:3; 2:3
Irritable (Range 0–18)	5.8 (4.4)	6.4 (4.1)	7.1 (4.7)	6.4 (4.9)	1.37	

**p* < .006

[(*F* (2, 281) = 22.97)], and Self-Injury/Stereotypic [(*F* (2, 281) = 10.84)] subscales.

Average subscale scores and standard deviations presented by age and level of adaptive functioning combined for both parent and teacher ratings can be obtained by contacting the author.

Cluster Solutions

Parent Ratings

Parent ratings suggested that a six-cluster solution fit the data best. Table 4 shows mean subscale scores and standard deviations for this cluster solution. Clusters were characterized as follows: (a) *Cluster 1, Problem Free:*

Members of this cluster (31% of the sample) had scores within ± SD from the sample average on all eight subscales. The average age was 9.2 years (SD = 3.9) and average SIB-R composite score was 55.7 (SD = 30.9). (b) *Cluster 2, Well Adapted:* Compared to other clusters, members of this cluster (21% of the sample) had the highest scores on both prosocial subscales and the lowest scores on all six problem behavior subscales. The average age was 9.5 years (SD = 4.3) and the average SIB-R composite score was 57.2 (SD = 37.5). (c) *Cluster 3, Ritualistic and Hyperactive:* Members of this cluster comprised 13% of the sample. Compared to other clusters, they had the lowest scores on the Adaptive/Social subscale. They also obtained scores roughly one SD above the sample average on both the Self-Isolated/Ritualistic and

Table 3 Summary of the main effect of adaptive behavior level for parent and teacher versions of the NCBRF

	Adaptive Behavior Composite Score			<i>F</i> (2, 312)	Source of significance
	Below 35 (<i>n</i> = 102) Mean (SD)	35–70 (<i>n</i> = 109) Mean (SD)	Above 70 (<i>n</i> = 109) Mean (SD)		
<i>Parent subscales</i>					
Compliant/Calm (Range 0–18)	8.4 (2.8)	8.7 (2.9)	9.5 (2.7)	4.74	
Adaptive Social (Range 0–12)	4.0 (2.1)	4.8 (2.2)	6.2 (1.9)	30.09*	1:2; 1:3; 2:3
Conduct Problem (Range 0–48)	11.5 (8.7)	11.4 (7.8)	11.4 (8.4)	.01	
Insecure/Anxious (Range 0–45)	3.5 (4.2)	7.0 (7.0)	8.5 (7.5)	16.42*	1:2; 1:3
Hyperactive (Range 0–27)	13.8 (7.2)	12.9 (5.9)	10.7 (5.3)	7.03*	1:3; 2:3
Self-Injury/Stereotypic (Range 0–21)	4.1 (3.9)	2.1 (2.6)	1.4 (2.4)	21.63*	1:2; 1:3
Self-Isolated/Ritualistic (Range 0–24)	7.5 (4.7)	7.7 (4.5)	5.3 (3.6)	10.92*	1:3; 2:3
Overly Sensitive (Range 0–15)	4.8 (2.6)	5.3 (2.8)	6.0 (3.1)	5.16	
	Below 35 (<i>n</i> = 95) Mean (SD)	35–70 (<i>n</i> = 86) Mean (SD)	Above 70 (<i>n</i> = 103) Mean (SD)	<i>F</i> (2, 276)	Source of significance
<i>Teacher subscales</i>					
Compliant/Calm (Range 0–15)	6.6 (2.6)	7.9 (2.8)	9.2 (3.1)	19.71*	1:2; 1:3; 2:3
Adaptive Social (Range 0–15)	5.0 (2.5)	6.4 (2.8)	7.9 (3.2)	26.48*	1:2; 1:3; 2:3
Conduct Problem (Range 0–49)	9.2 (7.7)	6.2 (7.0)	5.2 (5.9)	8.45*	1:2; 1:3
Insecure/Anxious (Range 0–45)	3.9 (5.0)	7.0 (6.6)	8.7 (7.7)	13.51*	1:2; 1:3; 2:3
Hyperactive (Range 0–24)	10.6 (5.5)	9.1 (5.2)	7.5 (5.1)	8.93*	1:3
Self-Injury/Stereotypic (Range 0–27)	6.4 (5.4)	3.7 (3.5)	2.4 (3.5)	22.97*	1:2; 1:3
Self-Isolated/Ritualistic (Range 0–33)	8.5 (5.6)	8.4 (5.3)	6.4 (4.9)	5.08	
Irritable (Range 0–18)	7.6 (4.8)	5.7 (3.9)	4.8 (4.4)	10.84*	1:2; 1:3

**p* < .006

Hyperactive subscales. The average age was 8.8 years (SD = 4.2) and the average SIB-R composite score was 27.4 (SD = 26.8). (d) *Cluster 4, Hyperactive with Conduct Problems*: Members of this cluster (14% of the sample) had scores one SD above the sample average on the Hyperactive subscale. They also had elevated scores on the Conduct Problem subscale. Their average age was 9.1 years (SD = 3.5) and average SIB-R composite score was 41.3 (SD = 27.8). (e) *Cluster 5, Anxious*: Members of this cluster comprised 13% of the sample. They were characterized by scores one SD above sample average on the Insecure/Anxious subscale. With the exception of scores on the Overly Sensitive subscale, all other subscale scores were within \pm SD of the sample average. Average age was 11.0 years (SD = 3.3) and average SIB-R composite score was 71.6 (SD = 23.7). (f) *Cluster 6, Undifferentiated*

Behavior Disturbance: Compared to other clusters, members of this cluster (9% of the sample) had the lowest scores on the Compliant/Calm subscale. With the exception of the Self-Isolated/Ritualistic subscale score, all other problem behavior subscale scores were above one SD from the sample average. Average age was 10.2 years (SD = 3.3) and average SIB-R composite score was 52.5 (SD = 33.2).

In sum, parent ratings suggested that (a) 52% of the sample was free of significant behavior and emotional problems (clusters 1 and 2), (b) 35% of the sample were in a cluster having an average score above one SD of the sample average on the Hyperactive subscale (clusters 3, 4, and 6), (c) 22% were in a cluster having an average score above one SD of the sample average on the Insecure/Anxious subscale (clusters 5 and 6), and (d) 9% of the

Table 4 Means and standard deviations by NCBRF subscale for the six-cluster solution based on parent ratings (*N* = 353)

	Clusters						Sample average
	1 <i>n</i> = 107 30.6%	2 <i>n</i> = 74 21.0%	3 <i>n</i> = 45 12.7%	4 <i>n</i> = 49 13.9%	5 <i>n</i> = 47 13.3%	6 <i>n</i> = 31 8.8%	
<i>Subscales</i>							
Compliant/Calm (Range 0–18)	9.5 (2.3)	11.3 (2.7)	8.2 (2.6)	7.2 (1.9)	8.3 (2.3)	6.3 (2.5)	9.0 (2.8)
Adaptive/Social (Range 0–12)	5.4 (2.3)	6.3 (2.2)	3.4 (1.7)	4.1 (1.8)	5.6 (1.8)	4.5 (1.8)	5.1 (2.2)
Conduct Problem (Range 0–48)	8.2 (4.4)	3.7 (3.0)	9.7 (4.1)	18.3 (4.5)	14.8 (4.7)	28.8 (7.8)	11.5 (8.4)
Insecure/Anxious (Range 0–45)	3.8 (3.5)	1.9 (2.3)	3.6 (3.6)	7.9 (4.6)	13.8 (5.3)	18.2 (9.2)	6.5 (6.9)
Hyperactive (Range 0–27)	11.0 (3.0)	5.6 (2.9)	18.0 (4.0)	18.9 (4.1)	9.3 (3.8)	19.9 (4.6)	12.4 (6.3)
Self-Injury/Stereo. (Range 0–21)	1.8 (2.2)	0.7 (1.3)	3.4 (3.5)	3.7 (3.2)	1.9 (2.3)	7.3 (4.5)	2.5 (3.2)
Self-Isolated/Ritua. (Range 0–24)	5.4 (3.7)	3.2 (2.3)	11.9 (3.2)	7.3 (3.3)	7.9 (3.7)	10.5 (5.3)	6.8 (4.5)
Overly Sensitive (Range 0–15)	4.7 (1.9)	3.0 (2.0)	5.2 (2.4)	6.8 (2.6)	7.3 (2.4)	9.2 (3.0)	5.5 (2.9)

Cluster 1 = Problem Free; Cluster 2 = Well Adapted; Cluster 3 = Ritualistic and Hyperactive; Cluster 4 = Hyperactive with Conduct Problems; Cluster 5 = Anxious; Cluster 6 = Undifferentiated Behavior Disturbance

sample had several subscale scores above one SD of the average (cluster 6).

The following comparisons involved an attempt to examine the validity of the clusters. First, cluster membership was examined as a function of age, gender, and level of functioning. No significant differences were found across clusters for age or gender. As in previous analyses, adaptive behavior was treated as a categorical variable. There were significant differences at the *p* < .001 level between clusters as a function of adaptive behavior [(*F* (5, 314) = 9.96)]. Members of cluster 3 (Ritualistic and Hyperactive) had significantly lower adaptive behavior scores than members from all other clusters, except those of cluster 4 (Hyperactive with Conduct Problems). Members of cluster 5 (Anxious) had significantly higher adaptive behavior scores than members from all other five clusters. Psychotropic medicine use was compared across clusters. This variable was treated dichotomously and cross-tabulated with the six clusters. The resulting Chi Square was significant [χ^2 (5, *n* = 353) = 34.6,

p < .0001]. Follow-up analyses indicated that members from clusters 1 and 2 (Problem Free and Well Adapted) were taking significantly less psychotropic medicine than members from clusters 4 and 6 (Hyperactive with Conduct Problems, and Undifferentiated Behavioral Disturbance). Finally, the presence of a behavior plan in the students' IEP was compared across clusters. This variable was also treated dichotomously and cross-tabulated with the six empirical clusters. The resulting Chi Square was significant [χ^2 (5, *N* = 299) = 11.6, *p* < .05] and follow-up analyses indicated that members from cluster 6 (Undifferentiated Behavioral Disturbance) had significantly more behavior plans than members from cluster 1 (Problem Free).

Teacher Ratings

Teacher ratings suggested that an eight-cluster solution fit the data best. Table 5 shows mean subscale scores and standard deviations for this cluster solution. Clusters were

Table 5 Means and standard deviations by NCBRF subscale for the eight-cluster solution based on teacher ratings (*N* = 437)

	Clusters								Sample average
	1 <i>n</i> = 86 19.7%	2 <i>n</i> = 100 22.9%	3 <i>n</i> = 56 12.8%	4 <i>n</i> = 37 8.4%	5 <i>n</i> = 51 11.7%	6 <i>n</i> = 48 11.0%	7 <i>n</i> = 36 8.2%	8 <i>n</i> = 23 5.2%	
<i>Subscales^a</i>									
Compliant/Calm	7.3 (1.9)	11.4 (2.3)	7.4 (2.5)	5.9 (2.0)	7.0 (2.2)	8.0 (2.0)	4.8 (2.1)	4.6 (1.8)	7.8 (3.1)
Adaptive/Social	5.2 (1.9)	9.5 (2.5)	4.4 (2.2)	6.5 (2.1)	5.8 (2.3)	7.5 (2.5)	3.9 (2.1)	6.0 (2.1)	6.5 (3.0)
Conduct Problem	3.9 (3.1)	1.7 (2.2)	5.7 (3.8)	16.7 (5.1)	7.3 (3.7)	5.5 (4.2)	18.3 (4.7)	23.7 (7.5)	7.5 (7.5)
Insecure/Anxious	1.5 (1.9)	5.3 (3.4)	4.4 (3.7)	7.1 (4.5)	5.6 (4.0)	18.0 (5.2)	6.2 (4.7)	22.6 (5.5)	7.0 (7.0)
Hyperactive	6.3 (3.0)	4.8 (3.0)	10.7 (4.2)	7.6 (4.2)	15.1 (3.5)	10.0 (4.5)	15.2 (3.7)	13.7 (5.6)	9.2 (5.3)
Self-Injury/Stereo.	3.4 (3.3)	1.2 (1.7)	8.3 (4.4)	1.8 (1.8)	3.0 (2.8)	2.6 (2.4)	12.8 (4.7)	5.6 (4.0)	4.1 (4.6)
Self-Isolated/Ritua.	5.5 (3.4)	4.8 (3.1)	14.6 (4.6)	6.7 (3.4)	6.1 (3.2)	11.0 (5.2)	12.0 (4.3)	12.1 (5.2)	8.2 (5.3)
Irritable	4.6 (2.9)	2.0 (1.9)	6.4 (3.6)	9.0 (3.8)	7.1 (3.3)	7.1 (3.0)	13.5 (3.2)	13.1 (2.2)	6.4 (4.6)

Cluster 1 = Problem Free; Cluster 2 = Well Adapted; Cluster 3 = Ritualistic; Cluster 4 = Conduct Problem; Cluster 5 = ADHD; Cluster 6 = Anxious; Cluster 7 = Undifferentiated Behavior Disturbance with Stereotypy; Cluster 8 = Undifferentiated Behavior Disturbance with Anxiety. ^a Ranges of subscale scores are the same as in Tables 2 and 3

characterized as follows: (a) *Cluster 1, Problem Free*: Members of this cluster (20% of the sample) had scores slightly below the sample average on both prosocial behavior subscales. They also had scores below the sample average on all six problem behavior subscales. The average age was 8.1 years (SD = 3.7) and average SIB-R composite score was 37.0 (SD = 29.6; $n = 31$). (b) *Cluster 2, Well Adapted*: Compared to other clusters, members of this cluster (23% of the sample) were characterized by the highest prosocial subscale scores. They also had the lowest scores for five of the six problem behavior subscales. Average age was 9.8 years (SD = 3.8) and average SIB-R composite score was 74.2 (SD = 27.9; $n = 36$). (c) *Cluster 3, Ritualistic*: Members of this cluster comprised 13% of the sample. They were characterized by scores that fell more than one SD above the sample average on the Self-Isolated/Ritualistic subscale. With the exception of scores on the Self-Injury/Stereotypic subscale, all other scores fell at \pm SD from the sample average. Average age was 9.6 years (SD = 4.1) and average SIB-R composite score was 39.8 (SD = 25.5; $n = 71$). (d) *Cluster 4, Conduct Problem*: Members of this cluster comprised 8% of the sample. They were characterized by scores more than one SD above the sample average on the Conduct Problem subscale. With the exception of the scores on the Self-Injury/Stereotypic subscale, all other subscales fell within \pm SD of the sample average. Average age was 9.1 years (SD = 2.8) and average SIB-R composite score was 50.2 (SD = 28.4; $n = 63$). (e) *Cluster 5, Hyperactive*: Members of this cluster comprised 12% of the sample. They obtained scores more than one SD above sample average on the Hyperactive subscale and scores within \pm SD of the sample average on all other subscales. The average age was 8.3 years (SD = 3.0) and the average SIB-R composite score was 44.6 (SD = 34.6; $n = 31$). (f) *Cluster 6, Anxious*: Members of this cluster (11% of the sample) were characterized by scores that were more than one SD above sample average on the Insecure/Anxious subscale. With the exception of the Self-Isolated/Ritualistic subscale, all other subscales fell within \pm SD from the sample average. The average age was 11.7 years (SD = 3.5) and the average SIB-R composite score was 72.0 (SD = 28.8; $n = 16$). (g) *Cluster 7, Undifferentiated Behavior Disturbance with Stereotypy*: Members of this cluster comprised 8% of the sample. They obtained the lowest scores on the Adaptive/Social subscale and scores roughly one SD below sample average on the Compliant/Calm subscale. They also obtained scores significantly above one SD from the sample average on the Conduct Problem, Hyperactive, Self-Injury/Stereotypic, and Irritable subscales. Average age was 10.6 years (SD = 3.3) and average SIB-R composite score was 27.0 (SD = 29.2; $n = 23$). *Cluster 8, Undifferentiated Behavior Disturbance with Anxiety*: Compared to

other clusters, members of this cluster (5% of the sample) had the lowest scores on the Compliant/Calm subscale. They were also characterized by scores that fell above one SD of the sample average on the Conduct Problem, Insecure/Anxious, Hyperactive, and Irritable subscales. Average age was 10.1 years (SD = 3.1) and average SIB-R composite score was 74.1 (SD = 34.1; $n = 13$).

In sum, when including members of the Ritualistic cluster, teacher ratings suggested that 55% of the sample was relatively free of behavior/emotional problems (clusters 1, 2, and 3). They also indicated that (a) 25% of the sample were in clusters having an average score above one SD from the sample average on the Hyperactive scale (clusters 5, 7, 8), (b) 16% of the sample were in clusters having an average score above one SD from the sample average on the Insecure/Anxious subscale (clusters 6 and 8), and (c) 13% of the sample had several subscale scores above one SD of the average (clusters 7 and 8).

The external validity of the clusters derived from teacher ratings was examined in the same fashion as with the parent ratings. There were no significant differences across clusters as a function of gender. There were significant differences at the $p < .001$ level across clusters as a function of age [($F(7, 418) = 5.92$)]. Members of cluster 6 (Anxious) were significantly older than members of clusters 1 (Problem Free), 5 (Hyperactive), and 6 (Conduct Problem). Members of cluster 1 (Problem Free) were younger than members from clusters 2 (Well Adapted) and 7 (Undifferentiated Behavior Disturbance with Stereotypy). There were also significant differences at the $p < .001$ level across clusters as a function of adaptive behavior level [($F(7, 276) = 13.79$)]. Members of cluster 6 (Anxious) had higher adaptive behavior scores than members of cluster 1 (Problem Free), 3 (Ritualistic), 5 (Hyperactive), and 7 (Undifferentiated Behavioral Disturbance with Stereotypy). Members from cluster 2 (Well Adapted) had more adaptive skills than members from clusters 1 (Problem Free), 3 (Ritualistic), 5 (Hyperactive), and 7 (Undifferentiated Behavioral Disturbance with Stereotypic Behavior). Finally, members from cluster 8 (Undifferentiated Behavioral Disturbance with Anxiety) had more adaptive skills than members from clusters 1 (Problem Free) and 7 (Undifferentiated Behavioral Disturbance with Stereotypy). Psychotropic medicine use was compared across the eight clusters. The resulting Chi Square was significant [$\chi^2(7, N = 303) = 18.1, p < .05$]. Follow-up analyses indicated that members from clusters 1 (Problem Free) and 2 (Well Adapted) were taking less psychotropic medicine than members from cluster 8 (Undifferentiated Behavioral Disturbance with Anxiety). Finally, the presence of a behavior plan in the students' IEP was compared across clusters. The resulting Chi Square was significant [$\chi^2(7, n = 419) = 22.8, p < .001$] and

follow-up analyses indicated that members from Clusters 2 (Well Adapted) and 5 (Hyperactive) had significantly fewer behavior plans than members from clusters 7 and 8 (both Undifferentiated Behavioral Disturbance groups).

Discussion

Prevalence

This was the first study to report on prevalence rates of specific behavior problems in a large non-clinically referred sample of young people with PDDs and to use a quantitative method for typing psychopathology. The sample was probably representative of the population of young people with PDDs attending public schools in the American Midwest. Participants were not selected on the basis of any behavioral or academic criteria. They were recruited from a good range of urban and rural communities and had an ethnic composition similar to the state's. The project was done in close collaboration with the Ohio Department of Education and response rates were very high, with 79.7% of the students initially targeted being rated by at least one informant (i.e., 487 out of 611). Supportive of this is the fact that the distributions of gender and levels of adaptive functioning were consistent with previous epidemiological studies of PDDs (e.g., see Bryson & Smith, 1998) and patterns of psychotropic medicine use were very similar to other reports for this population (e.g., Aman, Lam, & Collier-Crespin, 2003).

We recognize that some behaviors are an inherent part of PDDs (e.g., social withdrawal or repetitive behaviors), while others are mediated by cognitive and language abilities (e.g., threatening or arguing). The purpose of this study was not to determine if behavior and emotional problems are manifestations of the PDD diathesis, separate clinical entities, or non-specific symptoms secondary to the core feature of PDDs. Regardless of one's position in this debate, results indicated that young people with PDDs experience high rates of behavior and emotional problems. The *nature* of reported problems was consistent with other reports in the literature (e.g., see Gadow et al., 2004; Gadow et al., 2005; Lainhart, 1999). However, the NCBRF does not provide categorical classifications, which renders direct comparisons with many studies difficult.

Subject Characteristics

As far as subject characteristics go, the only age effect observed for both informants was on the Insecure/Anxious subscale. This effect was also found in a sample of clinically referred children with MR on both parent and teacher

versions of the NCBRF (Tassé et al., 1996). The absence of age effects on the Hyperactive subscale was somewhat surprising, as it has been observed in children with MR and in their typically developing peers (e.g., Arnold, 2000; Brown et al., 2002; Rojahn & Helsel, 1991; Tassé et al., 1996). The absence of gender effects has been reported in several studies with developmentally disabled populations, including children with PDDs (Brown et al., 2002; Gadow et al., 2004; Gadow et al., 2005; Tassé et al., 1996). This seems to be a consistent difference between populations with and without developmental disabilities. Level of adaptive functioning was the subject characteristic that impacted subscale scores the most. Overall, lower adaptive skills were associated with less prosocial behaviors and symptoms of anxiety, and higher scores on other subscales, indicating more problems. These findings are consistent with the MR literature (e.g., Brown et al., 2002; Einfeld & Tonge, 2002; Marshburn & Aman, 1992; Rojahn, Borthwick-Duffy, & Jacobson, 1993). Taken as a whole, these data underscore the importance of having appropriate comparison groups in clinical and research contexts. Subject characteristics impacted subscale scores differently than in MR and typically developing populations.

Empirical Classification

Cluster analyses complemented other descriptive analyses in at least two ways. First, they allow one to examine the severity and overlapping nature of behavioral dimensions simultaneously. Second, by using the person as the basic unit of analysis, they enabled us to study *all* cases on behavioral dimensions (including prosocial behaviors). Overall, parent and teacher ratings yielded similar solutions that were clinically meaningful and were supported with data external to the analyses. They indicated that slightly more than half of the students were relatively behavior- and emotional-problem free (52% for parent ratings and 55% for teacher ratings). For comparison purposes, Brown et al. (2004) reported that 64% of their sample of non-clinically referred children with MR were free of problem behavior. Kamphaus and colleagues reported rates of 53% and 63% in typically developing children based on teacher and parent ratings, respectively. Results of the cluster analysis also indicated that a substantial minority of young people presented with severe and undifferentiated behavior disturbances (8% and 13% for parent and teacher ratings, respectively). These rates were much higher than those reported in Brown et al. (4%) as well as in the Kamphaus et al. studies (3% and 4% for parents and teachers, respectively). Interestingly, participants with low and high adaptive skills were in relatively problem-free and severe psychopathology clusters. The cluster analyses also showed that different patterns of prosocial behaviors were

associated with different cluster memberships. Finally, the cluster solutions obtained here were similar to that obtained by Brown et al. (2004) in a non-clinically referred sample of children with MR. Brown and colleagues reported an eight-cluster solution with two problem-free clusters, as well as Hyperactive, Conduct Problem, Shy/Inactive, and Undifferentiated Behavior Disturbance clusters (the other two clusters were Social Withdrawal with Agitation and Autistic-like).

The differences between the solutions obtained for parent and teacher ratings were threefold. First, teachers reported two types of undifferentiated behavior disturbances (as compared with one for parents), based on level of functioning. Second, teacher ratings produced two separate clusters for hyperactive and conduct problems, instead of just one found with the parent data. Finally, teacher ratings produced two separate clusters for Ritualistic and Hyperactive behaviors, whereas parents only had one. These differences could be an artifact of the different samples, factor structures, raters, and patterns of medication use. In the only other studies with overlapping samples and different raters using the same instrument, Kamphaus and colleagues reported nine- and seven-cluster solutions for parent and teacher ratings, respectively.

Study Limitations

Results need to be interpreted within the context of the methodology. One caveat relates to the preciseness of the diagnostic information. Without evaluating every child individually, it is impossible to confirm who met criteria for Autistic Disorder, Asperger's Disorder, or PDDNOS. Despite the fact that a third of the sample was not labeled autistic by their IEP team, the actual presence of PDDs is likely not a significant concern for several reasons. School districts in Ohio have been apprehensive about offering educational services without a psychiatric diagnosis. Furthermore, Ohio does not have a primary and secondary disability data system. Many children who present with comorbid disorders will be labeled "multiply handicapped" rather than autistic. Many students in the current study were preschoolers and classified as "preschoolers with a disability." Finally, there are financial pressures for districts not to label students with autism, as there are assumptions that the autism disability category will lead to expensive services and litigation.

Given the context of the data collection, we feel confident that the 2% of participants for whom no parental confirmation of a diagnosis was available also had some type of PDD. They were treated as having PDDs by their schools and, at the very least, would have to present with sufficient autistic features to make them eligible to receive educational services.

Two recent epidemiological studies are relevant to the issue of educational classification. Yeargin-Allsopp et al. (2003) reported 100% agreement between CDC expert case review and previous PDD diagnosis used in concert with previous autism special education eligibility. This is not to say that all children with PDDs were treated as such in their IEP. In this study, only 41% of the students were classified as autistic, while 59% had other special education eligibility categories such as MR and speech and language impairments. A study by Bertrand et al. (2001) found that 50% of PDD students received services under various classifications such as language/communication impairments, multiple disabilities, or preschoolers with a disability.

Another limitation pertains to the instrument used. Although there were significant advantages to using a standardized behavior rating scale, the subgroups derived were necessarily limited by the range of items. An instrument with only 76 items will clearly not identify all possible behavioral and emotional problems. Furthermore, because only a moderately large community sample was used, it is likely that significant but uncommon subgroups were not detected.

Along the same lines, results were dependent on the accuracy of information obtained through informants. The investigator had no control over the accuracy of the information provided or over school policies (e.g., absence of behavior plans as part of the IEP does not necessarily mean that there were no behavior problems).

A final comment relates to the cluster analysis. Results of cluster analysis are often treated skeptically, and we recognize that different clustering methods could have yielded different results (e.g., see Milligan, 1996; Stevenson, 1989). Having said that, the clusters were clinically meaningful, showed some internal and interrater reliability, and they were partially supported by variables external to the analyses. There is always a threat of artifactual findings with cluster analysis, and the number of subgroups identified is a function of the subgroups present in the sample. No direct assessment of accuracy is possible; if the number of groups were known, there would be no need for cluster analysis.

Conclusions

Despite the limitations to the findings, the results have important implications. Although a substantial proportion of young people with PDDs present with relatively low rates of behavior and emotional problems, many of them have serious behavior problems. Furthermore, psychiatric symptomatology does not present in a single fashion. For instance, symptoms of ADHD can or cannot be

accompanied by other behavior and emotional problems. The data also suggested that there are similarities and differences in the dimensions of psychopathology as compared to populations of children with and without MR.

Dimensional approaches provide an opportunity to study behavioral phenomena that are not captured with categorical systems. They promote a greater understanding of the full range of child behavior by virtue of their sensitivity to sub-clinical symptoms. Some studies have shown that they have better predictive validity than categorical approaches in typically developing children (e.g., Fergusson & Horwood, 1995), which could lead to a better use of existing prevention and intervention strategies. The relative merit of dimensional approaches requires future research, especially with populations in which there is limited evidence that the traditional categorical systems used are even valid.

Although behavior and emotional problems have received more empirical attention as of late, they remain significantly understudied in this field. It would be helpful if future studies replicated these findings and examined the stability and predictive validity of empirically derived clusters. Additional research is also needed on the role of prosocial behaviors in the development of behavioral and emotional problems in this population. These individual differences may have important clinical implications in terms of etiology, responsiveness to treatment, and outcome. In all likelihood, a better comprehension of psychopathology in the PDD population will come by bridging the knowledge from the typically developing and MR literatures and by using multiple empirical approaches in concert such as family approaches, pharmacological probes, and neuroimaging techniques.

Acknowledgments This research was supported in part by funding from the Ohio Department of Education. The author wishes to thank all the project coordinators, teachers, and parents for their cooperation. He is especially grateful to Shiling Ruan and Tom Bishop for their assistance with statistical analyses and to Michael G. Aman for comments on earlier versions of the manuscript.

References

- Aman, M. G., Lam, K. S., & Collier-Crespin, A. (2003). Prevalence and patterns of use of psychoactive medicines among individuals with autism in the Autism Society of Ohio. *Journal of Autism & Developmental Disorders*, *33*, 527–534.
- Aman, M. G., Tassé, M. J., Rojahn, J., & Hammer, D. (1996). The Nisonger CBRF: A child behavior rating form for children with developmental disabilities. *Research in Developmental Disabilities*, *17*, 41–57.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.): *DSM-IV*. Washington, DC: Author.
- Arnold, L. E. (2000). *Contemporary diagnosis and management attention-deficit/hyperactivity disorder*. Newtown, PA: Handbooks in Health Care.
- Bertrand, J., Mars, A., Boyle, C., Bove, F., Yeargin-Allsopp, M., & Decoufle, P. (2001). Prevalence of autism in a United States Population: The Brick Township, New Jersey, Investigation. *Pediatrics*, *108*, 1155–1161.
- Brown, E. C., Aman, M. G., & Havercamp, S. M. (2002). Factor analysis and norms for parent ratings on the Aberrant Behavior Checklist-Community for young people in special education. *Research in Developmental Disabilities*, *23*, 45–60.
- Brown, E., Aman, M. G., & Lecavalier, L. (2004). Empirical classification of behavioral and psychiatric problems in children and adolescents with mental retardation. *American Journal on Mental Retardation*, *109*, 445–455.
- Bruininks, R. H., Woodcock, R. W., Weatherman, R. F., & Hill, B. K. (1996). *Scales of independent behavior – revised: Manual*. Boston: Riverside Publishing Company.
- Bryson, S. E., & Smith, I. M. (1998). Epidemiology of autism: Prevalence, associated characteristics, and implications for research and service delivery. *Mental Retardation and Developmental Disabilities Research Reviews*, *4*, 97–103.
- Einfeld, S. L., & Tonge, B. J. (2002). *Manual for the developmental behaviour checklist* (2nd ed.). Clayton, Melbourne, and Sydney: Monash University Center for Developmental Psychiatry and School of Psychiatry, University of South Wales.
- Fergusson, D. M., & Horwood, J. (1995). Predictive validity of categorically and dimensionally scored measures of disruptive childhood behaviors. *Journal of the American Academy of Child and Adolescent Psychiatry*, *34*, 477–487.
- Gadow, K. D., DeVincent, C. J., & Azizian, A. (2004). Psychiatric symptoms in preschool children with PDD and clinic and comparison samples. *Journal of Autism and Developmental Disorders*, *34*, 379–393.
- Gadow, K. D., DeVincent, C. J., Pomeroy, J., & Azizian, A. (2005). Comparison of DSM-IV symptoms in elementary school-aged children with PDD versus clinic and community samples. *Autism*, *9*, 392–415.
- Gillberg, C., & Billstedt, E. (2000). Autism and Asperger syndrome: Co-existence with other clinical disorders. *Acta Psychiatrica Scandinavica*, *102*, 321–330.
- Hammer, D., & Lecavalier, L. (2003). *Successful models and demonstration projects in autism spectrum disorders*. Columbus, OH: Ohio Department of Education.
- Kamphaus, R. W., Huberty, C. J., DiStehano, C., & Petoskey, M. D. (1997). A typology of teacher-rated child behavior for a national U.S. sample. *Journal of Abnormal Child Psychology*, *25*, 453–463.
- Kamphaus, R. W., Petoskey, M. D., Cody, A. H., Rowe, E. W., & Huberty, C. J. (1999). A typology of parent rated child behavior for a national U.S. sample. *Journal of Child Psychology and Psychiatry*, *40*, 607–616.
- Lainhart, J. E. (1999). Psychiatric problems in individuals with autism, their parents and siblings. *International Review of Psychiatry*, *11*, 279–298.
- Lecavalier, L. (2005). An evaluation of the Gilliam Autism Rating Scale. *Journal of Autism and Developmental Disorders*, *35*, 795–805.
- Lecavalier, L., Aman, M. G., Hammer, D., Stoica, W., & Matthews, G. L. (2004). Factor analysis of Nisonger Child Behavior Rating Form in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *34*, 709–721.
- Marshburn, E. C., & Aman, M. G. (1992). Factor validity and norms for the Aberrant Behavior Checklist in a community sample of children with mental retardation. *Journal of Autism Developmental Disorders*, *22*, 357–373.
- Milligan, G. W. (1996). *Clustering validation: Results and implications for applied analyses* (Working Paper Series No. 92–98). Columbus: Ohio State University, College of Business.

- Milligan, G. W., & Cooper, M. C. (1987). Methodology review: Clustering methods. *Applied Psychological Measurement, 11*, 329–354.
- Rojahn, J., Borthwick-Duffy, S. A., & Jacobson, J. W. (1993). The association between psychiatric diagnoses and severe behavior problems in mental retardation. *Annals of Clinical Psychiatry, 5*, 163–170.
- Rojahn, J., & Helsel, W. J. (1991). The Aberrant Behavior Checklist in children and adolescents with dual diagnosis. *Journal of Autism and Developmental Disorders, 21*, 17–28.
- Rojahn, J., Matson, J. L., Lott, D., Esbensen, A. J., & Smalls, Y. (2001). The Behavior Problems Inventory: An instrument for the assessment of self-injury, stereotyped behavior and aggression/destruction in individuals with developmental disabilities. *Journal of Autism and Developmental Disorders, 31*, 577–588.
- Stevenson, J. (1989). Multivariate statistics: II. Using different types of multivariate statistics to identify types of psychopathology: Cluster analysis and multidimensional scaling. *Nordisk Psykiatrisk Tidsskrift, 43*, 325–332.
- Sverd, J. (2003). Psychiatric disorders in individuals with pervasive developmental disorders. *Journal of Psychiatric Practice, 9*, 111–127.
- Tassé, M. J., Aman, M. G., Hammer, D., & Rojahn, J. (1996). The Nisonger Child Behavior Rating Form: Age and gender effects and norms. *Research in Developmental Disabilities, 17*, 59–75.
- Tassé, M. J., Girouard, N., & Morin, I. N. (1999). *La grille d'évaluation comportementale pour enfants Nisonger: manuel de l'utilisateur*. Département de psychologie, Université du Québec à Montréal .
- Tassé, M. J., Morin, I. N., & Girouard, N. (2000). French canadian translation and validation of the Nisonger child behavior rating form. *Canadian Psychology, 41*, 116–123.
- Tonge, B. J., & Einfeld, S. L. (2003). Psychopathology and intellectual disability: The Australian child to adult longitudinal study. In L. M. Glidden (Eds.), *International review of research in mental retardation* (vol. 26, pp. 61–91). San Diego, CA: Academic Press.
- World Health Organization. (1992). *The ICD-10 Classification of Mental and Behavioural Disorder. Clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization.
- Yeargin-Allsopp, M., Rice, C., Karapurkar, T., Doernberg, T., Boyle, C., & Murphy, C. (2003). Prevalence of autism in a US metropolitan area. *Journal of the American Medical Association, 289*, 49–55.