*Journal of the International Neuropsychological Society* (2007), **13**, 50–58.

DOI: 10.1017/S1355617707070087

# Taxonicity of nonverbal learning disabilities in spina bifida

M. DOUGLAS RIS,[1] ROBERT T. AMMERMAN,[1] NIELS WALLER,[2] NICOLAY WALZ,[1] SONYA OPPENHEIMER,[1] TANYA MAINES BROWN,[4] BENEDICTA G. ENRILE,[4] AND KEITH OWEN YEATES[4]

[1]Department of Pediatrics and Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio
[2]Department of Psychology, University of Minnesota, Minneapolis, Minnesota
[3]Division of Neuropsychology, Medical College of Wisconsin, Milwaukee, Wisconsin
[4]Department of Pediatrics and Columbus Children's Hospital, Columbus, Ohio

(Received April 17, 2006; Final Revision July 14, 2006; Accepted August 3, 2006)

**Abstract**

As currently defined, it is not clear whether Nonverbal Learning Disabilities (NLD) should be considered a matter of kind or magnitude (Meehl, 1995). The taxonicity of NLD, or the degree to which it is best construed as discrete *versus* continuous, has not been investigated using methods devised for this purpose. Latent Class Analysis (LCA) is a method for finding subtypes of latent classes from multivariate categorical data. This study represents an application of LCA on a sample of children and adolescents with spina bifida myelomeningocele (SBM) ($N = 44$), those presenting with features of NLD ($N = 28$) but no medical condition, and control volunteers ($N = 44$). The two-class solution provided evidence for the presence of a taxon with an estimated base-rate in the SBM group of .57. Indicator validities (the conditional probabilities of indicator endorsement in each latent class) suggest a somewhat different priority for defining NLD than is typically used by researchers investigating this disorder. A high degree of correspondence between LCA classifications and those based on a more conventional algorithm provided evidence for the validity of this approach. (*JINS*, 2007, *13*, 50–58.)

**Keywords:** Hydrocephalus, Learning disorders, Classification, Myelomeningocele, Neural tube defects, Developmental disabilities

## INTRODUCTION

Nonverbal learning disabilities (NLD) have been described from psychoeducational (Myklebust, 1975), neurological (Denckla, 1978 ; Voeller, 1986) and neuropsychological (Pennington, 1991; Rourke, 1987) perspectives. Though not identical, most of these descriptions emphasize the prominence of deficits in visual-spatial, motor, and math abilities purported to be related to the abnormal functioning of the right cerebral hemisphere and are also often associated with poor interpersonal adjustment (Ris & Nortz, in press). Independent evidence confirming structural or functional brain impairment is often lacking, but it is inferred based on a characteristic pattern of neurobehavioral deficits. In other cases, the critical pattern of strengths and weaknesses may be found in children with documented neurologic disorders or injuries, such as cranial radiation for brain tumors (Buono et al., 1998), traumatic brain injury (Ewing-Cobbs et al., 1993), and hydrocephalus (Donders et al., 1991).

Byron Rourke's (1987) conceptualization of NLD is clearly the most developed, both theoretically and empirically. Central to Rourke's model is the equipoise of *assets* and *deficits* in understanding the development and features of NLD. Classification criteria and rules have been offered that are meant to aid the researcher and clinician in operationalizing the disorder (Pelletier et al., 2001). As originally presented (Rourke, 1987), with later elaboration (Rourke, 1995), the pathophysiology of NLD is to be understood in terms of the integrated functioning of the brain, particularly as it pertains to systems of, or access to, the right cerebral hemisphere. Thus, the NLD constellation may result from dysfunction/dysgenesis of white matter outside

Correspondence and reprint requests to: Dr. M. Douglas Ris, Division of Behavioral Medicine & Clinical Psychology, Children's Hospital Medical Center, 3333 Burnet Ave., Cincinnati, OH 45229. E-mail: douglas.ris@cchmc.org

of the right hemisphere if commissural fibers are involved. A number of recent studies on NLD have investigated "neurobehavioral phenotypes" of patients with various disorders affecting the white matter. Rourke (1995) has grouped these into Levels depending on the extent of conformity to the NLD profile. Level 1 disorders (e.g., hydrocephalus) meet virtually all of the NLD criteria; Level 2 (e.g., congenital hypothyroidism), a majority of them; Level 3 (e.g., traumatic brain injury), many of them; and Level 4 (e.g., neurofibromatosis) are suggestive of NLD. Because actual diagnoses of NLD are often not reported, these studies of high-risk conditions provide evidence of the degree to which such children, *as a group*, are at risk *for particular features*, of NLD, but they cannot determine the *individual risk or base rate* of NLD.

Early shunted hydrocephalus is considered a Level 1 disorder and so, as a group, should show all or most of the signs/symptoms of NLD. Fletcher et al. (1995) present evidence that generally supports this postulate, and the work of the Houston and Toronto groups (Dennis et al., 2005) is exemplary in mapping deficits in motor timing, covert attention, voluntary attention, and perceptual processing to abnormalities in the cerebellum, midbrain, and parietal regions in children with spina bifida myelomeningocele (SBM). However, because NLD has not been an organizing construct for this line of research, it is not clear what proportion of children with SBM could be so characterized and what the anatomic correlates of NLD might be.

In a recent paper by Yeates et al. (2003), the incidence of NLD was compared for groups of children with SBM and healthy siblings. Classification as NLD was based on performances across 11 measures of assets and 17 measures of deficits. Whereas features of NLD were more prevalent in the SBM group, there was also a high degree of phenotypic variability, and only 45% of the children with SBM were classified as NLD.

Another means by which base-rates of a disorder can be estimated is through classification procedures such as Latent Class Analysis. Using binary indicators, LCA defines classes by the criterion of "conditional independence," [i.e., each indicator within a class is statistically independent of every other variable (Clogg, 1995)]. Such procedures can explore the continuous *versus* discontinuous nature of disorders, a fundamental but elusive distinction often determined by fiat rather than empirically (Meehl, 1995). Yet, the taxonic nature of a disorder can have important implications for models of pathogenesis/pathophysiology, genetic contribution, subtype analysis, and distinctions from similar conditions.

The "taxonic question," then, is whether the latent property is a single distribution or is composed of two or more groups (Meehl, 2004). It is a common misconception that taxonicity is reflected in bimodality of the indicator distributions. It has been established that this is true only in extreme cases in which taxon and complement distributions are separated by at least two standard deviations (Murphy, 1964), and so apparently unimodal distributions may obscure taxa. Other statistical classification methods, such as cluster analysis and inverse factor analysis have also been shown to be problematic for distinguishing disorders of "kind" from "magnitude" (Cleland et al., 2000; Golden & Mayer, 1995).

The taxonic/class nature of disorders such as autism (Szatmari et al., 1995), dissociative disorder (Waller et al., 1996), schizophrenia (Erlenmeyer-Kimling et al., 1989), antisocial personality disorder (Bucholz et al., 2000), and bulimia nervosa (Duncan et al., 2005) have been investigated as well as infant attachment patterns (Fraley & Spieker, 2003). However, appropriate statistical methods for determining the discrete *versus* continuous nature of neurobehavioral syndromes/disorders have rarely been applied.

In this study, we used LCA to investigate the taxonic *versus* spectral nature of NLD in a mixed sample of: (1) children with features of NLD but no medical condition; (2) volunteers without an identified developmental or psychological condition; and (3) children with SBM. We then compare the classifications *via* LCA with classifications achieved by a more traditional approach (Yeates et al., 2003). Finally, for participants with SBM, we compared NLD with Non-NLD groups for neurologic risk factors. Children with SBM are known to have a number of brain abnormalities involving gray and white matter (Fletcher et al., 2000) including deformation of the cerebellar tonsils, elongation of the pons and medulla, agenesis of the corpus callosum, aqueduct abnormalities, and tectal "beaking." Eighty to ninety percent of individuals with SBM have hydrocephalus requiring shunt placement (Reigel & Rotenstein, 1994). According to Rourke (1995), this then represents a Level 1 disorder, and so such children are considered to be at high risk for the NLD constellation of symptoms.

## METHOD

### Participants

The sample was comprised of participants in a longitudinal study of the neuropsychological and psychosocial functioning of pre-adolescents and adolescents (ages 10–18 years) with SBM. The SBM participants were recruited from longstanding specialized clinics at two major pediatric care centers (Cincinnati Children's Hospital Medical Center (CCHMC) and Columbus Children's Hospital). Inclusion criteria for this study stipulated that participants have IQs at or above 70 on either the Verbal or Performance scales of the age-appropriate Wechsler Scale. For the purposes of the LCA reported here, participants with SBM ($N = 44$) were pooled with those ($N = 28$) having *features* of NLD (VIQ>PIQ by at least 10 points, Reading>Math by at least 10 points, and evidence of fine motor deficits on visual-motor testing) but no specific medical disorder. There were two recruitment pathways for the NLD features group. First, children who had the Verbal-Performance split were identified through clinician referrals and review of charts at three CCHMC clinics: the Cincinnati Center for Develop-

https://doi.org/10.1017/S1355617707070087 Published online by Cambridge University Press

mental Disorders, the Division of Psychology, and the Division of Psychiatry. In this way, we were able to capture patients referred because of learning concerns as well as those referred for behavioral/emotional concerns. Some participants so identified had testing available for review that allowed us to determine whether they met the other two criteria (Reading>Math, fine motor deficits) whereas for other participants, this testing had to be completed prior to inclusion in the study.

The identification of the NLD features group was based on the design of the larger study, but serves the purposes of this latent class analysis by increasing the range of indicator values to include those that would fall in the clinical range. It is important to note that the screening methods described before do not constitute a diagnosis of NLD. Many more symptoms/indicators of NLD are required for such a diagnosis, as described by Pelletier et al. (2001) and Yeates et al. (2003). Our screening method, though, served the purpose of identifying participants among a heterogeneous clinically referred population with a greater than chance probability of having NLD, thus ensuring a minimally sufficient base rate of the disorder for application of latent class procedures.

Forty-four controls were also recruited from area pediatric practices. The parents of these volunteers were administered a brief screening questionnaire over the phone in an attempt to exclude youths with diagnosed psychological or developmental problems. Therefore, the overall sample size was 116 and the sample was quite heterogeneous in regards to neurobehavioral functioning. All three groups met age (10–18 years) and IQ (at or above 70 on either the Verbal or Performance Scales) inclusion criteria. As can be seen in Table 1, groups differed in age, gender make-up, and SES.

Data included in this study was obtained in compliance with guidelines of the Helsinki Declaration and was approved by the Cincinnati Children's Hospital Medical Center Institutional Review Board (CHMC #99-11-21).

**Table 1.** Sample characteristics

| | SBM ($N = 44$) | NLD Features ($N = 28$) | Healthy Controls ($N = 44$) |
|---|---|---|---|
| Mean Age (SD)* | 13.4 (2.57) | 12.7 (2.11) | 12.2 (2.29) |
| Mean Family SES[a] (SD)* | 48.2 (10.7) | 41.9 (11.1) | 50.2 (9.5) |
| Gender (% males)* | 43.2 | 71.4 | 47.7 |
| Number shunted (%) | 40 (91%) | | |
| Ethnicity: | | | |
| White | 41 | 25 | 42 |
| African-American | 2 | 3 | 2 |
| Asian | 1 | 0 | 0 |

*Note*. [a]Socioeconomic status according to Hollingshead's Four Factor Index of Social Status. There were three cases (one in each group) for which incomplete data precluded derivation of SES.
*$p < .05$.

## Latent Class Analysis

LCA is a method for finding subtypes of latent classes from multivariate categorical data. Mathematically, it is closely related to a form of cluster analysis called multivariate mixture estimation (Titterington et al., 1985). Its goal is to find the minimum number of latent classes accounting for variation in observed indicators (Dayton, 1999; McCutcheon, 1987; Rindskopf & Rindskopf, 1986; Young, 1983).

This goal is accomplished by partitioning a set of response vectors into $k$ latent classes by maximizing the likelihood of the data given the conjectured model. A model will include 2 sets of parameters: (1) the latent class prevalence rates and (2) the conditional indicator probabilities for each class. The overall likelihood for the model is the product of the $N$ individual likelihoods (or the sum of the log likelihoods). The method of maximum likelihood (ML) searches for parameter estimates that maximize the overall likelihood (see Everitt, 1984, for details). These so-called ML parameter estimates can then be used to assign individuals to a latent class.

In the current study class indicators were identified *a priori* in the following manner. Guided by Rourke's (1995) published research, we preselected from a larger neuropsychological battery indicators from each of the domains of NLD. These measures were part of a larger battery administered to the participants and their parents during a five to six hour visit to the medical center. Testing was conducted under controlled conditions by experienced psychometrists and neuropsychology postdoctoral fellows. More information about the test battery and assessment procedures are available on request from the first author (MDR).

1. Wechsler Intelligence Scale for Children, Third Edition (WISC-III: Wechsler, 1991) Verbal IQ>Performance IQ by at least 10 points (designated "IQ" indicator). The Wechsler Adult Intelligence Scale, Third Edition (WAIS-III: Wechsler, 1997) was used instead of the WISC-III for five subjects who were over 16 years of age.

2. Wechsler Individual Achievement Test (WIAT: Wechsler, 1992) Basic Reading Subtest > Numerical Operations by at least 8 points (designated "ACH" indicator).

3. Judgment of Line Orientation (Benton et al.,1983) at least 1 SD below age-mean (designated "JLO" indicator).

4. Children's Category Test (Boll, 1993) at least 1 SD below the mean (designated "CCT" indicator). There were five subjects age 17 years for which the 16-year-old norms (the oldest available for this test) were applied.

5. Grooved Pegboard Test (Klove, 1963) at least 1 SD below the mean for the two hands combined (designated "Gpegs" indicator).

6. Fingertip Number Writing Test (Reitan & Wolfson, 1985) at least 2 SD below the mean bilaterally (designated "FTNW" indicator).

7. Paralanguage score on the Developmental Assessment of Nonverbal Accuracy-2 (Nowicki & Duke, 1994) at least 1 SD below the mean (designated "DANVA" indicator).

8. Behavioral Assessment System for Children (Reynolds & Kamphaus, 1992) Internalizing T score of at least 60 (designated "BASC" indicator).

Wherever possible (i.e., WISC-III, WIAT), cutoffs could be easily drawn or extrapolated from Rourke's published works (1995). In other cases, judgments were made based on the obtained distributions when they were non-normal (i.e., FTNW), and also based on what would be considered of clinical significance (i.e., GPegs).

The resulting set of indicators is quite consistent with contemporary conceptualizations of NLD, and represent the diverse domains and patterns of scores believe to be central to the disorder (Rourke, 1995). While acknowledging the pitfalls of dichotomous indicators (MacCallum et al., 2002), in this case, these were preferred over continuous ones because we were attempting to approximate the clinical diagnostic situation in which discrete decisions are made, and this approach provided the best comparison to already published reports on the classification of NLD (Pelletier et al., 2001; Yeates et al., 2003).

## Concurrent Validity

The classifications achieved using LCA were compared to those using a more conventional method described in a recent publication on NLD in SBM (Yeates et al., 2003). These authors grouped neuropsychological tests into assets and deficits, and NLD classification was based on the proportion of one to the other. Somewhat different tests were used by us, but we attempted to follow the Yeates et al. approach as closely as possible. Listed in Table 2 are the tests representing NLD assets and deficits for our study.

Consistent with the Yeates et al. (2003) approach, an "NLD Total Score" was obtained by adding the percent assets with the percent deficits. An "NLD Difference Score" was generated by subtracting the percent deficits from the percent assets. A subject was classified NLD if both the NLD Total Score was $>.9$ *and* their NLD Difference Score was $<.5$.

## Neurologic Risk

Finally, for the SBM group, those classified as NLD were compared to those not classified as NLD on a neurologic risk index composed of the following: (1) lesion level (thoracic, lumbar, sacral corresponding to assigned values of 3, 2, 1, respectively); (2) seizures (1 or 0); (3) shunt infection (1 or 0); (4) number of shunt revisions (0 to 9); and (5) number of oculomotor/visual deficits (e.g., strabismus, nystagmus, papilledema, optic atrophy) (0 to 4). See Table 3 for more information about the composition of this index. Our approach to scaling overall neurologic risk borrows from Hommeyer et al. (1999), as well as Yeates et al. (2003), and has been used by us in a study of executive functioning in SBM (Brown et al., 2007). Information for the neurologic risk index was obtained from the parent and the medical chart, and was therefore acquired indirectly and retrospectively.

# RESULTS

## Latent Class Analysis

All analyses reported in this section were conducted with the LCA 1.1 package for latent class analysis (Waller, 2004). This library is a collection of *R* (R Development Core Team, 2005) functions for exploring typological models with binary items. The LCA library was chosen for this study because it includes a number of unique features that are useful for analyzing neurological symptom data. Foremost among these are parametric and nonparametric bootstrapping options (Efron, 1984; Langeheine et al., 1996) that allow for the accurate estimation of test statistics (Cressie & Read, 1984) in moderately sized samples, and the ability to execute mul-

**Table 2.** Tests Used to Measure NLD Assets and Deficits

| Assets | Deficits |
|---|---|
| Verbal Fluency | Judgment of Line Orientation |
| WISC-III[a]/WAIS-III[b] Vocabulary | WISC-III/WAIS-III Block Design |
| WISC-III/WAIS-III Similarities | WISC-III/WAIS-III Object Assembly |
| WISC-III/WAIS-III Information | WISC-III Coding/WAIS-III Digit Symbol |
| WISC-III Digit Span | CPT Commissions[c] |
| CMS Stories[d]/WMS Logical Memory[e] | CMS Faces/WMS Family Pictures |
| CMS Word Pairs/WMS Verbal Paired Associates | Children's Category Test |
| WIAT Basic Reading[f] | WIAT Numerical Operations |
| WIAT Spelling | Fingertip Number Writing |
| | Grooved Pegboard |

*Note*. [a]Wechsler Intelligence Scale for Children—III, [b]Wechsler Adult Intelligence Scale—III, [c]Conners Continuous Performance Test, [d]Children's Memory Scale, [e]Wechsler Memory Scale, [f]Wechsler Individual Achievement Test.

**Table 3.** Composition of the Neurologic Risk Index

| | Number | Mean (SD) | Range |
|---|---|---|---|
| Lesion level[a] | | | |
| Thoracic | 9 (23%) | | |
| Lumbar | 25 (64%) | | |
| Sacral | 5 (13%) | | |
| Seizures[a] | 7 (18%) | | |
| Shunt infection[b] | 3 (8%) | | |
| Number of shunt revisions[b] | | 1.53 (1.9) | 0–9 |
| Oculomotor/visual deficits[a] | | .9 (1.0) | 0–4 |

*Note*. [a]data available on 39 of the 44 SBM participants
[b]data available on 36 of the 44 SBM participants.

tiple runs from random starting points. This latter feature represents a useful approach to avoid local maxima in the maximum likelihood solution (for a discussion of local maxima in latent class analyses see Aitkin et al., 1981 and Goodman, 1974). In the following analyses we used a parametric bootstrap because, of the two options, research (De Menezes, 1999) suggests that the parametric method is preferable in moderately sized samples.

Three models were tested and compared for their ability to account for the observed data: a 1-class, 2-class and 3-class model. All criteria indicated that the 2-class model provided a superior fit to the data. To gauge this fit we consulted three statistics: (1) the likelihood ratio chi-square, (2) the Pearson chi square, and (3) the Cressie-Read index (Cressie & Read, 1984). Importantly, we did not rely on their putative chi-square distributions to assess significance levels as research indicates that in realistically sized samples these indices are not distributed as chi-square variates (Collins et al., 1993; Cressie & Read, 1984). Rather, we used the results from 200 parametric bootstrap samples to construct empirically justifiable null distributions. Moreover, each of the 200 bootstrap samples was analyzed 10 times from a random starting configuration to avoid local maxima. Thus, for each model (1-, 2-, or 3-class model) we performed 2000 latent class analyses.

For the 1-class model all fit indices were highly significant (Pearson $\chi^2 = 302.76$, $df = 247$, parametric bootstrap $p < .01$; Likelihood Ratio $\chi^2 = 239.57$, $df = 247$; parametric bootstrap $p < .001$; Cressie-Read: 280.92, $\lambda = .67$, $df = 247$; parametric bootstrap $p < .001$) suggesting that this model could not account for the observed data. Similarly, the 3-class model provided a poor fit to the data. Importantly, we did not rely on the aforementioned chi-square statistics to make this judgment. Rather we noticed that in the 3-class model four parameter estimates were located at a boundary value (i.e., at either 0 or 1). In latent class models, solutions with so-called boundary estimates are *prima facie* evidence that the data have been over-fit by extracting too many latent classes. Consequently, although the fit indices suggested that the 3-class model provided an acceptable level of fit (Pearson $\chi^2 = 159.21$, $df = 229$, parametric bootstrap $p = .27$; Likelihood Ratio $\chi^2 = 150.92$, $df = 229$, parametric bootstrap $p = 0.1$; Cressie-Read $= 161.41$, $\lambda$: .67, $df = 229$, parametric bootstrap $p = .19$), the boundary values indicated that the chi-square values should not be trusted.

Overall, as judged by the chi-square measures of fit with the robust (bootstrap) significance levels, the 2-class model provided an excellent fit to the data (Pearson $\chi^2 = 199.27$, $df = 238$, parametric bootstrap $p = .3$, Likelihood Ratio $\chi^2 = 178.64$, $df = 238$, LR Parametric bootstrap $p = .08$; Cressie-Read $= 195.87$, $\lambda$: .67, $df = 238$, parametric bootstrap $p = .16$). Further analyses of the results bolstered this conclusion. For example, in the 2-class model none of the Freeman-Tukey residuals (Freeman & Tukey, 1950) was larger than 1.96 in contrast to 8% of the residuals for the 1-class model and 3% of the residuals for the 3-class model. Figure 1 provides a graphical summary of the 2-class solution. Notice in this figure that, for each indicator, we have plotted the class membership endorsement probabilities with the accompanying 95% confidence interval. Therefore, JLO, CCT, Gpegs, and FTNW were significantly different for the taxon and complement groups, whereas IQ, ACH, DANVA, and BASC were not. These indicator probabilities denote the likelihood of each indicator being present for participants in the two classes (NLD, non-NLD).

The estimated taxon base-rate was .38 for the entire sample. In the SBM group, 25 (57%) subjects were assigned to the taxon. Of the four participants with SBM who were not shunted, two were assigned to the taxon and two were assigned to the complement group. In the NLD features group, 16 (57%) were assigned to the taxon, and for the control group, 3 (7%) were assigned to the taxon. This latter finding is quite interesting, given that these participants were recruited to be without known developmental/psychological abnormalities. This likely reflects the underidentification of NLD, which is not represented in any educational or psychiatric nosologies (Ris & Nortz, in press). To determine if LCA classifications differed by age, gender, and SES, simple contrasts (one-way ANOVA for age and SES, Chi Square for gender) were carried-out for each of the 3 groups (SBM, NLD features, and controls). For all such analyses, differences were not significant ($p > .05$).

## Concurrent Validity

As can be seen in Table 4, there is good correspondence between the classifications resulting from application of the LCA and Yeates et al. (2003) systems.

**Table 4.** LCA and Yeates et al. Classifications

| | Yeates et al. | | |
|---|---|---|---|
| | Not NLD | NLD | Total |
| LCA | | | |
| Not NLD | 63 (88%/82%)[a] | 9 (13%/23%) | 72 (62%) |
| NLD | 14 (32%/18%) | 30 (68%/77%) | 44 (38%) |
| Total | 77 (66%) | 39 (34%) | 116 |

*Note*. [a]percentages correspond to row (LCA)/column (Yeates et al.)

https://doi.org/10.1017/S1355617707070087 Published online by Cambridge University Press

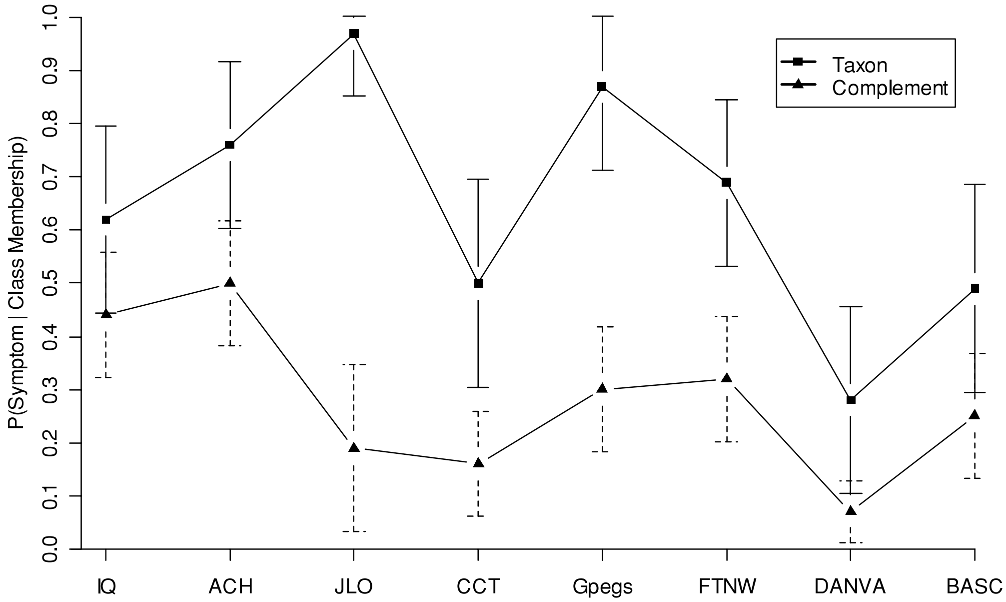


**Fig. 1.** Within class indicator probabilities for entire sample. IQ = VIQ > PIQ, ACH = Reading > Math, JLO = Judgment of Line Orientation, CCT = Children's Category Test, Gpegs = Grooved Pegboard Test, FTNW = Fingertip Number Writing Test, DANVA = Development Test of Nonverbal Accuracy, BASC = Behavioral Assessment System for Children.

The overall classification agreement was 80% with a significant Kappa statistic (.57, $p < .001$). Out of 23 disagreements, in 14 cases LCA classified as NLD whereas the Yeates et al. method classified as Not-NLD. In 9 cases, the reverse was true. In addition to suggesting that the Yeates et al. method is somewhat more conservative, further examination of the individual indicators in these disagreements suggests the following:

1. In the Yeates et al. scheme, IQs and subtest scores determined the classification to a greater extent than in the LCA method, whereas in the LCA classification, visual perceptual and motor scores played a more prominent role. In other words, disagreements were characterized by systematic differences between the two classification systems as to which tests defined NLD.
2. Because the Yeates et al. scheme did not include an indicator reflecting internalizing symptoms, this was relevant only for the LCA classification, being present in 10 of the 14 disagreements in which LCA classified as NLD.
3. Interestingly, the two schemes classified three *different* controls as NLD. On further analysis, this seemed to be related to the aforementioned greater weighting of the Yeates et al. system for IQ and subtest indicators, and a greater weighting of the LCA method for somatosensory and motor indicators of NLD.

## Neurologic Risk

In the SBM group, a neurologic risk index could not be calculated for eight subjects because of missing data on one or more of the five variables that comprise this index. For the remaining 36 SBM participants, one-way ANOVA contrasting the LCA taxon and complement groups on the neurologic risk index was non-significant ($F(1,34) = .486$, $p = .491$) indicating no difference in the NLD and non-NLD groups in terms of presence of neurologic abnormalities/risk factors. The same result was arrived at using a non-parametric test (Mann-Whitney $U = 152$, $p = .948$).

# DISCUSSION

This mixed sample comprised of participants with a high risk condition (SBM) for NLD, participants with some features of NLD, and volunteers without any identified psychological or developmental conditions provided an ideal opportunity for application of latent class procedures. The results of these analyses support the taxonicity of NLD. A distinct class distribution is evidence for a natural boundary between those with and without the disorder. Its distinctiveness from similar disorders such as Aspergers, however, awaits further study.

These results neither support nor refute Rourke's White Matter Hypothesis. Half of the SBM group, a condition at high risk for white matter abnormalities, was classified as having NLD. But as has already been pointed out, this condition is characterized by multiple brain abnormalities and significant phenotypic heterogeneity. Without radiographic evidence, it cannot be determined which such abnormalities are related to the NLD phenotype in particular. The nature of the neurobehavioral phenotype conveys little about the nature (discrete *vs*. continuous) of its substrate. It would not necessarily follow, for example, that a taxonic pheno-

https://doi.org/10.1017/S1355617707070087 Published online by Cambridge University Press

type would imply a discrete neural substrate as opposed to one that is dimensional (i.e., a gradient of white matter abnormalities). One could postulate, for example, a threshold effect consistent with Satz's theory (Satz, 1993) whereby a threshold of white matter disturbance must be reached before clear impairment is manifested on neurobehavioral measures, resulting in a taxonic phenotype. Alternatively, the distinctiveness of the taxon may be related to specific brain abnormalities or even factors other than underlying neuropathology, such as genetic risk or contextual factors. Given the small number of participants with SBM who were not shunted, their even distribution across taxon and complement groups demonstrates little other than the likely complex relationship between phenotype and neural substrate. Further research into the correspondence between empirically-valid classifications and such factors would advance our understanding of root and contributing causes of this neurobehavioral disorder.

Although a few of the indicator probabilities are not significantly different from one another for the taxon and complement groups (see confidence bounds in Fig. 1), the power to detect significant differences is modest in our sample size. The relative validity estimates of the indicators are consistent with Rourke's proposed "Primary Deficits" in his developmental model with sensory-motor and visual-perceptual deficits being precursors of the downstream deficits reflected in IQ and achievement patterns (Rourke, 1995). In this respect, our results using LCA seem to converge with those reported by Harnadek and Rourke (1994) who used discriminant function analysis on a sample comprised of NLD, reading/spelling disabled, and non-disabled children. Motor and visual-spatial deficits may therefore constitute cardinal features of NLD that are present early in development and have only stochastic relationships with other putative features of NLD (e.g., academic patterns, IQ patterns, and socioemotional deficits). This would imply that many of the indicators of NLD frequently used in the literature are not the most reliable and so would contribute to diagnostic imprecision. For example, the diagnostic priority assigned by Pelletier et al. (2001) to performance on the Grooved Pegboard Test is lower than found in our analysis. One benefit of further latent class research of this type would be to establish which indicators are most valid and therefore can be used to improve diagnostic accuracy.

The good correspondence reported here between the LCA and a more traditional approach (Yeates et al., 2003) provides further support for the validity of the LCA results. It is also noteworthy that the base-rate of NLD in the SBM group was very similar to that reported by Yeates et al. in a non-overlapping sample of children with SBM. The somewhat higher base-rate reported here (.57) compared with that reported by Yeates et al. (.45) likely reflects sampling differences; we excluded subjects with IQs below 70 whereas Yeates et al. did not. Furthermore, as reported by Yeates et al., there was no significant relationship found between NLD and an index of neurologic risk for the SBM group. At first, this might seem to conflict with reports of more intracranial imaging abnormalities, broader neuropsychological impairment, and greater metacognitive impairment associated with higher lesion levels (Brown, et al., 2006; Fletcher et al., 2005). But, this relationship pertains to degree of neuropsychological impairment, whereas NLD is defined by a distinct configuration of neuropsychological strengths and weaknesses. As such, NLD should not bear a direct (linear) relationship to neurologic abnormality. Indeed, severe central nervous system compromise would likely "wash out" the pattern that characterizes NLD. Whereas the precision of a composite neurologic index could also be called into questioned, the fact that it relates in a meaningful way to metacognitive functioning (Brown, et al., 2006) provides some evidence of its validity in scaling neurologic risk.

Of the NLD features group, 57% were assigned to the taxon. This is not surprising because the three tests used to select this group were similar to three of the eight class indicators. Because only one of these three (a measure of fine motor speed and dexterity) was among the more discriminating indicators for the NLD class, improved "screening" could be achieved by a better set of indicators, such as GPegs, JLO, and FTNW.

The limitations of this study should be acknowledged. First, whereas the sample size was sufficient, it remains desirable to have larger samples to yield the most robust results. Second, the indicators selected, though true to the NLD concept, were not exhaustive. Other indicators, or other ways of measuring the latent variables of interest, may result in different outcomes. Third, the participants studied do not represent the full age-range to which NLD diagnoses are applied. Therefore, the taxonicity of NLD and corresponding indicator validities cannot be assumed to generalize to younger and older populations. Fourth, validation of the LCA results against an external criterion (i.e., a variable not included here along which NLD and non-NLD groups differ) would lend further credence to these results. Finally, 93% of our participants with SBM were White, and so generalization to other ethnic groups should be made with caution in light of the Fletcher et al. (2005) report of different phenotypes in Whites and Hispanics.

Whereas the best evidence of taxonicity is to be found in further consistency testing (i.e., replication with classification procedures other than those used here), there is reason to believe that the NLD taxon described in this paper is valid. This assertion is based on the close correspondence achieved between two different procedures as well as the use of indicators that, though dichotomized, were drawn from tests that are not known to be "peaked," [i.e., are able to discriminate across a wide range of ability levels (Golden & Mayer, 1995)].

It should also be noted that the inclusion of an NLD features group provided a conservative test of this latent class approach by demonstrating that a distinct class could be discerned, not only among a non-referred sample but also clinical cases that resembled NLD. This methodology, therefore, simulates the diagnostic situation in which the clinician must identify a specific disorder and not just distinguish normality from abnormality.

This study represents a first attempt at applying LCA to a disorder that is still in the scientifically "formative" stage. These results provide support for some of the major tenets of Rourke's NLD model, but also point to some areas needing refinement. Moreover, they suggest that NLD should be considered a disorder with discernable boundaries and not just a gradient of severity with no "natural" demarcation from normality. Further research distinguishing essential from non-essential features of NLD in various risk groups would contribute much to future modeling efforts by directing attention to specific neural mechanisms, refining classification methods, and suggesting critical functions for early intervention.

## ACKNOWLEDGMENTS

The information in this manuscript and the manuscript itself is new and original and has never been published either electronically or in print. There are no financial conflicts of interest. This research was supported by a grant from the U.S. Department of Education (H133G000134) and in part from a grant to the CCHMC General Clinical Research Center (M01 RR08084-09)

## REFERENCES

Aitkin, M., Anderson, D., & Hinde, J. (1981). Statistical modeling of data on teaching styles (with discussion). *Journal of the Royal Statistical Society A.*, *144*, 419–461.

Benton, A., Hamsher, K., Varney, N., & Spreen, O. (1983). Judgement of line orientation. In *Contributions to neuropsychological assessment: A clinical manual* (pp. 53–64). Oxford: Oxford University Press.

Boll, T. (1993). *Children's Category Test*. San Antonio, TX: Psychological Corporation.

Brown, T.M., Ris, M.D., Beebe, D.W., Ammerman, R.T., Oppenheimer, S., Enrile, B.G., & Yeates, K.D. (2007). Factors of biological risk and reserve associated with executive behaviors in children and adolescents with spina bifida myelomeningocele. *Child Neuropsychology* (in press).

Bucholz, K.K., Hesselbrock, V.M., Heath, A.C., Kramer, J.R., & Schuckit, M.A. (2000). A latent class analysis of antisocial personality disorder symptom data from a multi-centre family study of alcoholism. *Addiction*, *95*, 553–567.

Buono, L.A., Morris, M.K., Morris, R.D., Krawiecki, N., Norris, F.H., Foster, M.A., & Copeland, D.R. (1998). Evidence for the syndrome of nonverbal learning disabilities in children with brain tumors. *Child Neuropsychology*, *4*, 144–158.

Cleland, C.M., Rothschild, L., & Haslam, N. (2000). Detecting latent taxa: Monte Carlo comparison of taxometric, mixture model, and clustering procedures. *Psychological Reports*, *87*, 37–47.

Clogg, C.C. (1995). Latent class models. In G. Arminger, C.C. Clogg, & M.E. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp. 311–359). New York: Plenum.

Collins, L.M., Fidler, P.L., Wugalter, S.E., & Long, J.D. (1993). Goodness of fit testing for latent class models. *Multivariate Behavioral Research*, *28*, 375–389.

Cressie, N. & Read, T.R.C. (1984). Multinomial goodness of fit tests. *Journal of the Royal Statistical Society, B*, *46*, 440–464.

Dayton, C.M. (1999). *Latent class scaling analysis*. Sage University Paper Series on Quantitative Applications in the Social Sciences. 07–126. Thousand Oaks, CA: Sage.

De Menzes, L.M. (1999). On fitting latent class models for binary data. *British Journal of Mathematical and Statistical Psychology*, *52*, 149–168

Denckla, M.B. (1978). Minimal brain dysfunction. In J. Chall & A. Mirsky (Eds.), *Education and the brain* (pp. 223–268). Chicago: National Society for the Study of Education and University of Chicago Press.

Dennis, M., Edelstein, K., Frederick, J., Copeland, K., Francis, D., Blaser, S.E., Kramer, L.A., Drake, J. M., Brandt, M., Hetherington, R., & Fletcher, J.M. (2005). Peripersonal spatial attention in children with spina bifida: Associations between horizontal and vertical line bisection and congenital malformations of the corpus callosum, midbrain, and posterior cortex. *Neuropsychologia*, *43*, 2000–2010.

Donders, J., Rourke, B.P., & Canaday, A.I. (1991). Neuropsychological functioning of hydrocephalic children. *Journal of Clinical and Experimental Neuropsychology*, *13*, 607–613.

Duncan, A.E., Neuman, R.J., Kramer, J., Kuperman, S., Hesselbrock, V., Reich, T., & Bucholz, K.K. (2005). Are there subgroups of bulimia nervosa based on comorbid psychiatric disorders? *International Journal of Eating Disorders*, *37*, 19–25.

Efron, B.S. (1984). A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician*, *37*, 36–48.

Erlenmeyer-Kimling, L., Golden, R.R., & Cornblatt, B.A. (1989). A taxometric analysis of cognitive and neuromotor variables in children at risk for schizophrenia. *Journal of Abnormal Psychology*, *98*, 203–208.

Everitt, B.S. (1984). A note on parameter estimation for Lazarsfeld's latent class model using the EM algorithm. *Multivariate Behavioral Research*, *19*, 79–89.

Ewing-Cobbs, L., Fletcher, J.M., Levin, H.S., & Boudousquie, A. (1993). Nonverbal learning disabilities in children and adolescents following closed-head injury. *Journal of Clinical and Experimental Neuropsychology*, *15*, 41.

Fletcher, J.M., Brookshire, B.L., Bohan, T.P., Brandt, M.E., & Davidson, K.C. (1995). Early Hydrocephalus. In B.P. Rourke (Ed.), *Syndrome of nonverbal learning disabilities* (pp. 206–238). New York: Guildford Press.

Fletcher, J.M., Dennis, M., & Northrup, H. (2000). Hydrocephalus. In K.O. Yeates, M.D. Ris, & H.G. Taylor (Eds.), *Pediatric neuropsychology: Theory, research, and practice* (pp. 25–46). New York: Guilford Press.

Fletcher, J.M., Copeland, K., Frederick, J.A., Blaser, S.E., Kramer, L.A., Northrup, H., Hannay, H.J., Brandt, M.E., Francis, D.J., Villarreal, G., Drake, J.M., Laurent, J.P., Townsend, I., Inwood, S., Boudousquie, A., & Dennis, M. (2005). Spinal lesion level in spina bifida: A source of neural and cognitive heterogeneity. *Journal of Neurosurgery*, *102*, 268–279.

Fraley, R.C. & Spieker, S.J. (2003). Are infant attachment patterns continuously or categorically distributed? A taxometric analysis of strange situation behavior. *Developmental Psychology*, *39*, 387–404.

Freeman, M.F. & Tukey, J.W. (1950). Transformations related to the angular and square root. *Annals of Mathematics and Statistics*, *21*, 607–611.

Golden, R.R. & Mayer, M.J. (1995). Peaked indicators: A source of pseudotaxonicity of a latent trait. In D. Lubinski & R.V.

Dawis (Eds.), *Assessing individual differences in human behavior* (pp. 93–115). Palo Alto, CA: Davies-Black Publishing.

Goodman, L.A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, *61*, 215–231.

Harnadek, M.C. & Rourke, B.P. (1994). Principal identifying features of the syndrome of nonverbal learning disabilities in children. *Journal of Learning Disabilities*, *27*, 144–154.

Hommeyer, J.S., Holmbeck, G.N., Willis, K.E., & Coers, S. (1999). Condition severity and psychosocial functioning in preadolescents with spina bifida: Disentangling proximal functional status and distal adjustment outcomes. *Journal of Pediatric Psychology*, *24*, 499–509.

Klove, H. (1963). Clinical neuropsychology. In F.M. Forster (Ed.), *The medical clinics of North America* (pp. 1647–1658). New York: Sanders.

Langeheine, R., Pannekoek, J., & van de Pol, F. (1996). Bootstrapping goodness of fit measures in categorical data analysis. *Sociological Methods and Research*, *24*, 492–516.

MacCallum, R.C., Zhang, S., Preacher, K.J., & Rucker, D.D. (2002). On the practice of dichotomization of qualitative variables. *Psychological Methods*, *7*, 19–40.

McCutcheon, A.C. (1987). *Latent class analysis*. Beverly Hills: Sage Publications.

Meehl, P.E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist*, *50*, 266–274.

Meehl, P.E. (2004). What's in a taxon? *Journal of Abnormal Psychology*, *24*, 499–509.

Murphy, E.A. (1964). One cause? Many causes? The argument from the bimodal distribution. *Journal of Chronic Disease*, *17*, 301–324.

Myklebust, H.R. (1975). Nonverbal learning disabilities: Assessment and intervention. In H.R. Myklebust (Ed.), *Progress in learning disabilities* (Vol. 3, pp. 85–121). New York: Grune & Stratton.

Nowicki, S., Jr. & Duke, N.P. (1994). Individual differences in the non-verbal communication of affect: The Diagnostic Analysis of Nonverbal Accuracy (DANVA) Scale. *Journal of Nonverbal Communication*, *18*, 9–18.

Pelletier, P.M., Ahmad, S.A., & Rourke, B.P. (2001). Classification rules for basic phonological processing disabilities and nonverbal learning disabilities: Formulation and external validity. *Child Neuropsychology*, *7*, 84–98.

Pennington, B.F. (1991). *Diagnosing learning disorders: A neuropsychological framework*. New York: Guilford.

R Development Core Team (2005). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Reigel, D.H. & Rotenstein. (1994). Spina bifida. In W.R. Cheek (Ed.), *Pediatric Neurosurgery* (3rd ed.) (pp. 51–76). Philadelphia: W. B. Saunders.

Reitan, R.M. & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation*. Tuscon, AZ: Neuropsychology Press.

Reynolds, C.R. & Kamphaus, R.W. (1992). *Behavioral assessment system for children manual*. Circle Pines, MN: American Guidance Service, Inc.

Rindskopf, R. & Rindskopf, W. (1986). The value of latent class analysis in medical diagnosis. *Statistics in Medicine*, *5*, 21–27.

Ris, M.D. & Nortz, M. (in press). Nonverbal learning disorder. In J. Morgan & J. Ricker (Eds.), *Handbook of clinical neuropsychology*. New York: Taylor & Francis Publishing.

Rourke, B.P. (1987). Syndrome of nonverbal learning disabilities: The final common pathway of white matter disease/dysfunction. *The Clinical Neuropsychologist*, *1*, 209–234.

Rourke, B.P. (1995). *Syndrome of nonverbal learning disabilities: Neurodevelopmental manifestations*. New York: The Guildford Press.

Satz, P. (1993). Brain reserve capacity on symptom onset after brain injury: A formulation and review of evidence for threshold theory. *Neuropsychology*, *7*, 273–295.

Szatmari, P., Volkmar, F., & Walter, S. (1995). Evaluation of diagnostic criteria for autism using latent class models. *Journal of the American Academy of Child and Adolescent Psychiatry*, *34*, 216–222.

Titterington, D.M., Smith, A.F.M., & Makov, U.E. (1985). *Statistical analysis of finite mixture distributions*. New York: Wiley.

Voeller, K.K. (1986). Right-hemisphere deficit syndrome in children. *American Journal of Psychiatry*, *143*, 1004–1009.

Waller, N.G., Putnam, F.W., & Carlson, E.B. (1996). Types of dissociation and dissociative types: A taxometric analysis of dissociative experiences. *Psychological Methods*, *1*, 300–321.

Waller, N.G. (2004). *LCA* 1.1 An R package for exploratory latent class analysis. *Applied Psychological Measurement*, *28*, 141–142.

Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children–Third Edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1992). *The Wechsler Individual Achievement Test*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997). *The Wechsler Adult Intelligence Scale*–3rd ed. San Antonio, TX: The Psychological Corporation.

Yeates, K.O., Loss, N., Colvin, A.N., & Enrile, B.G. (2003). Do children with myelomeningocele and hydrocephalus display nonverbal learning disabilities? An empirical approach to classification. *Journal of the International Neuropsychological Society*, *9*, 653–662.

Young, M.A. (1983). Evaluating diagnostic criteria: A latent class paradigm. *Journal of Psychiatric Research*, *17*, 285–296.