Sensorimotor performance in school-age children with autism, developmental language disorder, or low IQ

David E Mandelbaum* MD PhD, Departments of Clinical Neurosciences and Pediatrics, Brown Medical School, Providence, RI;
Michael Stevens PhD, Clinical Cognitive Neuroscience Laboratory, Olin Neuropsychiatry Research Center, Hartford Hospital, Hartford, CT;
Eric Rosenberg MS, Independent Statistical Consultant, New York, NY;
Max Wiznitzer MD, Department of Neurology, Rainbow Children's Hospital and Case Western Reserve University School of Medicine, Cleveland, OH;
Mitchell Steinschneider MD PhD, Saul R Korey Department of Neurology, Albert Einstein College of Medicine, Bronx, NY;
Pauline Filipek MD, Department of Pediatrics, University of California at Irvine, Irvine, CA;
Isabelle Rapin MD, Saul R Korey Department of Neurology, Albert Einstein College of Medicine, Bronx, NY, USA.

**Correspondence to first author at* 110 Lockwood Street, Suite 342, Providence, RI 02903, USA.
E-mail: David_Mandelbaum@brown.edu

The purpose of the study was to determine the prevalence of 'soft' motor deficits in school-aged children with either developmental language disorder (DLD), autism (with high IQ [HiAD] or low IQ [LoAD]), or low IQ without autism (LoIQ), and to evaluate the utility of a refined neurological examination to discriminate between these groups. A total of 242 children (74% male), aged 7 or 9 years, were evaluated as part of a longitudinal, multi-institutional study, with a standardized neurological examination that included Denckla's Physical and Neurological Examination for Soft Signs. Most of the scores separated children into two groups defined by nonverbal IQ, with the DLD and HiAD groups performing better than the LoAD and LoIQ groups. Exceptions included motor impersistence and stereotypies, which were more likely in the autistic groups. The neurologists' summary clinical impressions indicated better sensory/motor skills, oromotor skills, and praxis in the HiAD than in the DLD children. Inability/unwillingness to perform tasks was much more frequent in LoAD than LoIQ children.

See end of paper for list of abbreviations.

There is a long-standing tradition of the assessment of so-called 'soft signs' in developmental disorders in an effort to enhance the sensitivity of the classical neurological examination for detecting subtle, global neurological, or sensorimotor dysfunction (Denhoff et al. 1968, Touwen and Prechtl 1970, Peters et al. 1975). Denckla (1977) distinguished two classes of soft signs: those that are very mild but classic neurological deficits, such as ataxia or minimal lateralized motor deficit (the so-called 'pastel classic' signs), and immature sensorimotor skills characteristic of a younger child ('developmental soft signs'). The Physical and Neurological Examination for Soft Signs (PANESS) quantifies subtle deficits of gross and fine motor function ('soft signs') such as adventitious movements and impersistence during stressed gaits and the maintenance of posture, and movement overflow from one body part to another during timed, rapid, repetitive motor tasks (Denckla 1985). Satisfactory completion of timed tasks requires the integration of attention, memory, response preparation, and inhibition; it is, therefore, useful for evaluating children with executive function disorders.

The autistic spectrum disorders (pervasive developmental disorders in the Diagnostic and Statistical Manual of Mental Disorders 4th edition [DSM IV; American Psychiatric Association 2000] and the International Classification of Diseases [ICD-10; World Health Organization 1992]), referred to as autism or autistic disorder (AD) in this report, are neurodevelopmental disorders defined behaviorally by impaired sociability and empathy, deficient verbal and nonverbal communication and imaginative play, and rigidity, perseveration, and a narrow range of interests. Children with autism, even those without a defined etiology, are likely to have a variety of sensorimotor deficits, namely stereotypies, low muscle tone/increased joint laxity, clumsiness, and apraxia, the neurological basis of which is undefined (Bauman 1992, Rapin 1996b). Children with mental retardation* regularly have a variety of motor deficits, some of which overlap those of children with autism. Some children with developmental language disorders (DLDs), without the social deficits or behavioral abnormalities characteristic of autism, are reported to have delayed motor milestones and mild motor deficits including deficient oromotor function (Haynes and Naidoo 1991).

So far, few comparisons of sensorimotor function among groups of school-aged children diagnosed with autism or language impairment have been reported. A large, multi-institutional study of non-autistic children with DLD or low IQ (LoIQ), and children with autism with high (HiAD) or low nonverbal IQ (LoAD; < 80) provides the opportunity to determine the prevalence of 'soft' motor deficits in each of these developmental disorders, and evaluate the utility of a refined neurological examination, the PANESS, for discriminating between these groups. The study also investigated whether children with DLD, whose language difficulty cannot be attributed to impaired social functioning, might have a greater degree of motor impairment than autistic children with high IQ.

Method

PARTICIPANTS

Four groups of children were evaluated at the ages of 7 years, 9 years, or both. For the present report, 9-year examinations were used, i.e. those of children evaluated at the age of 9 years

*UK usage: learning disability.

only (n=73) and those of the children evaluated at both 7 and 9-years (n=81); 7-year evaluations were used for the 88 children who were studied at the age of 7 years only. When tested by analysis of variance (ANOVA), the groups differed by age ($F_{3,238}$=5.89; p<0.001). Bonferroni posthoc tests showed that the only significant age difference between the groups was that the DLD group was younger than the LoAD group. Seventy-four percent of the children were males, 26% were females. Handedness was similar across the groups (Table I).

Because this was part of a longitudinal study, children remained in the group to which they had been assigned at their preschool assessment. Criteria for preschool group assignments have been described in detail (Rapin 1996a). Children with autism were divided into two groups, HiAD and LoAD. DLD children, all of whom had a nonverbal IQ of 80 or better, scored at least 1 SD below the mean in mean length of utterance for age (Miller et al. 1981) or on the Test of Early Language Development (Hresko et al. 1981), and did not fulfill criteria for an autistic spectrum disorder. Non-autistic LoIQ children had a nonverbal IQ below 80 and did not meet criteria for an autistic spectrum disorder. A test of nonverbal IQ (NVIQ; Stanford Binet nonverbal IQ) or, for nonverbal children, the Leiter International Performance Scale and the Wing Autistic Disorder Interview Checklist (WADIC; Wing 1996) were administered again at school age. The mean scores of the groups at school age conformed well to the classification given on entry into the study, with significant differences between the DLD and HiAD groups compared with the LoAD and LoIQ groups in NVIQ, and significant differences between the AD and non-AD groups on the WADIC p<0.001 on Bonferroni posthoc tests; Table II).

Preschool exclusionary criteria for all children were a hearing deficit that exceeded 20dB in either ear, a major malformation, a gross sensorimotor deficit, such as a hemiparesis or obvious movement disorder, uncontrolled epileptic seizures, large doses of antiepileptic or psychotropic drugs, or a family whose primary language was not English.

Recruitment, evaluation, and follow-up assessments were performed at four sites: Boston, MA, the Bronx, NY, Cleveland, OH, and Trenton, NJ. Sites differed in the type of child they recruited at preschool, except that all sites were to recruit a contingent of LoIQ children. Boston and Trenton recruited children with autism, Cleveland recruited children with DLD, and the Bronx recruited children with either DLD or autism. Although the examining neurologists were aware of the type of child recruited at their sites they were blind to preschool group assignments and to all quantitative psychological, language, and behavioral test data. Parents of all the children signed an informed consent form at each age, and children aged 7 years or above able to do so gave an oral assent or signed an informed assent form. All procedures were approved by the Institutional Review Boards of all four institutions.

NEUROLOGICAL EXAMINATION

At preschool the children had undergone a standardized neurological examination (Rapin 1996a). At the 7-year and 9-year reassessments, much the same standardized neurological examination was administered, with additions appropriate for age, including the PANESS (Denckla 1985). To insure intersite reliability, all examining neurologists had been trained together for the preschool examination and were retrained for the school-age examination. The items of the structured neurological examination were broken down into three categories: (1) tasks that the children were asked to perform on verbal command or to imitate (that is, tasks that required active cooperation, including gross motor, fine motor, and oromotor skills, and sensory tests); (2) observations made by the examining neurologists; and (3) neurologists' summary impressions of their evaluations (Table III). Because no significant cranial nerve abnormality was detected in the preschool assessment, with the exception of oromotor function and facial movement, they are not reported here; nor are the results of the mental status examination. Among the oromotor-timed tasks from the PANESS, tongue side-to-side scores were abnormal in all except four children (and two of the authors); it was, therefore, excluded from the analysis.

Table I: Age, sex, and handedness of four groups of children

Parameter	*DLD (n=89)*	*HiAD (n=42)*	*LoAD (n=74)*	*LoIQ (n=37)*
Age mean (SD), mo	98.0 (12.00)	105.4 (10.93)	107.89 (22.98)	105.62 (10.97)
Male, *n* (%)	63 (70.8)	33 (78.6)	60 (80.8)	22 (61.1)
Right-handed[a], *n* (%)	69 (78.4)	31 (75.6)	43 (78.6)	30 (83.3)

[a]In LoAD group, handedness data in only 74.3% of patients; for other groups, data in more than 97% of patients.
DLD, developmental language disorder; HiAD, high-IQ autistic disorder; LoAD, low-IQ autistic disorder; LoIQ, low IQ without autism.

Table II: Nonverbal IQ and autism scores of four groups of children

Parameter	*DLD (n=89)*	*HiAD (n=42)*	*LoAD (n=74)*	*LoIQ (n=37)*
NVIQ[a]	97 (17)	102 (16)	62 (20)	69 (17)
WADIC[b]	1.11 (2.75)	7.56 (6.31)	12.87 (4.62)	4.86 (5.78)

Results are shown as mean (SD).
[a]$F_{3,226}$=68.114, p<0.001. [b]$F_{3,187}$=64.119, p<0.001.
DLD, developmental language disorder; HiAD, high IQ autistic disorder; LoAD, low IQ autistic disorder; LoIQ, low IQ without autism; NVIQ, nonverbal IQ ratio; WADIC, Wing Autistic Disorders Interview Checklist, total score range 0 to 21.

SCORING OF THE DATA

For each item of the standardized neurological examination, the child's performance was scored 0 for normal,1 for mildly abnormal, and 2 for unequivocally abnormal. For timed standardized tests with age norms, namely the Purdue Pegboard, PANESS, and syllable repetition, we assigned a score of 0 if the child's time was less than 1 SD below the norm, 1 if it was between 1 and 2 SDs below the norm, and 2 if it was at or below 2 SDs from the norm.

Neurologists assigned a specific code to items that the children would not or could not perform. These 'wouldn't do/couldn't do' (WD/CD) scores were merged for the analysis because distinguishing unwillingness from inability to perform was considered too subjective to be reliable. The WD/CD scores were not included in the calculations of mean task scores; giving the children a score of 3 for the WD/CD items would have inflated the deficit scores, perhaps inaccurately, for an unknown fraction of the children. Omitting the WD/CD items from the mean scores used in the analysis probably underestimates the severity of deficits for the tasks requiring active cooperation. This underestimation of severity almost certainly affects the four groups unequally, because the LoAD group had the largest number of WD/CD scores. Scores in the tables, therefore, represent the functional abilities of only those children who performed the tasks.

As summarized in Table III, there were four major domains for which subdomain summary scores were constructed: gross motor skills, fine motor skills, oromotor skills, and sensory evaluation. Each subdomain score was created by grouping and averaging scores of related items from the neurological examination. Each subdomain item (scored 0, 1, or 2) was given equal weight in calculating mean subdomain scores. Adjustment for missing scores (WD/CD or no data) was performed as follows. If a child completed between 50 and 100% of the tasks within a particular subdomain, the mean subdomain score was calculated on the basis of those tasks completed. If scores for fewer than 50% of tasks were available, this subdomain was scored as 'missing'. This rule was adopted to ensure that subdomain scores were representative and not overweighted by one or two tasks, and also to avoid spuriously reduced subdomain scores because of missing data from one or two uncompleted tasks. All related subdomains were averaged into a total domain score for gross motor, fine motor, oromotor, and sensory. Because missing data were dealt with at the subdomain level, these data represent the average degree of impairment (on a scale of 0 to 2) for tasks attempted in that domain; they are, therefore, weighted in favor of the subdomains for which there was greater participation. An example of the impact of this scoring method can be seen in the scores for oromotor function: because of the lower percentage of LoAD children performing the more difficult subdomain of timed oromotor tasks compared with the easier subdomains of non-timed tasks and facial movements, the domain score for the LoAD group is more reflective of the easier tasks than for the DLD and HiAD groups. This probably explains the fact that the oromotor domain score of the LoAD group was superior to that of children with HiAD and DLD (Table IV).

Neurologists were required to score their overall clinical impressions with the same metric as the individual items: 0 (normal or trivial impairment of no clinical significance), 1 (mild to moderate abnormality), or 2 (severe abnormality). On the basis of the presumption that the children in the high-IQ groups had less global dysfunction than those in the low-IQ groups, the presence of any abnormality in the high-IQ groups was considered to be significant. A dichotomous analysis of normal (0) versus any abnormality (1 or 2) was, therefore, also performed to compare the DLD and HiAD groups.

The proportion of WD/CD scores for items requiring active participation of the child (stressed gaits, station, timed and

Table III: Summary sensorimotor scores of neurological examination

A. *Gross motor skills*
- Stressed gaits[a]
  - Heel
  - Toes
  - Sides of feet
  - Tandem gait
  - Hopping
- Station (standing on one foot, eyes open)
- Balance (tendency to fall; feet together with eyes closed)
- Persistence (ability to maintain posture and keep the eyes closed for a given number of seconds)

B. *Fine motor skills*
- PANESS (timed; Denckla 1985)
  - foot taps (20 toe and 10 heel–toe taps)
  - hand pats (20 flat and 10 pronation–supination)
  - finger taps (20 index–thumb and five sets of sequential each finger–thumb)
- Purdue Pegboard (Gardner and Broman1979)
- Visuo/Graphomotor; untimed
  - building the five-block bridge
  - handwriting (name)
  - copying a sentence
  - drawing a diamond

C. *Oromotor skills*
- Timed (20 repetition of each of three syllables [pa, ta, ka], and 10 repetitions of all three sequentially; Fletcher 1972)
- Untimed (jaw and tongue side-to-side and in/out; tsk; tongue in cheek; purse; kiss)

D. *Sensory evaluation*[b]
- finger localization (Kinsbourne and Warrington 1964)
- graphesthesia
- two-point discrimination (penny versus dime with eyes closed)

E. *Neurologists' observations*
- Stereotypies
- Apraxia
- Oromotor apraxia
- Abnormal movements (tremor; tics; dysmetria on finger to nose test; dysrhythmia; overflow; dystonia; choreoathetosis; Prechtl movements (inability to maintain the fingers immobile with the eyes closed and hands outstretched)

F. *Neurologists' overall impressions*
- Sensorimotor function
- Oromotor function
- Apraxia

[a]Stressed gaits consisted of walking 10 steps on heels, 10 on toes, 10 on sides of the feet, 10 in tandem steps forward, 10 tandem steps backwards, time for 25 hops on each foot (at 7 years) or 50 hops (at 9 years). [b]The Face–Hand Test (Pollack and Goldfarb 1957) yielded only two scores: 0 (normal) and 'wouldn't do/couldn't do'; it was not included in sensory score. PANESS, Physical and Neurological Examination for Soft Signs.

untimed fine motor skills, timed and untimed oromotor skills, and sensory evaluation) were tallied and analyzed for group differences. For this analysis, all these items in the neurological database were rescored as 0 if any score was present and 1 for WD/CD. The sum of these scores was divided by the total number of items to obtain a mean WD/CD index with a range from 0 (attempted all tasks) to 1 (refused or could not do any of the tasks).

Several items of the neurological examination had separate entries for the right and left sides of the body. These included lateralized gross and fine motor tasks, tendon stretch reflexes, abnormal movements, facial movements, and sensory evaluation. A lateralized impairment index was created by collapsing abnormal scores (i.e. scores of 1 or 2) for all lateralized items.

DATA ANALYSIS

All subdomain scores were ordinal variables ranging from 0 to 2. For the purpose of this analysis, these scores were treated as continuous variables to examine mean group differences. This was justified not only by the rigorous inter-rater reliability (Rapin 1996b) but also because these ratings roughly reflect a linearly increasing impairment. One-way ANOVA was applied to evaluate group means for significant differences. Assumptions of normality were met for most variables, both across and within subgroups.

However, consistent skewness was noted for impression scores of timed oromotor, non-timed visual motor, facial movements, stereotypy, oromotor deficit impression, and autistic features. For these variables, the one-way ANOVA was complemented with a Kruskal–Wallis non-parametric ANOVA test to confirm group differences. A probability of $p \leq 0.05$ was considered significant. Bonferroni corrections for multiple comparisons were calculated to evaluate differences among the groups. χ^2 analysis was calculated to evaluate dichotomous data (neurologists' impression ratings and laterality abnormality index). Where appropriate, data were described by using 95% confidence intervals (CIs) around the mean score. Data analyses were performed with SPSS for Windows (version 10.0.7).

Table IV: Mean (SD) performance scores of four diagnostic groups and percentage of children with data in each subdomain

Domain	*Groups*								*F*	*Subgroup differences*
	DLD (n=89)		*HiAD (n=42)*		*LoAD(n=74)*		*LoIQ (n=37)*			
	Mean (SD)	*%n*	*Mean (SD)*	*%n*	*Mean (SD)*	*%n*	*Mean (SD)*	*%n*		
A. Gross motor	0.70 (0.42)		1.01 (0.41)		1.21 (0.48)		1.01 (0.43)		$F_{3,197}=14.38$	a, b, c
Stressed gaits	0.32 (0.32)	99	0.41 (0.33)	88	0.79 (0.54)	39	0.58 (0.34)	78	$F_{3,179}=13.80$	a, b, d
Station/tendency to fall	0.93 (0.57)	99	1.01 (0.64)	93	1.36 (0.68)	43	1.42 (0.43)	84	$F_{3,186}=7.98$	a, b, d
Impersistence	0.86 (0.86)	99	1.48 (0.76)	93	1.37 (0.58)	47	0.88 (0.76)	78	$F_{3,187}=8.01$	a, b, h
B. Fine motor skills	0.7 (0.42)		1.01 (0.41)		1.21 (0.48)		1.01 (0.43)		$F_{3,230}=45.99$	a, b, d, e
Fine-motor PANESS	0.69 (0.51)	99	0.77 (0.62)	100	1.56 (0.60)	66	1.23 (0.67)	86	$F_{3,207}=27.59$	a, b, e
Purdue Pegboard (both hands)	0.80 (0.90)	89	1.38 (0.89)	88	1.25 (0.92)	43	1.61 (0.78)	62	$F_{3,167}=7.11$	a, b
Non-timed Visuo/ Graphomotor	0.41 (0.45)	99	0.39 (0.50)	100	1.28 (0.68)	66	1.32 (0.67)	86	$F_{3,207}=43.06$	a, b, d, e
C. Oromotor	0.44 (0.20)		0.38 (0.22)		0.32 (0.31)		0.55 (0.36)		$F_{3,237}=6.56$	e, f, g
Timed	1.26 (0.58)	92	1.11 (0.59)	88	1.38 (0.51)	53	1.59 (0.43)	73	$F_{3,186}=4.17$	a, e
Untimed	0.17 (0.21)	99	0.17 (0.18)	100	0.45 (0.45)	100	0.34 (0.26)	100	$F_{3,237}=13.41$	a, b, d
Facial movements	0.01 (0.05)	96	0.01 (0.08)	95	0.01 (0.06)	97	0.00 (0.00)	95	$F_{3,228}=0.31$	
D. Sensory domain	0.78 (0.28)	88	0.68 (0.26)	86	0.86 (0.39)	24	1.05 (0.24)	46	$F_{3,145}=6.99$	b, e

Scores: 0, normal; 1, mildly abnormal; 2, unequivocally abnormal. Scores for timed tasks: 0, within 1 SD of norm; 1, more than 1 SD from norm; 2, more than 2 SD from norm. Subgroup differences: a, DLD<LoIQ; b, DLD<LoAD; c, DLD<HiAD; d, HiAD<LoAD; e, HiAD<LoIQ; f, LoAD<DLD; g, LoAD<LoIQ; h, LoIQ<HiAD. *%n*, percentage of children with data in each subdomain. DLD, developmental language disorder; HiAD, high-IQ autistic disorder; LoAD, low-IQ autistic disorder; LoIQ, low IQ without autism; PANESS, Physical and Neurological Examination for Soft Signs.

Table V: Neurologists' mean observation scores

Observations	*DLD (n=89)*		*HiAD (n=42)*		*LoAD (n=74)*		*LoIQ (n=37)*		*F*	*Subgroup differences*
	Mean (SD)	*%n*	*Mean (SD)*	*%n*	*Mean (SD)*	*%n*	*Mean (SD)*	*%n*		
Stereotypies	0.03 (0.24)	76	0.29 (0.62)	83	1.18 (0.85)	77	0.23 (0.62)	84	$F_{3,187}=40.37$	b, d
Abnormal movements	0.19 (0.16)	100	0.18 (0.14)	100	0.27 (0.24)	100	0.34 (0.23)	100	$F_{3,238}=7.41$	a, e
Apraxia	0.08 (0.17)	99	0.01 (0.06)	100	0.22 (0.36)	76	0.31 (0.35)	84	$F_{3,213}=12.22$	a, b, d, e
Oromotor apraxia	0.39 (0.51)	96	0.30 (0.46)	95	0.76 (0.74)	69	0.76 (0.61)	89	$F_{3,205}=8.15$	a, b, d, e

Scores: 0, normal; 1, mildly abnormal; 2, unequivocally abnormal. Subgroup differences: a, DLD<LoIQ; b, DLD<LoAD; c, DLD<HiAD; d, HiAD<LoAD; e, HiAD<LoIQ; f, LoAD<DLD; g, LoAD<LoIQ; h, LoIQ<HiAD. *%n*, percentage of children with data in each subdomain. DLD, developmental language disorder; HiAD, high-IQ autistic disorder; LoAD, low-IQ autistic disorder; LoIQ, low IQ without autism.

Results

Table IV displays the scores for the various domains and subdomains listed in Table III. For variables identified as non-normal, the Kruskal–Wallis ANOVA test confirmed all significant group findings. Most of the scores separated children into two NVIQ-defined groups, with the DLD and HiAD groups performing better than the LoAD and LoIQ groups. There was one exception: the PANESS impersistence score. Both AD groups (which did not differ from one another) were significantly more impaired than both the non-autistic control groups (whose scores were essentially similar; DLD mean 0.86, 95% CI 0.68 to 1.04; HiAD mean 1.48, 95% CI 1.24 to 1.73; LoAD mean 1.37, 95% CI 1.17 to 1.57; LoIQ mean 0.88, 95% CI=0.57 to 1.57). This indicates that impersistence is more likely in, but is not limited to, autism.

The results of fine motor skills, sensory data, and oromotor function must be interpreted keeping in mind the large number of WD/CD scores, especially in the LoAD group. In fact, one of the most salient findings of this study was the differences between the groups in the prevalence of WD/CD scores. Bonferroni posthoc tests showed that the WD/CD scores significantly distinguished the LoAD group from the LoIQ group, with more LoAD than LoIQ children either unable or unwilling to complete tasks (DLD mean 0.02, 95% CI 0.01 to 0.04; HiAD mean 0.04, 95% CI 0.01 to 0.08; LoAD mean 0.45, 95% CI 0.38 to 0.52; LoIQ mean 0.19, 95% CI=0.11 to 0.26). In contrast, this score did not differentiate the HiAD group from the DLD group, because both essentially performed all tasks.

The percentage of children in each group who could not or refused to perform at least 50% of tasks is displayed in Figure 1. The 50% cutoff was chosen because it was the same as that used to determine whether a subdomain score was included in the analysis or considered missing (see Method, above). Note the extremely high percentage of children in the LoAD group, a proportion about threefold that in the LoIQ group.

Tendon stretch reflexes were abnormal in only 3.7% of the children (30 individual scores, 24 hyperreflexic, 6 absent) and were not lateralized. Two-thirds of the abnormal reflex scores were recorded in the LoAD children. There was no difference between the groups in mean laterality score when all the laterality measures were summed.

The neurologists' observations of motor abnormalities summarized in Table V largely confirm the quantitative assessments indicating that IQ had the strongest influence on the scores. This provides some validation of the quantitative test results, because the proportion of children with missing observational data was considerably smaller than for tasks requiring cooperation. There were occasional missing data points that were due to the examiner failing to score a child on that item.

Apraxia is widely considered to characterize children with autism (Bauman 1992), yet the neurologists' scores suggest that low IQ rather than autism *per se* may be its correlate. In fact, when comparing the two high IQ groups, the HiAD group had a lower mean apraxia score than the DLD group, although this difference did not reach statistical significance when tested with the Bonferroni test. When we considered the presence or absence of any abnormality of praxis (namely a score of 1 or 2 vs a score of 0), a direct comparison of the DLD and HiAD groups by χ^2 did show a significantly higher prevalence of apraxia in the DLD group than in the HiAD group ($\chi^2=5.825$, $p<0.05$). The assessment of praxis involved such a large number of items on the examination that more than 95% of the children in both the DLD and HiAD groups were scored as being apraxic in at least one of them.

Sterotypies distinguished the DLD and HiAD groups from the LoAD group (DLD mean 0.03, 95% CI 0.00 to 0.09; HiAD mean 0.29, 95% CI 0.07 to 0.50; LoAD mean 1.18, 95% CI 0.95 to 1.40; LoIQ mean 0.23, 95% CI=0.00 to 0.045). When any stereotypy score above zero was counted as abnormal (i.e. scores of 1 or 2 vs 0), χ^2 analysis indicated that the difference in prevalence between the four groups was highly significant ($\chi^2=83.373$, $p<0.0001$). The DLD group was the only one in which stereotypies were virtually absent (DLD 1.5%; HiAD 20%; LoAD 72%; LoIQ 13%).

The neurologists' overall clinical impressions of the children are summarized in Table VI. Unlike the performance-based measures just discussed, there were few missing data points, with scores assigned to 97 to 100% of children. Because missing data were inconsequential, no correction was deemed necessary.

In a separate analysis, the neurologists' impressions of sensorimotor and oromotor abnormalities, and apraxia for the DLD and HiAD groups were dichotomized as normal (score of 0) versus any abnormality (score of 1 or 2). This analysis showed that, on average, more children with DLD than HiAD had deficits in sensorimotor skills ($\chi^2=8.450$; $p<0.05$) and oromotor function ($\chi^2=5.285$; $p<0.05$), and there was a trend toward the DLD children being scored as more apraxic ($\chi^2=2.795$; $p=0.095$).

The statistical power provided by the current sample size was examined to determine what degree of confidence to use when interpreting the results. This analysis shows that each *t*-test posthoc comparison has the power to accurately capture what Cohen (1988) terms 'medium'-effect size differences between diagnostic groups (assuming conventional parameters of $\alpha=0.05$ and $1-\beta=0.80$). We, therefore, cannot rule out the possibility that smaller, but nonetheless valid, differences may exist between the groups. For this reason, 95% CIs are presented in the results for the more significant findings. Because few of the non-significant group differences for impersistence, stereotypies, and CD/WD involve scores

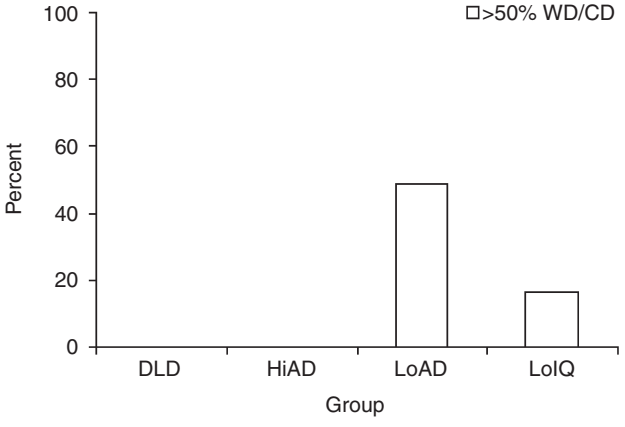

Figure 1: *Percentage 'wouldn't do/couldn't do' (WD/CD) by group (defined as more than 50% of tasks coded as 'WD/CD'). DLD, developmental language disorder; HiAD, high-IQ autistic disorder; LoAD, low-IQ autistic disorder; LoIQ, low IQ.*

that fall into a clinically meaningful range, using 95% CI, this supports the interpretation that non-significant results are not the result of Type II error.

Discussion

The purpose of the sensory/motor neurological examination is to detect focal, lateralized, or diffuse evidence for brain dysfunction, although it does not address deficits in the complex human abilities that differentiate one developmental disorder from another. As the children in the study were preselected not to have overt evidence of lateralized or diagnosable diffuse brain dysfunction as indexed by gross sensorimotor deficits or epilepsy, it is hardly surprising that NVIQ, as a surrogate for the severity of global brain involvement, was the main functional discriminator. One robust finding in this study was that the WD/CD scores sharply distinguished children with autism and low IQ not only from the non-retarded groups with NVIQs above 80, but also from non-autistic low IQ children. This is not to say that WD/CD scores were limited to the LoAD group, but clinicians would be well advised to consider an autistic spectrum disorder when confronted with a child unable or unwilling to perform tasks that are within the range of his physical and cognitive abilities, especially if that child also has difficulty in completing tasks; in other words, the child is impersistent. There are many putative, overlapping reasons for WD/CD scores in autism: severe mental retardation or the ubiquitous and persistent language comprehension deficits of many low-functioning children (Rapin and Dunn 2003), impaired perception of the usual signals of social rewards that encourage cooperation (Waterhouse et al. 1996), or severely impaired ability to pay attention jointly to a task introduced by someone else (Mundy et al. 1990). A longitudinal study by Stevens et al. (2000) of this same sample of children with autism suggests that a higher NVIQ had a moderating effect on the severity of their impairment of social interaction as they matured.

Stereotypy was the one observational score that distinguished the autistic from non-autistic groups. However, because the 95% CI for LoIQ stereotypies overlapped with scores seen in LoAD participants, no firm conclusion can be made with respect to the predictive value of this measure with non-autistic, low-IQ children. In the preschool study, the percentage of children in whom neurologists reported stereotypies was very similar to the school-age observations (prevalence of stereotypies in preschool children: DLD 2%, HiAD 41%, LoAD 65%, LoIQ 13%; Rapin 1996a). Others, too, have observed an increased prevalence of stereotypies in autistic children compared with children with mental retardation (Bodfish et al. 2000). Brasic (1999) proposed that stereotypies may serve as a fairly specific marker for autism, even though they are often seen in normal infants and may occasionally persist in children who are neither autistic nor mentally retarded. Tan et al. (1997) studied 10 children with stereotypies who had a history of mild to moderately delayed developmental milestones; five were hyperactive or had attention problems but none had autism or mental retardation. Only two of these 10 children stopped having stereotypies spontaneously; i.e. without medication. McCreary and Handley (1999) found that in rodents, blockage of dopamine receptors did not have an effect on a tic-like behavior (blinking induced by thyrotropin-releasing hormone), whereas antagonists of dopamine D_1 but not D_2 receptors attenuated forepaw licking induced by thyrotropin-releasing hormone, a repetitive behavior more akin to stereotypy. This apparent dissociation between tics and stereotypies and the evidence that at least some stereotypies are mediated by dopaminergic mechanisms implicate the role of subcortical striatal networks in the many stereotypies of autism.

The neurologists' summary clinical impressions indicated better sensorimotor and oromotor skills, and praxis in the HiAD children than in the DLD children. The subdomain test scores for apraxia were also better for the HiAD children than for the DLD children, whereas quantitative test scores for sensorimotor and oromotor tasks were comparable for the two groups (with the exception of the Purdue Pegboard, on which the DLD group performed better than the HiAD group). For all groups, scores for the final clinical impression of apraxia were higher (worse) than the observational apraxia scores. It seems that the neurologists' overall interaction with the child, perhaps supplemented by the histories obtained from the children's parent at the time of the school-age assessment, might have influenced their final clinical impression of apraxia beyond their observation score.

Studies of the KE family, with a genetically determined severe mixed language disorder, revealed a marked oromotor dyspraxia (Hurst et al. 1990) together with profound impairment of linguistic and grammatical skills (Watkins et al. 2002a). Affected individuals in this family have a point mutation in the *FOXP2* gene (Lai et al. 2001) on chromosome 7q31 (Fisher et al. 1998). It has been proposed that the verbal and nonverbal deficits in this family 'arise from a common impairment in the ability to sequence movement or in procedural learning' (Watkins et al. 2002a). Imaging studies suggest that these impairments involve abnormalities in the basal ganglia (Watkins et al. 2002b). In our study, the test

Table VI: Neurologist impression scores

F. Neurologist Impression	*DLD (n=89)*		*HiAD (n=42)*		*LoAD (n=74)*		*LoIQ (n=37)*		*F*	*Subgroup differences*
	Mean (SD)	*%n*	*Mean (SD)*	*%n*	*Mean (SD)*	*%n*	*Mean (SD)*	*%n*		
Sensorimotor	0.61 (0.53)	99	0.38 (0.49)	100	0.63 (0.61)	99	0.92 (0.73)	97	$F_{3,235}=5.42$	e
Oromotor	0.43 (0.58)	99	0.24 (0.43)	100	0.44 (0.63)	97	0.72 (0.66)	97	$F_{3,234}=4.45$	e
Apraxia	0.44 (0.54)	99	0.21 (0.42)	100	0.74 (0.67)	99	0.89 (0.75)	97	$F_{3,235}=11.58$	a, b, d, e

Scores: 0, normal; 1, mildly abnormal; 2, unequivocally abnormal. a, DLD<LoIQ; b, DLD<LoAD; c, DLD<HiAD; d, HiAD<LoAD; e, HiAD<LoIQ; f, LoAD<DLD; g, LoAD<LoIQ; h, LoIQ<HiAD; i, LoIQ<LoAD. *%n*, percentage of children with data in each subdomain. DLD, developmental language disorder; HiAD, high-IQ autistic disorder; LoAD, low-IQ autistic disorder; LoIQ, low IQ without autism.

scores indicating worse apraxia and the neurologists' impressions of worse sensorimotor, oromotor, and apraxic skills in the DLD group than in the HiAD group are compatible with the findings in the KE family.

Thus, the standard sensorimotor neurological examination has a limited utility for distinguishing the sensorimotor function of autism from that of other neurodevelopmental disorders, even when bolstered with the PANESS to provide a more quantitative assessment of soft signs. The two possible exceptions are stereotypies and impersistence. In children with impaired language without mental retardation, stereotypies should raise the suspicion of autism. The strongest finding of this study is that inability/unwillingness to perform tasks in a child with a low IQ is highly suggestive of a diagnosis of autism.

DOI: 10.1017/S0012162206000089

Accepted for publication 24th February 2005.

Acknowledgements
We acknowledge the support of the National Alliance for Autism Research (DEM) and NINDS grant 20489 (IR). We greatly appreciate willingness of the children and their parents to participate in this study. We acknowledge the work and many other contributions of the late Dr Doris A Allen and the collaboration of Dr Dorothy M Aram, Dr Michelle Dunn, Dr Deborah Fein, Dr Nancy Hall, Dr Robin Morris, and Dr Lynn Waterhouse in designing the original project and sharing behavioral data, and of additional neurological examiners Dr Susan Klein, Dr Karen Balaban-Gil, and Dr David Urion. We thank Danielle Drezek and many research assistants for their assistance in behavioral testing and data analysis.

References

American Psychiatric Association. (2000) *Diagnostic and Statistical Manual of Mental Disorders. 4th edn,* text revision: *DSM IV-TR.* Washington, DC: American Psychiatric Association.

Bodfish JW, Symons FJ, Parker DE, Lewis MH. (2000) Varieties of repetitive behavior in autism: comparisons to mental retardation. *J Autism Dev Disord* **30:** 237–243.

Bauman ML. (1992) Motor dysfunction in autism. In: Joseph A, Young RR, editors. *Movement Disorders in Neurology and Neuropsychiatry.* Boston: Blackwell. p 660–663.

Brasic JR. (1999) Movements in autistic disorder. *Med Hypoth* **53:** 48–49.

Cohen J. (1988) *Statistical Power Analysis for the Behavioral Sciences.* 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates.

Denckla MB. (1977) Minimal brain dysfunction and dyslexia: beyond diagnosis by exclusion. In: Blaw ME, Rapin I, Kinsbourne M, editors. *Topics in Child Neurology.* New York: Spectrum. p 243–261.

Denckla MB. (1985) Revised neurological examination for subtle signs. *Psychopharm Bull* **21:** 773–800.

Denhoff E, Siqueland ML, Komich MP, Hainsworth PK. (1968) Developmental and predictive characteristics of items from the Meeting Street School Screening Test. *Dev Med Child Neurol* **10:** 220–232.

Fisher SE, Bodfish JW, Symons FJ, Parker DE, Lewis MH. (1998) Localisation of a gene implicated in a severe speech and language disorder. *Nat Genet* **18:** 168–170.

Fletcher SG. (1972) Time-by-count measurement of diadochokinetic syllable rate. *J Speech Hear Res* **15:** 763–770.

Gardner RA, Broman M. (1979) The Purdue Pegboard: Normative data on 1334 school children. *J Clin Child Psychol* **8:** 156–162.

Haynes C, Naidoo S. (1991) *Children with Specific Speech and Language Impairment. Clinics in Developmental Medicine No. 119.* London: Mac Keith Press.

Hresko W, Reid D, Hammill D. (1981) *The Test of Early Language Development.* Austin, TX: Pro-ED.

Hurst JA, Baraitser M, Auger E, Graham F, Norell S. (1990) An extended family with a dominantly inherited speech disorder. *Dev Med Child Neurol* **32:** 352–355.

Kinsbourne M, Warrington EK. (1964) The development of finger differentiation. *Q J Exp Psychol* **15:** 132–137.

Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. (2001) A novel forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **413:** 519–523.

McCreary AC, Handley SL. (1999) The thyrotropin-releasing hormone analogue MK771 induces tic-like behaviors: the effects of dopamine D1 and D2 receptor antagonists. *Eur J Pharmacol* **369:** 1–9.

Miller JF, Chapman RS, MacKenzie H. (1981) Individual differences in the language acquisition patterns of mentally retarded children. In *Proceedings of the Symposium on Research in Child Language Disorders* **2:**130–146.

Mundy P, Sigman M, Kasari C. (1990) A longitudinal study of joint attention and language development in autistic children. *J Autism Dev Disord* **20:** 115–128.

Peters JE, Romine JS, Dykman RA. (1975) A special neurological examination of children with learning disabilities. *Dev Med Child Neurol* **17:** 63–78.

Pollack M, Goldfarb W. (1957) The face-hand test in schizophrenic children. *Arch. Neurol Psychiatry* **77:** 635–642.

Rapin I, editor. (1996a) *Preschool Children with Inadequate Communication Clinics in Developmental Medicine No. 139.* London: Mac Keith Press.

Rapin I, editor. (1996b) *Preschool Children with Inadequate Communication. Clinics in Developmental Medicine No. 139.* London: McKeith Press. p 98–122.

Rapin I, Dunn M. (2003) Update on the language disorders of individuals on the autistic spectrum. *Brain Dev* **25:** 166–172.

Stevens MC, Fein D, Dunn M, Allen D, Waterhouse LH, Feinstein C, Rapin I. (2000) Subgroups of children with autism by cluster analysis. A longitudinal examination. *J Am Acad Child Adolesc Psychiatry* **39:** 399–407.

Tan A, Salgado M, Fahn S. (1997) The characterization and outcome of stereotypical movements in nonautistic children. *Mov Disord* **12:** 47–52.

Touwen BCL, Prechtl HFR. (1970) *The Neurological Examination of the Child with Minor Nervous Dysfunction. Clinics in Developmental Medicine No. 38.* London: Spastics International Medical Publication.

Waterhouse L, Fein D, Modahl C. (1996) Neurofunctional mechanisms in autism. *Psychol Rev* **103:** 457–489.

Watkins KE, Dronkers NF, Vargha-Khadem F. (2002a) Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain* **125:** 452–464.

Watkins KE, Vargha-Khadem F, Ashburner J, Passingham RE, Connelly A, Friston KJ, Frackowiak RS, Mishkin M, Gadian DG. (2002b) MRI analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain* **125:** 465–478.

Wing L. (1996) Wing Autistic Disorder Interview Checklist (WADIC). In: Rapin I, editor. *Preschool Children with Inadequate Communication. Clinics in Developmental Medicine No. 139.* London: Mac Keith Press. p 247–251.

World Health Organization. (1992) *The ICD-10 Classification of Mental and Behavioural Disorders. Clinical Descriptions and Diagnostic Guidelines.* Geneva: World Health Organization.

List of abbreviations

AD	Autistic disorder
DLD	Developmental language disorder
HiAD	Autism with high IQ
LoAD	Autism with low IQ
LoIQ	Low IQ without autism
NVIQ	Nonverbal IQ
PANESS	Physical and Neurological Examination for Soft Signs
WD/CD	'wouldn't do/couldn't do