Wisconsin Card Sorting Test and Halstead Category Test Performances of Children and Adolescents who Exhibit the Syndrome of Nonverbal Learning Disabilities*

Nancy J. Fisher\(^1,2\), John W. DeLuca\(^2\), and Byron P. Rourke\(^1,3\)

\(^1\)University of Windsor, \(^2\)Wayne State University School of Medicine, Detroit, MI, and \(^3\)Yale University School of Medicine

**ABSTRACT**

Wisconsin Card Sorting Test (WCST) and Halstead Category Test (HCT) performances of 15 children/adolescents (age range 9 to 17 years) diagnosed with the syndrome of nonverbal learning disabilities (NLD) and 15 age-, gender-, and FSIQ-matched verbal learning-disabled (VLD) controls were examined. The VLD group made significantly fewer errors on the HCT than did the NLD group. In addition, the VLD controls significantly outperformed the NLD participants on seven of the nine WCST index scores (all except failure to maintain set and learning to learn scores). Consistent with the adult literature, the common variance between the HCT and WCST was modest, indicating that these two neuropsychological measures gauge different domains of functioning. It is suggested that the HCT assesses conceptualization/higher order reasoning abilities, whereas the WCST taps dimensions of executive functioning.

Children and adolescents with the syndrome of nonverbal learning disabilities (NLD) display a pattern of neuropsychological assets and deficits characterized by well developed rote verbal skills within the context of poor psychomotor, tactile-perceptual, visual-spatial-organizational, and problem-solving abilities (Harnadek & Rourke, 1994; Rourke, 1989, 1995). Many psychometric studies of those with the syndrome have been conducted (see Rourke 1989, 1995, for reviews). Past studies investigating the performance of individuals with the syndrome of NLD on the Halstead Category Test (HCT; Reitan & Davison, 1974), a problem-solving measure thought to involve higher order deductive reasoning (e.g., rule learning) skills (Reitan & Wolfson, 1992, 1993), have reported impaired performance among those with this subtype of learning disability (e.g., Strang & Rourke, 1983). However, a neglected area of study is the performance of those with NLD on problem-solving measures of so-called “executive” functioning (EF); that is, EF as distinct from higher order reasoning (HOR) (Dencsla, 1996a). As a global domain of neuropsychological functioning, EF is thought to include components such as the following: task adherence, planning, response inhibition, working memory, and cognitive flexibility (DeLuca, Butkus, & Margolis, 1995; DeLuca, Henderson, & Tojek, in press; Dencsla, 1996b; Levin et al., 1991; Taylor, Schatschneider, Petrill, Barry, & Owens, 1996; Welsh, Pennington, & Groisser, 1991).

In the current investigation, we examined the performances of children and adolescents with NLD on the Wisconsin Card Sorting Test (WCST; Heaton, Chelune, Talley, Kay, & Curtiss, 1993), a test recognized as the most widely utilized neuropsychological EF measure (Butler, Retzlaff, & Vanderploeg, 1991), in addition to examining their performance on the HCT (Reitan & Davison, 1974; Reitan & Wolf-
son, 1992, 1993). In doing so, our goals were two-fold. First, we sought to determine whether children and adolescents with NLD demonstrate impairment on the WCST and HCT. Second, we wished to delineate the amount of shared variance between the HCT and the WCST, as one approach to investigating the validity of the notion that EF encompasses a domain of functioning distinct from higher order deductive reasoning. Before stating our hypotheses, it is instructive to review the theoretical models and past research that gave rise to our predictions.


Rourke (1989) applied the Goldberg-Costa (1981) model of hemispheric specialization in his theory of NLD as a manifestation of white matter dysfunction, which is hypothesized to eventuate in compromise of right hemisphere systems moreso than systems within the left hemisphere. The Goldberg-Costa (1981) model has its basis in observations of neureanatomy organizational differences between the right and left cerebral hemispheres. Most important, Goldberg and Costa (1981) reviewed evidence documenting the diffuse organization of the right hemisphere, in contradistinction to the more focal organization of the left hemisphere. In this regard, Goldberg and Costa (1981) noted that the right hemisphere has a lower grey-to-white-matter ratio than does the left, (i.e., implying more long myelinated fibres in the right hemisphere) and has a greater representation of the association areas. On the other hand, they noted that the left hemisphere is more focally organized, with distinct modality-specific cortical areas (i.e., the opercula) appearing more prominently. Goldberg and Costa go on to argue that the structural organization and patterns of connectivity of the right hemisphere would seem to make it more efficient at interregional integration, whereas that of the left hemisphere implies specialization for intraregional activity.

On the basis of these structural distinctions between the hemispheres, in addition to the unilateral brain-damage literature, Goldberg and Costa (1981) argued that systems within the right hemisphere have a greater capacity to deal with complex information (i.e., due to greater representation of association areas in this hemisphere) and also have the capability to process several modes of representation within a single cognitive task (i.e., due to interregional connectivity). On the other hand, it was theorized that the left hemisphere would function more effectively on tasks requiring the utilization of a single mode of processing because of its greater representation of unimodal sensory and motor areas and its characteristic pattern of intramodal connectivity.

Taking this one step further, Goldberg and Costa (1981) theorized that right hemisphere systems are specialized for exploratory processing of novel situations, wherein pre-existing strategies or set solutions are not available to the individual. Left hemisphere systems, on the other hand, are hypothesized to be more efficient at processing that involves storage and stereotypic application of cognitive strategies/routines already formulated, available, and familiar. Thus, it was theorized that whereas systems within the right hemisphere are responsible for assembling new behavioural rules/routines (i.e., codes), systems within the left hemisphere store the codes once established, and are specialized to apply them in a routinized manner when the need arises in future similar situations.

In accordance with the above, Rourke (1982, 1989) hypothesized that the syndrome of NLD may be a manifestation of white matter dysfunction, which would (i.e., with application of the Goldberg-Costa model) compromise the functioning of right hemisphere systems moreso than left hemisphere systems. This line of thought leads one to predict that individuals with NLD (i.e., suspected white matter dysfunction) would not fare well on measures such as the WCST, which appears to require strategy generation and set-switching within a novel, unstructured setting. Indeed, from the Goldberg and Costa (1981) theoretical stance, right hemisphere systems would be better able to deal with novelty and ambiguity, especially when no explicit instructions or cognitive framing are provided. Thus, on the WCST, one would expect individu-
als with NLD to produce a large number of stereotypic perseverative responses, yet demonstrate minimal difficulty in maintaining a correct pattern of responding, once obtained. In short, such individuals would be predicted to have difficulty switching their approach in response to changing task demands.

The Distinction Between HOR and EF: HCT versus WCST
Past studies employing adult neurological, psychiatric, and normal samples have suggested minimal overlap between the HCT and WCST in terms of shared variance (i.e., 12 to 34%) (Crockett, Bilsker, Hurwitz, & Kozak, 1986; Donders & Kirsch, 1991; King & Snow, 1981; Pendleton & Heaton, 1982; Perrine, 1993). The second goal of this investigation involved a determination of whether the general finding of low shared variance between the HCT and WCST among adults could be extended to children and adolescents with learning disabilities. Furthermore, past research (e.g., Perrine, 1993) has suggested that the HCT is more a measure of rule learning/conceptualization whereas the WCST gauges attribute identification (Perrine, 1993), is less conceptually difficult, and is more sensitive to perseverative tendencies (Donders & Kirsch, 1991; Pendleton & Heaton, 1982; Perrine, 1993). The WCST may be more sensitive to perseverative tendencies because, in the administration of this measure, the examinee is not explicitly informed that the categorizing rule may change, as he/she is repeatedly reminded during the HCT (Donders & Kirsch, 1991; King & Snow, 1981). As such, the WCST may be more likely to reflect internally driven control/switching (i.e., “executive”) functions. On the other hand, the HCT may not reflect some aspects of internally initiated mechanisms, because administration of this measure requires the examiner to interact with the test taker, “utilizing prompts to facilitate the patient’s exploration of his/her abilities in organization and problem-solving” (Finlayson, Sullivan, & Alfano, 1986, p. 709).

Statement of Hypotheses
In this investigation, we sought to determine whether children and adolescents with the syndrome of NLD exhibit impaired performance on the HCT and WCST. In terms of HCT performance, we predicted that our findings would replicate those of Strang and Rourke (1983), demonstrating impairment on this measure for an NLD group but not for a control group of children and adolescents with verbal learning disabilities (VLD). In particular, we expected the VLD group to perform in a significantly superior manner on subtests 4, 5, and 6, (i.e., the most complex HCT subtests in terms of conceptual difficulty and incidental memory) but not on subtests 1, 2, and 3, as was found by Strang and Rourke (1983).

Similarly, on the basis of the Rourke white matter model, we predicted that our NLD sample would perform poorly on the WCST as compared to VLD controls. Finally, in accord with past research, we expected that the shared variance between the HCT and WCST would be negligible, reflecting the different underlying aspects of cognitive functioning that are inferred to be measured by each.

METHOD

Subjects
Archival WCST and HCT data from 15 children/adolescents (age range 9 to 17 years) with NLD and 15 children/adolescents (age range 9 to 17 years) with verbal learning disabilities (VLD; i.e., controls) were studied in this investigation. The WCST was administered in the standard fashion as outlined in Heaton et al. (1993). Those aged 9 through 14 were administered the older children’s version of the HCT (Reitan & Wolfson, 1992), while those aged 15 through 17 were administered the adult HCT (Reitan & Wolfson, 1993). All subjects were diagnosed with NLD or VLD by the same licensed clinical neuropsychologist (J.W.D.), between the years 1987 and 1996, after having been administered an extended Halstead-Reitan neuropsychological test battery. All had been referred to one of three metropolitan outpatient clinics for evaluation regarding academic, behavioral, and/or socioemotional difficulties. Eighty percent of the subjects (24/30) were right-handed. There were 6 girls and 9 boys in each of the NLD and VLD groups.
The NLD participants met the majority of neuropsychological test performance criteria for classification of potential NLD subjects outlined by Harnadek and Rourke (1994) (see Table). All members of the NLD sample displayed a Wexscher Intelligence Scale for Children (WISC-R or WISC-III; Wechsler, 1974, 1991) VIQ > PIQ pattern; the mean VIQ > PIQ difference was 15.53 (SD = 8.66). As well, all participants in the NLD group demonstrated inferior performances on academic achievement tests of arithmetic (i.e., Wide Range Achievement Test - 2nd or 3rd version; Jastak & Wilkinson, 1984; Wilkinson, 1993) as compared to their performances on achievement tests of spelling and reading; the mean difference score between the standard scores for spelling + reading and that for arithmetic (i.e., [(S+R) - A] / 2) was 25.3 (SD = 10.99).

The VLD controls were diagnosed with one of the following: phonological processing disorders, verbal expressive disorder, developmental speech and language disorder, dyslexia, or mixed developmental disorder involving speech and/or language. The VLD subjects were selected to match the NLD participants in terms of gender, age, and FSIQ. There were no significant differences between the NLD and VLD groups in terms of Full Scale WISC-R/WISC-III IQ (NLD mean FSIQ = 81.67 [SD = 8.3]; VLD mean FSIQ = 86.53 [SD = 7.97], p = .113) or age (mean age of NLD group = 13.93 years [SD = 2.34]; mean age of VLD controls = 13.17 years [SD = 2.3], p = .373). The majority of VLD controls (14/15) displayed a WISC-R/WISC-III PIQ > VIQ pattern; the mean PIQ minus VIQ index was 16.53 (SD = 11.67).

In determining levels of performance on the WCST, normative data provided by Heaton et al. (1993) were utilized for the following variables: categories completed; failure to maintain set; learning to learn; trials to first category. For the remaining WCST score variables, normative data were provided by the WCST computer scoring system (Heaton, Curtiss, & Tuttle, 1993). Impaired performance was indicated by T scores ≤ 40 (M = 50; SD = 10; higher scores indicating better performance) for all variables. Normative data provided by Knights and Norwood (1980) were utilized in determining levels of performance on the HCT for participants up to 14 years of age; for those 15 and older, normative data provided by Fromm-Auch and Yeudall (1983) were utilized.

**Measures**

**HCT**

In administration of the HCT, the examinee is seated in front of a screen onto which stimulus figures (i.e., shapes, letters, designs) are presented. Below the screen lies a response board comprised of buttons labelled with the numbers 1, 2, 3, and 4. Examinees are instructed to observe the stimulus on the screen, decide whether it suggests the number 1, 2, 3, or 4, and then pull the corresponding response key. If the examinee answers correctly, a bell sound ensues; if he/she delivers an incorrect response, a harsh buzz is emitted. Only one response is permitted per item. The HCT contains several subtests; the examinee is informed that once a rule is deduced, continued application of this principle should produce correct responses throughout the remainder of the particular subtest, despite changes in features of the stimuli (e.g., from shapes to letters). They are also informed that across one subtest to another, the response rule may stay the same or may change. The examiner is never informed of the principle for any subtest (Reitan & Wolfson, 1992, 1993; see these texts for more complete descriptions).

**WCST**

The WCST requires the examinee to sort up to 128 response cards displaying one to four blue, yellow, red, or green shapes [either cross(es), circle(s), triangle(s), star(s)], to four unchanging “key” (i.e., stimulus) cards. The four key cards depict the following: (1) one red triangle; (2) two green stars; (3) three yellow crosses; (4) four blue circles. Examinees are instructed to take one response card at a time and place it below the key card that they think it matches. As the response cards can be matched to the key cards in terms of colour, shape, or number, the test-taker must figure out the sorting principle currently in effect by trial and error, modulating responses to evaluative feedback. After each card is sorted, the examinee is informed as to the correctness of their sort.

The to-be-derived sorting principles are colour, form (i.e., shape), and number — in that order. After 10 consecutive correct sorts, the criterion sorting principle is changed without warning, requiring the subject to shift cognitive set in deducing a new principle. When each of the three sorting principles has been derived and maintained to criterion (i.e., 10 consecutive correct sorts are
Table. Criteria for Potential NLD Subjects (Harnadek & Rourke, 1994) Met by the Current Sample.

<table>
<thead>
<tr>
<th>Harnadek &amp; Rourke (1994) Criteria for Potential NLD Subjects</th>
<th>Subject Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>WISC VIQ &gt; 79</td>
<td>1  2  3  4  5  6  7  8  9  10  11  12  13  14  15</td>
</tr>
<tr>
<td>SSPT* or Auditory Closure&lt; 1 SD below mean</td>
<td>+  +  +  -2  +  +  +  +  +  +  +  +  +  +  +  +</td>
</tr>
<tr>
<td>WRAT Reading &amp; Spelling &gt; Arithmetic by 10 or more SS</td>
<td>+  +  +  +  -1  +  +  +  +  +  +  +  +  +  +  +</td>
</tr>
<tr>
<td>Target Tests &gt; 1 SD below mean</td>
<td>N/A + N/A N/A N/A N/A N/A N/A + + + N/A + N/A</td>
</tr>
<tr>
<td>VIQ &gt; PIQ by 10 or more points</td>
<td>-1  +  +  +  +  +  +  +  +  +  +  +  +  +  +  +</td>
</tr>
<tr>
<td>Grooved Pegboard Test* (either hand) &gt; 1 SD below mean</td>
<td>+  +  +  +  +  +  +  +  +  +  +  +  +  +  +  +</td>
</tr>
<tr>
<td>Dysgraphesohia, finger agnosia or astereognosia &gt; 1 SD below mean</td>
<td>+  +  +  +  +  +  +  +  +  +  +  +  +  +  +  +</td>
</tr>
<tr>
<td>Percent of available criteria met</td>
<td>83  100 100 67 83 83 67 100 100 100 100 100 100 83</td>
</tr>
</tbody>
</table>

Note. N/A = Not Administered; NLD = Nonverbal Learning Disabled; PIQ = Performance Intelligence Quotient; SD = Standard Deviation; SS = Scaled Score; SSPT = Speech Sounds Perception Test; VIQ = Verbal Intelligence Quotient; WISC = Wechsler Intelligence Scale for Children; WRAT = Wide Range Achievement Test; *meets criterion; -does not meet criterion; 1VIQ > PIQ by 9 points; 2VIQ = 78; 3Reading = 105, Spelling = 129, Arithmetic = 96; 4Reitan & Davison, 1974; 5Kass, 1964; 6Klove, 1963.
made to each), they are repeated in the same order. The examinee is never informed of the sorting principles, their order of presentation, or the criterion number of consecutively correct responses required for each rule before a new rule is implemented. The test is terminated when all six categorization rules have been derived and maintained to criterion or when all 128 cards have been sorted (Heaton et al., 1993).

The following nine WCST score variables were examined in this study: (1) categories completed; (2) trials to first category; (3) total errors; (4) perseverative responses; (5) perseverative errors; (6) nonperseverative errors; (7) conceptual level responses; (8) failure to maintain set; and (9) learning to learn (see Heaton et al., 1993, for descriptions of each).

Overview of data analysis
The first goal of the current investigation was to determine whether or not children and adolescents with NLD (as distinct from those with VLD) demonstrate impairment on the HCT and WCST. We predicted that those with NLD would perform in an inferior fashion relative to those diagnosed with VLD, because the HCT and WCST would appear to be more reliant upon the integrity of systems within the right hemisphere. To evaluate these predictions, the mean HCT total errors score and the mean WCST index scores from the NLD and VLD groups were submitted to between-group t tests for independent samples (Norusis, 1993). For the subjects aged 9–14, we also conducted t tests between the two learning-disabled groups for each of the six HCT subtest scores. In accord with the findings of Strang and Rourke (1983) we predicted that for this reduced sample, the VLD group would make significantly fewer errors as compared to the NLD group on subtests 4, 5, and 6, but not on the first three HCT subtests.

The second goal of this study was to determine whether the common variance shared by the HCT and the WCST among our sample of children and adolescents with learning disabilities is comparable to that reported in studies of adults. On the basis of these past studies with adults (e.g., Donders & Kirsch, 1991; Perrine, 1993), we predicted only a modest amount of shared variance between the two measures. To evaluate this prediction, correlations were computed between HCT total errors scores and scores on the nine WCST variables examined in this investigation. Squaring the resulting coefficients allowed us to determine the proportion of shared variance between the two tests.

RESULTS
In accordance with the first goal of this study, we compared the HCT and WCST performances of the NLD and VLD groups via t tests for independent samples. The mean number of HCT total errors for the VLD group (M = 48.8) was significantly lower than that for the NLD sample (M = 76.4) [(t)_{28} = 4.24, p < .0001]. Whereas 14 of the 15 NLD subjects were impaired in terms of HCT total errors scores, only 5 of the 15 VLD controls were impaired (see Figure 1). Similarly, the VLD group outperformed the NLD sample on all nine WCST scores (see Figures 1 and 2). These differences were significant at the .05 level for all but the failure to maintain set and learning to learn scores.

When we reduced the NLD and VLD samples to include only those aged 9–14 (i.e., those administered the older children's version of the HCT), and conducted t tests between the group mean scores on each particular HCT subtest, significant differences in favor of the VLD group obtained for subtests four [(t)_{17} = 2.54, p < .05], five [(t)_{17} = 3.26, p < .05], and six [(t)_{17} = 2.23, p < .05]. While the VLD group also outperformed the NLD sample on subtests 1, 2, and 3, these differences were not statistically significant. This pattern of HCT subtest findings is consistent with those of Strang and Rourke (1983) (see Figure 3).

With respect to the second goal of this investigation, Pearson product-moment correlations computed between the HCT total errors score and raw scores on the nine WCST indices for the combined sample (i.e., NLD and VLD) were significant for the categories completed [(r)_{28} = -.452, p < .05], percent conceptual level [(r)_{28} = -.444, p < .05], learning to learn [(r)_{28} = -.374, p < .05], perseverative errors [(r)_{28} = .408, p < .05], perseverative responses [(r)_{28} = .385, p < .05], trials to first category [(r)_{28} = .415, p < .05], and total errors [(r)_{28} = .436, p < .05] variables. These correlations on average suggest 17.4% shared variance. Correlations between HCT errors and the failure to maintain set and nonperseverative errors scores were not significant. With respect to the relationship between total errors on both the WCST and HCT,
Fig. 1. Percentage of NLD and VLD subjects impaired on the nine WCST variables and the HCT. CATERR = Number of errors on the HCT; CC = Categories Completed; CONLEV = Percent Conceptual Level Responses; FTMS = Failure to Maintain Set; LL = Learning to Learn; NPERR = Nonperseverative Errors; NLD = Nonverbal Learning Disabled; PERR = Perseverative Errors; PRS = Perseverative Responses; TTFC = Trials to First Category; TOTERR = Total Errors on WCST; VLD = Verbal Learning Disabled.

Fig. 2. Mean T score profiles on the nine WCST variables. CC = Categories Completed; CONLEV = Percent Conceptual Level Responses; FTMS = Failure to Maintain Set; LL = Learning to Learn; NPERR = Nonperseverative Errors; NLD = Nonverbal Learning Disabled; PERR = Perseverative Errors; PRS = Perseverative Responses; TTFC = Trials to First Category; TOTERR = Total Errors on WCST; VLD = Verbal Learning Disabled; WCST = Wisconsin Card Sorting Test.
it appears that the tests share approximately 19% common variance.

DISCUSSION

Individuals with NLD display a well-documented pattern of neuropsychological assets and deficits involving adept rote verbal skills, relatively impaired visual-spatial-organizational and problem-solving skills, in addition to other cognitive difficulties. No prior studies have investigated the performance of NLD individuals on measures of EF as distinct from HOR. In this investigation, we examined the performance of children and adolescents with NLD on both the HCT and WCST, and the degree of shared variance between WCST and HCT. In doing so, we indirectly evaluated Rourke's (1982, 1989) application of the Goldberg-Costa model in his theory of NLD.

The results of this study demonstrated that children and adolescents with the syndrome of NLD exhibit impairment on both the HCT and WCST. Our findings converge with those of Strang and Rourke (1983), who reported a T score profile hovering around 40 for their NLD sample on the HCT subtests, and reported superior performance of those with VLD over those with NLD on the last three subtests of the HCT (for ages 9–14) yet no significant differences between the groups in terms of performance on the first three subtests of this measure. Thus, we replicated the Strang and Rourke (1983) findings of impairment on the HCT by individuals with NLD. Moreover, our demonstration of inferior WCST performance by the NLD sample extends the domain of problem-solving performance impairment among those with this type of learning disability to include EF. These extended findings support Rourke’s (1982, 1989, 1995) model of NLD as a syndrome involving deficiencies in the capacity to deal with novelty, ambiguity, and nonroutinized situations.

The correlation between HCT total errors and WCST total errors indicates that there is less than one-fifth shared variance between the two error summary scores. Other WCST variables
correlating significantly with HCT total errors had squared correlation coefficients suggesting between 14 and 20% shared variance. These data are generally consistent with those reported in the adult literature (Donders & Kirsch, 1991; King & Snow, 1981; Pendleton & Heaton, 1982; Perrine, 1993) and extend the finding to younger age-groups. As with adults, it is clear that the two tests cannot be used interchangeably for clinical or other purposes. It appears that, as the adult literature has demonstrated, although the HCT and WCST are related, these tests measure different abilities in children and adolescents. Although at first glance the HCT and WCST may appear to be similar tests, we agree with others that the HCT seems to be a measure mostly of rule deduction/concept formation, and suggest that the WCST would appear to gauge executive-type functioning. Clearly, factor analytic research is needed to test these suggested differences between the HCT and WCST. Such studies are currently underway in our laboratories.

REFERENCES


