Screening for neuropsychological impairment in children using Reitan and Wolfson’s preliminary neuropsychological test battery

Jana L. Vanderslice-Barr a, Julie K. Lynch b,*, Robert J. McCaffrey a, b

a University at Albany, State University of New York, Albany, NY, United States
b Albany Neuropsychological Associates, Albany, NY, United States

Accepted 3 January 2008

Abstract

This study examined the efficacy of two screening batteries developed by Reitan and Wolfson (Reitan, R. M., & Wolfson, D. (2006). Issues in neuropsychology: Where have we succeeded? Where have we failed? What are the clinical and forensic implications? Presented at the 26th annual conference of the National Academy of Neuropsychology) for predicting neuropsychological impairment on the Halstead–Reitan Neuropsychological Test Battery for Older Children (ages 9–14) and the Reitan–Indiana Neuropsychological Test Battery (ages 5–8). Sixty-eight subjects were selected from archival data collected from cases in a private practice setting. Using Reitan and Wolfson’s recommended screening battery neuropsychological deficit score (SBNDS) cut-off of 16/17 correctly classified 82.5% of older children as normal or impaired according to total GNDS scores. The screening battery for young children (n = 28) resulted in 100% correct classification using Reitan and Wolfson’s recommended SBNDS cut-off of 22/23.

Published by Elsevier Ltd on behalf of National Academy of Neuropsychology.

Keywords: Halstead–Reitan Neuropsychological Test Battery for Older Children; Reitan–Indiana Neuropsychological Test Battery; Neuropsychological screening battery

The development of valid neuropsychological screening batteries for children has been an ongoing goal for clinical researchers (Donders, 1992; Horton, 1995, 1996, 1998; Reitan & Herring, 1985; Reitan & Wolfson, 2004a, 2004b; Slater & VanWagoner, 1988). Comprehensive neuropsychological testing requires numerous hours of administration time. Utilization of a short-length screening battery would enable clinicians to distinguish between children who may or may not require further testing, instead of conducting full-length neuropsychological assessments on every child where cognitive and/or behavioral difficulties are reported.

The Halstead–Reitan Neuropsychological Test Battery for Older Children (HRNB-OC, ages 9–14) and the Reitan–Indiana Neuropsychological Test Battery for Younger Children (RINTB, ages 5–8) are two commonly used neuropsychological test batteries for children (Nussbaum & Bigler, 1997). Reitan and Wolfson have developed a screening battery for the purpose of identifying children who have a greater likelihood of evidencing neuropsychological impairment on comprehensive neuropsychological testing with the Reitan batteries. Two screening batteries have been developed; one for younger children that is derived from the RINTB, and another for older children that
is derived from the HRNB-OC. In their initial validation studies of these screening instruments, Reitan and Wolfson (2006) have found the younger children screening battery to have a sensitivity of 85% and specificity of 85%; the screening battery for older children was found to have a sensitivity of 83% and specificity of 91%. Since these are recently introduced screening batteries, additional study of the predictive validity of the measures with other samples is important. The predictive validity of the screening batteries was examined in a sample of younger and older children referred for neuropsychological evaluation because brain damage was suspected, although not definitively confirmed by medical procedures as was the case with the Reitan and Wolfson validation samples. In routine clinical practice, it is very typical to conduct a neuropsychological evaluation in the absence of medically confirmed brain damage, and therefore, our sample is representative of the type of referrals that clinical neuropsychologists commonly encounter. Information regarding the utility of these screening measures with our sample will therefore provide useful information relevant to typical clinical practice.

1. Study 1: Older Children

1.1. Method

1.1.1. Participants
Approval for analysis of archival data was obtained from a human subjects institutional review board. Participants were selected from a pool of patients evaluated in a private practice setting. Of the 41 total participants, one was excluded due to incomplete data. The subjects ranged in age from 9 to 14 years ($M = 11.15$, S.D. = 1.642). Subjects were administered the full HRNB-OC as well as other measures.

1.1.2. Instruments and procedure
Total general neuropsychological deficit scale (GNDS) scores were computed for each subject using 45 variables from the HRNB-OC (see Reitan & Wolfson, 1992a, 1992b). Total neuropsychological deficit scale (NDS) scores for the screening battery (SBNDS) were also computed for each subject by converting raw scores of 12 variables to NDS scores per Reitan and Wolfson (2006). The SBNDS has two phases; Phase I includes one measure (i.e., Trail Making Test, Part B) and is designed to provide an initial identification of children who should have Phase II testing. Our study only examined the predictive validity of Phase II of the screening battery, which includes Part B of the Trail Making Test. Phase II of the screening battery is comprised of the following variables: total errors and greater number of errors minus lesser number of errors on the Bilateral Sensory Imperception Test; right hand and left hand errors and greater number of errors minus lesser number of errors on the Tactile Finger Recognition Test; right hand and left errors and greater number of errors minus lesser number of errors on the Finger-Tip Number Writing Test; total number of taps with dominant and nondominant hand and 1 minus nondominant hand score divided by dominant hand score on the Finger Tapping Test; total score on Reitan–Indiana Aphasia Test; total time on the Trail Making Test, Part B. For each of these variables raw scores were converted into NDS scores ranging from zero (normal range) to three (severe impairment). See Reitan and Wolfson (1992a, 2006) for a complete description.

1.2. Results
Correlational analyses between the GNDS and SBNDS were completed. Since these two variables are not independent, the measurement error of the variables are correlated which results in the spurious inflation of the correlation coefficients. According to Kaufman (1977), correction of the coefficients is not required when the purpose of the screening battery is to identify individuals requiring administration of the complete battery. Alternatively, if the purpose of a shortened test battery is to measure the same underlying trait or construct as the full test battery, then correction of the correlation coefficients would be required. Given that the SBNDS has been developed as a preliminary battery to determine the need for the complete HRNB-OC, correction for correlated error variances was not undertaken.

The Pearson correlation between the older children’s GNDS and SBNDS scores was .87 ($p \leq .01$). Using logistic regression with the GNDS from the full HRNB-OC battery as the criterion for determining neuropsychological impairment, the screening measure correctly classified 79.2% of children as unimpaired (GNDS score of ≤42) and 87.5% of children as impaired (GNDS score of ≥43). The overall classification rate was 82.5%.
Table 1

Matrix of true and false positives and negatives

<table>
<thead>
<tr>
<th>True state</th>
<th>Results of screening battery for older children</th>
<th>Total true states</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impaired</td>
<td>Impaired</td>
<td>14 (TP)</td>
</tr>
<tr>
<td>Normal range</td>
<td>Normal range</td>
<td>2 (FN)</td>
</tr>
<tr>
<td>Total test results</td>
<td>Total test results</td>
<td>19</td>
</tr>
</tbody>
</table>

Note. Values determined using a cut-off score of 16 on the screening battery and 43 on the full HRNB-C. TP, true positives; FN, false negatives; FP, false positives; TN, true negatives.

In calculating sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). Participants’ scores were classified using GNDS scores of 43 or greater as indicative of impairment and SBNDS scores of 16 or greater as suggestive of impairment on the GNDS. Reitan and Wolfson have recommended a cut-off score of 16/17 for the screening battery, and describe scores at or above this cut-off as indicating possible impairment of neuropsychological functioning (Reitan & Wolfson, 2006). Using these cut-off scores, subjects were either classified as true negatives (normal GNDS and SBNDS scores; \( n = 19 \)), true positives (impaired GNDS and SBNDS scores; \( n = 14 \)), false negatives (impaired GNDS and normal SBNDS scores; \( n = 2 \)), or false positives (normal GNDS and impaired SBNDS scores; \( n = 5 \)) (see Table 1). The screening battery was found to have a sensitivity of 87.5% and a specificity of 79.2%. The PPV of the screening battery was determined to be .737, signifying that if the screening battery predicts impairment, there is a 74% chance that the participant will be considered impaired on the full HRNB-OC. The NPV of the screening battery was found to be .905, indicating a 91% chance that a participant deemed normal on the screening battery will fall within the unimpaired range on the full HRNB-OC. While the screening battery identified the majority of children who evidenced neuropsychological impairment on the full battery, two children who showed impairment on the comprehensive battery were misclassified as unimpaired. In the case of screening batteries, the sensitivity of the instrument is more important than specificity, so those individuals with actual neuropsychological impairment are not missed. The predictive validity values for the screening battery were calculated using a cut-off of 14, which Reitan and Wolfson have suggested as a score at which consideration should be given for conducting a comprehensive neuropsychological evaluation. Using this cut-off score, participants were classified as true negatives (normal GNDS and SBNDS scores; \( n = 13 \)), true positives (impaired GNDS and SBNDS scores; \( n = 16 \)), false negatives (impaired GNDS and normal SBNDS scores; \( n = 0 \)), or false positives (impaired GNDS and normal SBNDS scores; \( n = 11 \)). Using the stated classifications, the screening battery was found to have a sensitivity of 100% and a specificity of 54.2%. The PPV of the screening battery was determined to be .593, suggesting that if the screening battery predicts impairment, there is a 59% chance that the participant will be considered impaired on the full HRNB-OC. The NPV of the screening battery was found to be 1.0, signifying a 100% chance. The sensitivity, specificity, PPV, and NPV were also calculated for a cut-off score of 15 on the screening battery. These values for each of the three cut-off scores are presented in Table 2.

A receiver operating characteristic (ROC) curve was generated (see Figs. 1–3) for each screening battery cut-off score (14, 15 and 16). The area under the curve (AUC) for a cut-off score of 16 was determined to be .833 (S.E. = .069); .760 (S.E. = .076) for the cut-off score of 15; .771 (S.E. = .074) for the cut-off score of 14. An ROC curve is the representation of the tradeoffs between sensitivity (false negative rate) and specificity (false positive rate). The higher the AUC score is, the better the test is at correctly classifying normal versus impaired subjects (Griner, Mayewski,

Table 2

Predictive utility of range of cut-off scores on screening battery for older children

<table>
<thead>
<tr>
<th>Score</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>1.00</td>
<td>.542</td>
<td>.593</td>
<td>1.00</td>
</tr>
<tr>
<td>15</td>
<td>.938</td>
<td>.583</td>
<td>.600</td>
<td>.933</td>
</tr>
<tr>
<td>16</td>
<td>.875</td>
<td>.792</td>
<td>.737</td>
<td>.905</td>
</tr>
</tbody>
</table>

PPV, positive predictive value; NPV, negative predictive value.
An AUC of .833 is considered to be a “B” (i.e., “good”) score (using the traditional academic point system: A = .90–1.0, B = .80–.90, C = .70–.80, etc.); AUC's of .760 and .771 are considered “C” scores (i.e., “fair”) (Black, Bordley, Tape, & Panzer, 1999).

2. Study 2: Young Children

2.1. Method

2.1.1. Participants
Following approval for analysis of archival data from the human subjects institutional review board, cases were selected from a pool of children evaluated in a private practice setting. From a total of 31 subjects, 3 were excluded from this study, as they did not complete all measures necessary for analyses. The children ranged in age from 5 to 8 years ($M = 6.78$, S.D. = 1.031). Each child was administered the RINTB as well as other measures.
2.1.2. Procedure

GNDS scores were computed for each subject using 52 variables from the RINTB (see Reitan & Wolfson, 1992b). Total NDS scores for the SBNDS were then computed for each subject by converting raw scores from 11 variables. This screening battery, similar to the older child screening battery, has two phases. Phase I consists of the Progressive Figures Test and is intended to provide an initial broad identification of children that should undergo Phase II screening. The focus of our study was the predictive validity of Phase II of the screening battery. The Phase II screening battery variables include: total errors and greater number of errors minus lesser number of errors on the Bilateral Sensory Imperception Test; right hand and left hand errors and greater number of errors minus lesser number of errors on the Tactile Finger Recognition Test; right hand and left errors and greater number of errors minus lesser number of errors on the Finger-Tip Symbol Writing Test; total number of taps with dominant and nondominant hand and 1 minus nondominant hand score divided by dominant hand score on the Finger Tapping Test; total score on the Reitan–Indiana Aphasia Test; total time on the Reitan–Indiana Color Form Test. For each of these variables raw scores were converted into scores ranging from zero (normal range) to three (severe impairment). See Reitan and Wolfson (1992b, 2006) for description.

2.2. Results

The Pearson correlation between the younger children’s GNDS and SBNDS scores was .92 ($p \leq .01$). The predictive utility of the screening battery was also investigated by calculating the sensitivity, specificity, PPV, and NPV. Participants’ scores were first classified using GNDS scores of 54 or greater as indicative of neuropsychological impairment and SBNDS scores of 22 or greater as suggestive of probable impairment on the GNDS, as recommended by Reitan and Wolfson (2006). Using these cut-off scores, participants were classified as true negatives (normal GNDS and SBNDS scores; $n = 25$), true positives (impaired GNDS and SBNDS scores; $n = 3$), false negatives (impaired GNDS and normal SBNDS scores; $n = 0$), or false positives (normal GNDS and impaired SBNDS scores; $n = 0$) (see Table 3). The screening battery was found to have a sensitivity of 100% and a specificity of 100%. The PPV of the screening battery was determined to be 1.0, indicating that if the screening battery predicts impairment, there is a 100% chance that the participant will be deemed impaired on the full RINTB. The NPV of the screening battery was found to be 1.0, indicating that there is a 100% chance that a participant judged as normal on the screening battery will be unimpaired on the full RINTB (Griner et al., 1981).

2.3. Discussion

The findings from the current study support the clinical utility of Reitan and Wolfson’s screening batteries with children referred for neuropsychological evaluation with possible acquired neurological damage due to a variety of
Table 3
Matrix of true and false positives and negatives

<table>
<thead>
<tr>
<th>True state</th>
<th>Impaired</th>
<th>Normal range</th>
<th>Total true states</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impaired</td>
<td>3 (TP)</td>
<td>0 (FN)</td>
<td>3</td>
</tr>
<tr>
<td>Normal range</td>
<td>0 (FP)</td>
<td>25 (TN)</td>
<td>25</td>
</tr>
<tr>
<td>Total test results</td>
<td>3</td>
<td>25</td>
<td>28</td>
</tr>
</tbody>
</table>

Note. Values determined using a cut-off score of 22 on the screening battery and 54 on the full RINTB. TP, true positives; FN, false negatives; FP, false positives; TN, true negatives.

causes. The screening battery for younger children had a perfect (100%) classification rate using a cut-off score of 22/23. All of the younger children who performed at or above this score on the screening battery also performed within the range of impaired neuropsychological functioning on the full GNDS of the Reitan Indiana Neuropsychological Test Battery; there were no false positives. The older children screening battery, using the recommended cut-off of 16/17 correctly identified 87.5% of the children whose neuropsychological functioning fell within the impaired range on the full Halstead–Reitan Neuropsychological Battery for Older Children, and 79.2% of the children who performed within the normal range on the full battery.

Our findings with the younger children battery are better than those of Reitan and Wolfson, who reported a sensitivity and specificity of 85%. The predictive values for the older children screening battery for our sample are similar to Reitan and Wolfson’s values of a sensitivity of 83% and specificity of 91%. In our sample of older children, the screening battery resulted in a higher number of false positive errors (specificity = 79%). These would be older children who, based on a low score on the screening battery, were administered the complete HRNB-OC and who ultimately performed within the normal range. Of more concern with a screening measure is the false negative rate. In this study, the false negative rate was low with just two older children (13%) incorrectly classified as having normal neuropsychological functioning based on the screening battery. Lowering the older child screening battery cut-off score to 14 identified all of the children who evidenced neuropsychological impairment of the full battery; however, the false positive rate rose considerably to 56%. The lower cut-off score of 14 may result in referral of a greater number of older children for a complete neuropsychological battery. This is arguably preferable to missing children who have neuropsychological impairment and who are in need of a complete evaluation for diagnosis and the determination of appropriate interventions to maximize development.

In order to support more effective clinical use of the screening battery for older children, the sensitivity, specificity, PPV, and NPV of the suggested cut-off scores were calculated for a variety of patient population base rates for neuropsychological impairment (see Table 4). Based upon our sample, a more conservative cut-off score of 14 would be recommended for children without confirmed positive neuroradiological findings as used in Reitan and Wolfson’s (2006) sample population.

In summary, the findings from this study provide evidence that the performance on the Reitan and Wolfson screening batteries for younger and older children are predictive of impairment on the HRNB-OC and the RINTB. While our

Table 4
Predictive power of range of screening battery scores and base rates for older children

<table>
<thead>
<tr>
<th>Score</th>
<th>.02</th>
<th>.10</th>
<th>.20</th>
<th>.50</th>
<th>.75</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PPV</td>
<td>NPV</td>
<td>PPV</td>
<td>NPV</td>
<td>PPV</td>
</tr>
<tr>
<td>14</td>
<td>.043</td>
<td>1.00</td>
<td>.195</td>
<td>1.00</td>
<td>.353</td>
</tr>
<tr>
<td>15</td>
<td>.044</td>
<td>.998</td>
<td>.200</td>
<td>.988</td>
<td>.360</td>
</tr>
<tr>
<td>16</td>
<td>.079</td>
<td>.237</td>
<td>.319</td>
<td>.983</td>
<td>.513</td>
</tr>
</tbody>
</table>

PPV, positive predictive value; NPV, negative predictive value.
findings represent a cross-validation of Reitan and Wolfson’s initial studies, we would recommend further validation of these two screening batteries with larger clinical samples.

Finally, it is important to bear in mind the proper use of screening batteries. Neuropsychological screening batteries are to be utilized solely to determine if comprehensive neuropsychological testing is warranted. The results from a screening battery cannot be used to arrive at a diagnosis of the presence of brain dysfunction, and do not provide sufficient information regarding a child’s cognitive strengths and weaknesses for guiding either treatment planning or interventions. In short, administration of the screening batteries is not a substitute for a comprehensive neuropsychological evaluation.

References


