Advances in Pediatric Neuropsychology Test Interpretation: Importance of Considering Normal Variability and Performance Variability

Brian L. Brooks, PhD
Alberta Children's Hospital
University of Calgary
Calgary, Alberta, Canada

and

Michael Kirkwood, PhD, ABPP/CN
Department of Rehab Medicine
Children’s Hospital Colorado
Aurora, Colorado, USA

National Academy of Neuropsychology, Annual Conference
Austin, Texas
November 5, 2015
Financial Disclosures

- **Employment**: Alberta Health Services
- **Consulting**: I have received payment as a consultant to Copeman Healthcare in Calgary, AB for TBI-related cases. Past consulting to Pearson Assessment as a beta tester for WMS-IV scoring program (reimbursement included a copy of the scoring program).
- **Stock ownership**: None relevant
- **Research support**: Principal investigator, co-investigator, or collaborator on grants funded by Alberta Children’s Hospital Research Institute, Canadian Institutes of Health Research, and Alberta Innovates Health Solutions. Past support from CNS Vital Signs (in-kind test credits for research) and collaborator on study funded by AstraZeneca Canada (>5 yrs ago)
- **Honoraria**: I have received honoraria and expense reimbursement from institutions for prior invited talks, including talks on normal variability and performance validity testing

---

Financial Disclosures

- I have financial relationships to disclose:
  
  1. **Co-author of two tests published by Psychological Assessment Resources, Inc.**
     - I receive royalties for the sale of these products
  
     - Brooks & Iverson chapter provides the basis for this talk
     - I receive royalties for the sale of this book
Objectives

1. Understand the difference between univariate and multivariate clinical interpretation.

2. Learn the 5 principles of multivariate base rates.

3. Appreciate how using multivariate base rates can reduce chances of over-interpreting isolated low scores.

Background

- Neuropsychology is well positioned to provide valuable information about whether a child’s abilities have been negatively affected by a disease or injury, to quantify the change in functioning, and to communicate the impact on day-to-day functioning.
  - 1. What is impacted?
  - 2. How much is it impacted?
  - 3. How does this impact real-world?
No other specialty has developed, normed, and validated measures of cognitive abilities in the same manner as neuropsychology.

The diligence of our field leads to lengthy assessments covering multiple cognitive domains and generating numerous scores.
Background

- Clinical neuropsychological assessments are estimated between 4.4-6.5 hours
  — Sweet et al., 2002

- The average forensic neuropsychological assessment is estimated at 9.5 hours
  — Sweet et al., 2002

- Our assessments result in a large amount of data that are gathered and subsequently analyzed
**24 scores**

- Motor Abilities
  - Motor Speed in Right Hand (Right Hand, CNS VS Finger Tapping)
  - Motor Speed in Left Hand (Left Hand, CNS VS Finger Tapping)
  - Left Hand Motor Dexterity (Purdue Pegboard)

- Executive Functioning
  - Impulsive Control - Verbal (NEPSY-II Inhibitory)
  - Impulsive Control - Verbal (NEPSY-II Inhibitory Combined Score)

- Learning and Memory for Verbal Information
  - Word List Learning (CVLT-C Trials 1-5)
  - Rate of Learning (CVLT-C Step Trials 1-5)
  - Long Delay Free Recall of Word List (CVLT-C LD/FR)

- Spatial Abilities
  - Visual-Spatial Skills (NEPSY-II Geometric Puzzles)

**33 scores**

- Executive Functioning
  - Visual-Motor Processing Speed (CVLT-C)
  - Visual-Motor Sensitivity (WAIS-IV Symbol Smoothing)
  - Visual-Motor Sensitivity (WISC-IV Coding)

- Learning and Memory for Visual Information
  - Word List Learning (CVLT-C Trials 1-5)
  - Rate of Learning (CVLT-C Step Trials 1-5)
  - Long Delay Free Recall of Word List (CVLT-C LD/FR)

- Spatial Abilities
  - Visual-Spatial Construction (WISC-IV Block Design)
  - Visual-Spatial Imagery (VMI)
Background

“Seeing the forest for the trees”

To discern an overall pattern from a mass of detail; to see the big picture, or the broader, more general situation

http://en.wiktionary.org/wiki/
Univariate Test Interpretation

- Univariate analyses: consideration of a single test score in isolation
- Bell curve generally applies

Assuming a normal distribution, what percent of the standardization sample obtains a score ≤5th percentile?

5%
Univariate Test Interpretation

- What about....?
  - If there are 2 scores?
  - If there are 5 scores?
  - If there are 50 scores?
  - If the person is low functioning?
  - If the person is high functioning?

  Reliance on the bell curve when interpreting multiple test scores will lead us astray...

Multivariate Test Interpretation

- Univariate clinical analyses: consideration of a single test score in isolation
  - This is not really what we do in neuropsychology

- Multivariate clinical analyses: consideration of multiple test scores simultaneously
  - This is neuropsychology
Two case examples to consider regarding multivariate interpretation

Case Example #1:

- 14-year-old previously healthy boy who sustained a concussion two years before assessment (slip and fall)
- Although family report vague, appears to be functioning similar to before the injury; similar academic performance
- Intellectual abilities estimated to be within the average range
- Due to complaints about memory problems, administered the CMS as part of assessment
Case Example #2:

- 11-year-old previously healthy girl who sustained a severe TBI in a high-speed MVC
- Lowest GCS 4/15, PTA and fluctuating orientation for 10 days, brain MR scan with diffuse and focal findings, numerous extra-cranial injuries
- Assessment 1.5 years after injury
- Patient was administered 17 subtests from the NEPSY-II as part of her assessment

### TABLE 4.2. Performance on the Children’s Memory Scale (CMS) Indexes in a 14-Year-Old Boy Who Sustained a Concussion

<table>
<thead>
<tr>
<th>CMS Index Scores</th>
<th>Standardized Performance and Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Index Score</td>
</tr>
<tr>
<td>Learning</td>
<td>103</td>
</tr>
<tr>
<td>Visual Immediate</td>
<td>103</td>
</tr>
<tr>
<td><strong>Visual Delayed</strong></td>
<td><strong>84</strong></td>
</tr>
<tr>
<td>Verbal Immediate</td>
<td>115</td>
</tr>
<tr>
<td>Verbal Delayed</td>
<td>106</td>
</tr>
<tr>
<td>Delayed Recognition</td>
<td>103</td>
</tr>
</tbody>
</table>
How can multivariate interpretation help? We will return to these examples later....
Multivariate Test Interpretation

- Historical context of multivariate test interpretation
  – Low scores/test-score scatter suggest something is wrong

- Are there empirical methods for interpreting multiple scores (multivariate clinical interpretation)?
  – Earliest work using the Halstead-Reitan Battery
    - Reitan & Wolfson, 1985, 1993;
“...it is a serious mistake to assume that one or more test scores beyond the acceptable cutoff scores always indicate the presence of an acquired cerebral disorder” (pp. 72-73).
Principles When Interpreting Multiple Scores

1. Test-score variability (scatter) is common
2. Having some low scores is common
3. The number of low scores is related to the cutoff score used
4. The number of low scores is related to the number of tests administered
5. The number of low scores varies by examinee characteristics

Principle 1

Test-score variability (scatter) is common

Also known as, an absence of variability or scatter in scores is uncommon (query pattern analysis)
Principle 1

Percent with 1, 2, 3, or 4SD spread between highest and lowest subtest scores on Wechsler tests

Data derived from Table B.9 in WPPSI-IV and WISC-V, and Table B.6 in WISC-IV and WAIS-IV administration and scoring manuals.

---

### Principle 1

- Scatter changes with level of functioning

*The Clinical Neuropsychologist*, 2013
Vol. 27, No. 6, 988-1003, [http://dx.doi.org/10.1080/13854046.2013.797502](http://dx.doi.org/10.1080/13854046.2013.797502)

**Normal Variability of Children's Scaled Scores on Subtests of the Dutch Wechsler Preschool and Primary Scale of Intelligence – Third Edition**

P.P.M. Hurks¹, J.G.M. Hendriksen², J.E. Dek³, and A.P. Koolj³

¹Maastricht University, Maastricht, The Netherlands
²Kempenhaeghe Center for Neurological Learning Disabilities, Heeze, The Netherlands
³Pearson Test Publishers, Amsterdam, The Netherlands
Principle 1

- Scatter changes with level of functioning

that large differences between highest and lowest scaled subtest scores (or subtest scatter) were common in this sample. Furthermore, degree of subtest scatter was related to: (a) the magnitude of the highest scaled subtest score, i.e., more scatter was seen in children with the highest WPPSI-III-NL scaled subtest scores, (b) Full Scale IQ (FSIQ) scores, i.e., higher FSIQ scores were associated with an increase in subtest scatter, and (c) sex differences, with boys showing a tendency to display more scatter than girls. In conclusion, viewing subtest scatter as an index for abnormality in WPPSI-III-NL scores is an oversimplification as this fails to recognize disparate subtest heterogeneity that occurs within a population of healthy children aged 4:0–7:11 years.

Summary: More scatter if higher subtest scores, in higher FSIQ, and in boys

Principle 2

Low scores are common

Also known as, an absence of low scores is uncommon
**Principle 2**

Percent with 1 or more scores at or below 5\textsuperscript{th} percentile on different pediatric batteries

![Bar chart showing percent with 1 or more scores at or below 5\textsuperscript{th} percentile for different batteries.](chart)

Univariate expectation

Figure 4.3; Brooks & Iverson, 2012

---

**Principle 3**

The number of low scores depends on cutoff

*Also known as, adjusting your cutoff score will adjust the number of low scores*
Principle 3

Percent with 1 or more low scores across different cutoff scores on three pediatric memory batteries

![Graph showing cumulative percent distribution of low scores across different cutoff scores.

Univariate expectations

Figure 4.4; Brooks & Iverson, 2012

Principle 4

Number of low scores depends on number of tests

Also known as, give more tests and get more low scores
Estimated percent of people with at least one low subtest score (<5th percentile) when varying the length of a battery

Note: Percent of people with low scores was estimated using the Crawford et al. (2007) Monte Carlo program and an average inter-subtest correlation of 0.3.

Figure 2; From Donders, Brooks, Sherman, & Kirkwood, in press

Percent with 1 or more scores at or below 5th percentile

Figure 4.5; Brooks & Iverson, 2012
Number of low scores depends on examinee’s characteristics

Also known as, examinee characteristics need to be considered

Percent with 1 or more WISC-IV subtest scores at or below 5th percentile by FSIQ categories

Figure 4.6; Brooks & Iverson, 2012
Percent with 1 or more Children's Memory Scale index scores at or below 5th percentile by WISC-IV FSIQ

Univariate expectation

Figure 4.7; Brooks & Iverson, 2012

Principle 5

Percent with 1 or more scores at or below 5th percentile by parent education

Univariate expectation

Figure 4.8; Brooks & Iverson, 2012
What is a clinician to do?

Multivariate analyses in pediatric neuropsychological evaluations

1. Knowledge is power
2. Use existing published tables (where available)
3. Compute your own multivariate base rates
What is a clinician to do?

- Knowing the prevalence of low scores can help to minimize the chance of misinterpretation of isolated low scores
  - Both misdiagnosis and missed diagnosis

- Multivariate analyses help determine if a certain number of low scores is uncommon

What is a clinician to do?

- Published tables with multivariate analyses are available for some pediatric neuropsychological tests
  - WISC-IV (Brooks, 2010; Brooks, 2011; Crawford et al., 2007)
  - Children’s Memory Scale (Brooks et al., 2009)
  - NEPSY-II (Brooks et al., 2010)
  - Child and Adolescent Memory Profile (Sherman and Brooks, 2015)
What is a clinician to do?

Brooks, 2010

Table 1
Base Rates of Low WISC-IV Subtest Scores by Impairment Cutoff, Level of Intelligence, and Parental Education

<table>
<thead>
<tr>
<th>Number of low WISC-IV scores</th>
<th>Total sample</th>
<th>Very low (&lt;30)</th>
<th>Below average (30-89)</th>
<th>Lower average (90-99)</th>
<th>Upper average (100-109)</th>
<th>Above average (110-119)</th>
<th>Very high (&gt;120)</th>
<th>Parental education (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤5th percentile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9-11</td>
</tr>
<tr>
<td>10 or more</td>
<td>0.5</td>
<td>5.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.5</td>
</tr>
<tr>
<td>9 or more</td>
<td>0.9</td>
<td>10.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.9</td>
</tr>
<tr>
<td>8 or more</td>
<td>1.3</td>
<td>15.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>3.8</td>
</tr>
<tr>
<td>7 or more</td>
<td>1.8</td>
<td>21.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>3.8</td>
</tr>
<tr>
<td>6 or more</td>
<td>2.2</td>
<td>26.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>3.7</td>
</tr>
<tr>
<td>5 or more</td>
<td>3.0</td>
<td>24.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>5.6</td>
</tr>
<tr>
<td>4 or more</td>
<td>4.7</td>
<td>38.8</td>
<td>0.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>9.3</td>
</tr>
<tr>
<td>3 or more</td>
<td>8.0</td>
<td>76.3</td>
<td>10.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>20.4</td>
</tr>
<tr>
<td>2 or more</td>
<td>14.2</td>
<td>93.5</td>
<td>34.1</td>
<td>3.5</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
<td>33.3</td>
</tr>
<tr>
<td>1 or more</td>
<td>31.7</td>
<td>98.9</td>
<td>79.4</td>
<td>32.2</td>
<td>11.7</td>
<td>2.2</td>
<td>59.3</td>
<td>62.0</td>
</tr>
<tr>
<td>No low scores</td>
<td>68.3</td>
<td>1.1</td>
<td>20.6</td>
<td>67.8</td>
<td>88.3</td>
<td>97.8</td>
<td>100</td>
<td>41.7</td>
</tr>
</tbody>
</table>

What is a clinician to do?

Sherman and Brooks, 2015

Table 10.1
Base Rates of Low Scores on the ChAMP

<table>
<thead>
<tr>
<th>Number of low subtest scores (SS ≤7 or ≤1 SD)</th>
<th>Total Sample</th>
<th>Parent education level (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>≤12</td>
</tr>
<tr>
<td>5 low scores</td>
<td>1.9</td>
<td>.8</td>
</tr>
<tr>
<td>7 or more</td>
<td>2.5</td>
<td>6.7</td>
</tr>
<tr>
<td>6 or more</td>
<td>6.5</td>
<td>11.5</td>
</tr>
<tr>
<td>5 or more</td>
<td>12.1</td>
<td>14.5</td>
</tr>
<tr>
<td>4 or more</td>
<td>18.6</td>
<td>19.0</td>
</tr>
<tr>
<td>3 or more</td>
<td>28.4</td>
<td>26.5</td>
</tr>
<tr>
<td>2 or more</td>
<td>41.6</td>
<td>40.2</td>
</tr>
<tr>
<td>1 or more</td>
<td>56.9</td>
<td>53.9</td>
</tr>
<tr>
<td>No low scores</td>
<td>43.1</td>
<td>46.1</td>
</tr>
</tbody>
</table>

Number of extremely low subtest scores (SS ≤4 or ≤2 SD)

<table>
<thead>
<tr>
<th>Number of low subtest scores (SS ≤4 or ≤2 SD)</th>
<th>Total Sample</th>
<th>Parent education level (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 low scores</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7 or more</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6 or more</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5 or more</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4 or more</td>
<td>1.5</td>
<td>3.3</td>
</tr>
<tr>
<td>3 or more</td>
<td>3.2</td>
<td>6.0</td>
</tr>
<tr>
<td>2 or more</td>
<td>9.1</td>
<td>9.5</td>
</tr>
<tr>
<td>1 or more</td>
<td>9.7</td>
<td>9.4</td>
</tr>
<tr>
<td>0 low scores</td>
<td>8.0</td>
<td>10.6</td>
</tr>
</tbody>
</table>

Note: N = 1,264. Analyses included scaled scores from Line, Line Delayed, Maze, Maze Instructions Delayed, Objects, Objects Delayed, Places, and Places Delayed. Line Recognition and Instructions Recognition are not included in these analyses. Cumulative percentages are reported for all values except for “0 low scores,” which is reported as an individual percentage.
What is a clinician to do?

- Can compute multivariate base rates for any group of scores using a Monte Carlo program if intercorrelations are known.

- Program publically available by Dr. John Crawford at http://homepages.abdn.ac.uk/j.crawford/pages/dept/psychom.htm

Webpage last accessed October 1, 2015
What is a clinician to do?

matrix entry:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.50</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.50</td>
<td>0.50</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>1.00</td>
</tr>
</tbody>
</table>

What is a clinician to do?

Results shown: PercentileNorm.RIV. Expected percentage of population with 1 or more abnormal scores and score difference.

- Estimated percentage of population with 1 or more abnormal low scores = 16.4359%
- Estimated percentage of population with 2 or more abnormal low scores = 1.5281%
- Estimated percentage of population with 3 or more abnormal low scores = 0.4909%
- Estimated percentage of population with 4 or more abnormal low scores = 0.1593%
- Estimated percentage of population with 5 or more abnormal low scores = 0.0000%

Results shown: PercentileNorm.RIV. Expected percentage of population with 1 or more abnormal scores and score difference.

Save Output  Clear Results  Return to Worksheet  Exit
What is a clinician to do?

- Monte Carlo estimation has good accuracy compared to actual base rates in standardization samples
- Caution with high or low functioning; subtest intercorrelations do not reflect

Applications of Multivariate Interpretation

Goal was to use the NIH pediatric sample to create a definition of “neuropsychological impairment” for future research comparisons
Applications of Multivariate Interpretation

24% of healthy children in the NIH sample had 1 or more scores more than 1.5SDs below the mean.

Thus, based on the frequency distributions presented in Table 6, the definition of neuropsychological impairment that best fits the NIHFD data and identifies approximately 95% of the population as “typically developing” is the following: “A neuropsychological impairment is present when an individual performs 1.5 standard deviations below the mean on two or more measures.” Our derived NPI rule identifies 5.1% of the total sample as impaired on two or more of the eight subtests in the assessment battery, which covers six domains of neuropsychological functioning. Applying this rule to individual age groups identifies between 3.0 and 7.2% of the population as impaired, suggesting that the case definition is appropriate for children between the ages of 6 and 18 years.

Applications of Multivariate Interpretation

Substantial risk of “Accidental MCI” in healthy older adults: Base rates of low memory scores in neuropsychological assessment

BRIAN L. BROOKS, GRANT L. IVESON, JAMES A. HOLDNACK

Potential for misclassification of mild cognitive impairment: A study of memory scores on the Wechsler Memory Scale-III in healthy older adults

BRIAN L. BROOKS, GRANT L. IVESON, JAMES A. HOLDNACK

25-30% of healthy adults would meet psychometric criteria for memory impairment based on 1 or more scores being 1.5SDs below the mean.
Table 3. Guidelines for determining memory impairment, based on level of functioning, when considering a cutoff of ≤50th percentile.

<table>
<thead>
<tr>
<th>Level of intelligence</th>
<th>n</th>
<th>Memory scores below cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Broadly normal</td>
</tr>
<tr>
<td>Unusually low (FSIQ ≤79)</td>
<td>40</td>
<td>0–3</td>
</tr>
<tr>
<td>Low average (79 ≤ FSIQ ≤ 89)</td>
<td>68</td>
<td>0–3</td>
</tr>
<tr>
<td>Average (90 ≤ FSIQ ≤ 109)</td>
<td>213</td>
<td>0</td>
</tr>
<tr>
<td>High average (FSIQ = 110–119)</td>
<td>83</td>
<td>0</td>
</tr>
<tr>
<td>Superior/very superior (FSIQ ≥120)</td>
<td>66</td>
<td>0</td>
</tr>
</tbody>
</table>

Level of estimated premorbid intelligence

<table>
<thead>
<tr>
<th>Level of intelligence</th>
<th>n</th>
<th>Memory scores below cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unusually low (WTAR-FLSIQ ≤79)</td>
<td>28</td>
<td>0–2</td>
</tr>
<tr>
<td>Low average (79 ≤ WTAR-FLSIQ ≤ 89)</td>
<td>62</td>
<td>0–2</td>
</tr>
<tr>
<td>Average (90 ≤ WTAR-FLSIQ ≤ 109)</td>
<td>255</td>
<td>0</td>
</tr>
<tr>
<td>High average (WTAR-FLSIQ = 110–119)</td>
<td>80</td>
<td>0</td>
</tr>
<tr>
<td>Superior/very superior (WTAR-FLSIQ ≥120)</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

Years of education

<table>
<thead>
<tr>
<th>Years of education</th>
<th>n</th>
<th>Memory scores below cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 years or less</td>
<td>194</td>
<td>0–1</td>
</tr>
<tr>
<td>9–11 years</td>
<td>70</td>
<td>0–1</td>
</tr>
<tr>
<td>12 years</td>
<td>151</td>
<td>0</td>
</tr>
<tr>
<td>13–15 years</td>
<td>71</td>
<td>0</td>
</tr>
<tr>
<td>16+ years</td>
<td>57</td>
<td>0</td>
</tr>
</tbody>
</table>

(≤50th percentile is a scaled score of 5 (mean = 10, SD = 3). The false positive rates in healthy older adults, which are presented in parentheses, are presumed because the healthy community-dwelling adult sample was not followed longitudinally to determine if some of them were experiencing prodromal AD. Intelligence is based on FSIQ scores from the WAIS-III [22]. Intellectual abilities are estimated using the WTAR-demographics prediction method [32].)
Case Example #1:

- 14-year-old previously healthy boy who sustained a concussion two years before assessment (slip and fall)
- Although family report vague, appears to be functioning similar to before the injury; similar academic performance
- Intellectual abilities estimated to be within the average range
- Due to complaints about memory problems, administered the CMS as part of assessment

<table>
<thead>
<tr>
<th>CMS Index Scores</th>
<th>Standardized Performance and Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning</td>
<td>Index Score 103 Percentile Rank 58 Classification Average</td>
</tr>
<tr>
<td>Visual Immediate</td>
<td>Index Score 103 Percentile Rank 58 Classification Average</td>
</tr>
<tr>
<td><strong>Visual Delayed</strong></td>
<td>Index Score 84 Percentile Rank 14 Classification Low Average</td>
</tr>
<tr>
<td>Verbal Immediate</td>
<td>Index Score 115 Percentile Rank 84 Classification High Average</td>
</tr>
<tr>
<td>Verbal Delayed</td>
<td>Index Score 106 Percentile Rank 66 Classification Average</td>
</tr>
<tr>
<td>Delayed Recognition</td>
<td>Index Score 103 Percentile Rank 58 Classification Average</td>
</tr>
</tbody>
</table>
Case #1 summary using multivariate:

- Obtained 1 index score at 14\textsuperscript{th} percentile on CMS
- According to Brooks et al. (2009), having 1+ index scores ≤16\textsuperscript{th} percentile is found in 37\% of healthy children and adolescents
- Considering only those with average intelligence, 1+ index scores ≤16\textsuperscript{th} percentile is found in 36\% of healthy children and adolescents
- Number of low index scores on the CMS would be considered ‘common’

Case Example #2

Case Example #2:

- 11-year-old previously healthy girl who sustained a severe TBI in a high-speed MVC
- Lowest GCS 4/15, PTA and fluctuating orientation for 10 days, brain MR scan with diffuse and focal findings, numerous extra-cranial injuries
- Assessment 1.5 years after injury
- Patient was administered 17 subtests from the NEPSY-II as part of her assessment
### Case #2 summary using multivariate:

- Several low scores found on the NEPSY-II
  - 8 scores ≤10th percentile
  - 4 scores ≤5th percentile
  - 2 scores ≤2nd percentile
- Having this many low scores is found in 0.9-5.2% of the standardization sample (range depends on cutoff selected) (Brooks et al., 2010)
- Number of low scores on NEPSY-II is ‘uncommon’

#### TABLE 4.3: Performance on Selected NEPSY-II Subtests in an 11-Year-Old Girl Who Sustained a Severe Traumatic Brain Injury

<table>
<thead>
<tr>
<th>NEPSY-II Domains and Subtests</th>
<th>Scaled Score</th>
<th>Percentile</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention and Executive Functioning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Animal Naming Total Correct</td>
<td>6</td>
<td>9</td>
<td>Borderline</td>
</tr>
<tr>
<td>Auditory Attention Total Correct</td>
<td>6</td>
<td>9</td>
<td>Borderline</td>
</tr>
<tr>
<td>Response Set Total Correct</td>
<td>5</td>
<td>5</td>
<td>Borderline</td>
</tr>
<tr>
<td>Inhibitions Naming Total</td>
<td>6</td>
<td>9</td>
<td>Borderline</td>
</tr>
<tr>
<td>Completion Time</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inhibitions Additions Total</td>
<td>4</td>
<td>2</td>
<td>Extremely Low</td>
</tr>
<tr>
<td>Completion Time</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inhibitions Switching Total</td>
<td>2</td>
<td>&lt;1</td>
<td>Extremely Low</td>
</tr>
<tr>
<td>Language</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension of Instructions Total</td>
<td>11</td>
<td>65</td>
<td>Average</td>
</tr>
<tr>
<td>Phonological Processing Total</td>
<td>9</td>
<td>37</td>
<td>Average</td>
</tr>
<tr>
<td>Speeded Naming Total</td>
<td>7</td>
<td>16</td>
<td>Low Average</td>
</tr>
<tr>
<td>Completion Time</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memory and Learning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memory for Designs total</td>
<td>9</td>
<td>5</td>
<td>Average</td>
</tr>
<tr>
<td>Memory for Designs Delayed total</td>
<td>8</td>
<td>25</td>
<td>Average</td>
</tr>
<tr>
<td>Narrative Memory Free &amp; Cond</td>
<td>6</td>
<td>9</td>
<td>Borderline</td>
</tr>
<tr>
<td>Recall Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Narrative Memory Free Recall Total</td>
<td>5</td>
<td>5</td>
<td>Borderline</td>
</tr>
<tr>
<td>Word List Interference Repetition</td>
<td>8</td>
<td>25</td>
<td>Average</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word List Interference Recall Total</td>
<td>7</td>
<td>16</td>
<td>Low Average</td>
</tr>
<tr>
<td>Visuospatial Processing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block Construction Total Score</td>
<td>10</td>
<td>40</td>
<td>Average</td>
</tr>
<tr>
<td>Geometric Puzzles Total Score</td>
<td>12</td>
<td>75</td>
<td>Average</td>
</tr>
</tbody>
</table>
Conclusions

- Interpretation of *multiple* test scores is different than interpretation of an isolated *single* test score
- Clinicians should appreciate the five principles of multivariate test interpretation

Conclusions

- Every test will have a “normal” amount of scatter and a “normal” amount of low scores that need to be accounted for when interpreting results
- Multivariate interpretation increases empirically-based conclusions on neuropsychological data
  — Provides empirical basis for “pattern analysis”
Multivariate Test Interpretation

Has the field moved forward?

A significant discrepancy between any two or more scores is the basic element of test score interpretation. Any single discrepant score or response error can usually be disregarded as a chance deviation. A number of errors or test score deviations may form a pattern. Marked quantitative discrepancies have the most difficulty. When the pattern of impaired functions or lowered test scores does not appear to be consistently associated with a known or neurologically meaningful pattern of cognitive dysfunction, discrepant scores may well be attributable to psychogenic, developmental, or chance deviations (L.M. Binder, Iverson, and Brooks, 2009).

Cautions and Caveats

- Multivariate analyses *supplement, but do not replace*, clinical judgment
- Presence of more low scores than expected is not diagnostic
- Having a low score may not be ‘uncommon’, but could still impact functioning and merit accommodation
- Caution against substituting tests with existing tables
  — See next figure
Is substitution of scores problematic?

Estimated percent of healthy people who would obtain at least one subtest score <5th percentile across different subtest intercorrelations

Note: Percent of people with low scores was estimated using the Crawford et al. (2007) Monte Carlo program.

Figure 3: From Donders, Brooks, Sherman, & Kirkwood, in press

Collaborators

- Primary collaborators for multivariate base rate research:
  - Dr. Grant Iverson
  - Dr. James Holdnack, Pearson (now University of Delaware)
  - Dr. Elisabeth Sherman
  - Dr. Travis White, PAR Inc.
  - Dr. Larry Binder

- Primary reference: