FINAL RESULTS OF THE WJ-IV ACH AND GORT-5 STUDY: A WOODCOCK INSTITUTE FUNDED PROJECT

A SPECIAL THANKS TO THE WOODCOCK INSTITUTE

- Small grants program aimed at advancing:
  - Interdisciplinary research into the cognitive profiles of individuals with diagnosed exceptionalities (learning disabilities, neuropsychological conditions, behavioral and psychiatric disorders, and giftedness),
  - Effective clinical assessment practices and the dissemination of research findings through direct professional development opportunities and publications and applied evidence-based assessment
  - https://twu.edu/woodcock-institute/
A SPECIAL THANKS TO OTHER CONTRIBUTORS

- Jackson County Schools, West Virginia
  - Mallory Frampton
  - Paula King
  - Rachel James
  - Mary Toler
  - Becky Wendell
- Gallipolis City Schools, Ohio
  - Heidi Creamer
  - Sandra Stroebel, PhD
  - All Student Participants

OVERVIEW OF TODAY’S PRESENTATION

- Purpose of Study
- Background Information
- Research Questions
- Method & Results
- Conclusion
Evaluate the relationship between the GORT-5 and the WJ-IV ACH reading tests.

Practice-driven study

Initiated to fill gaps in STAR Reading data in one school district’s multi-tiered system of support (MTSS)

---

Renaissance Learning™
Computer Adaptive Testing available since 1996
Assesses Five areas
- Word Knowledge and Skills
- Comprehension Strategies and Constructing Meaning
- Analyzing Literary Text
- Understanding Author’s Craft
- Analyzing Argument and Evaluating Text
Generates scaled scores, percentiles, and lexiles
STAR Early Reading used in grades K-2 in our district

STAR READING

HTTP://WWW.RTI4SUCCESS.ORG/RESOURCES/TOOLS-CHARTS/SCREENING-TOOLS-CHART

<table>
<thead>
<tr>
<th>Tool</th>
<th>Area</th>
<th>Classification Accuracy Rating</th>
<th>Generalizability</th>
<th>Reliability</th>
<th>Validity</th>
<th>Unaggregated Reliability, Validity, and Classification Data for Diverse Population</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAR</td>
<td>Reading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Administration &amp; Scoring Time, Scoring Key, Benchmarks / Norms</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Broad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10 Minutes, Computer Scored</td>
<td>Yes</td>
</tr>
</tbody>
</table>

[Graph showing reading scores over time]
Make it difficult to:
1) interpret individual student performance in a MTSS framework
2) subsequently recommend appropriate interventions and special education referrals.
Support need for an additional diagnostic instrument to guide referral decisions
- For select students at Tier 2 / Tier 3

<table>
<thead>
<tr>
<th>Referral Decision: Reject null hypothesis.</th>
<th>Null hypothesis ($H_0$) is, in fact, true</th>
<th>Null hypothesis ($H_0$) is, in fact, false</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO DISABILITY</td>
<td>Type I error False positive</td>
<td>Correct outcome True positive (Sensitivity)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Referral Decision: Accept null hypothesis.</th>
<th>Null hypothesis ($H_0$) is, in fact, false</th>
<th>Null hypothesis ($H_0$) is, in fact, true</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUSPECTED DISABILITY</td>
<td>Correct outcome True negative (Specificity)</td>
<td>Type II error False negative</td>
</tr>
</tbody>
</table>

Null hypothesis ($H_0$) is, in fact, true
No disability
Null hypothesis ($H_0$) is, in fact, false
Suspected disability

Referral Decision:
- Reject null hypothesis (Type I error: False positive)
- Accept null hypothesis (Type II error: False negative)

Statistical Conclusion
- Validity

Type I error: False positive
Type II error: False negative
Correct outcome True positive (Sensitivity)
Correct outcome True negative (Specificity)
GRAY ORAL READING TESTS - FIFTH EDITION (GORT-5)

- Requires direct interaction between the examiner and examinee
- Informs intervention and the extent to which attention and motivation impact reading results
- First published in 1963 by Dr. William S. Gray
- Selected for use because it was commonly used in initial referrals by two local school psychologists

Administration: Individual.
Time: 15-45 minutes.
Publisher: PRO-ED.
Two Parallel Forms: A and B

The Gray Oral Reading Tests, Fifth Edition (GORT-5) provides an objective measure of oral reading Rate, Accuracy, Fluency, and Comprehension.

Consists of 14 developmentally sequenced reading passages with five comprehension questions following each story.

Vocabulary in the questions was controlled to ensure the vocabulary would not be more difficult than that in the stories.

Age Range: 6 years 0 months to 23 years 11 months

Wiederholdt & Bryant, 2012
RESEARCH QUESTIONS

- What is the correlation between the GORT-5 Oral Reading Index (ORI) and the WJ-IV ACH Broad Reading score?
- Are the GORT-5 ORI standard scores for participants equivalent to the WJ-IV ACH Broad Reading, Reading Fluency, and Reading Comprehension standard scores?
- Does GORT-5 ORI reading performance at or below the tenth percentile predict WJ-IV ACH performance at or below the tenth percentile?
- What is the Area Under the Curve (AUC) and sensitivity/specificity of the GORT-5, as measured by a Receiver Operator Characteristic (ROC) Curve?
METHOD

Participants
- 118 students
- Referred sample
  - Includes gifted exceptionality in West Virginia
  - Females (n=41; 34.7%), Males (n=77; 65.3%),
- Elementary participants
  - Grades 2-5 (n=60; 50.8%)
- Middle level participants
  - Grades 6-8 (n=34; 28.8%)
- High school participants
  - Grades 9-12 (n=24; 20.3%)
METHOD

Participants

- 1.7% Hispanic
- 5.0% Multiple Races
- 90.8% White, non-Hispanic
- 2.5% Other Race/Ethnicities

METHOD

- Measures
  - GORT-5
  - WJ IV Tests of Achievement (all reading tests)
Procedure
- Trained professionals including school psychologists, diagnosticians, and school psychology candidates
- Administered both assessments under standard conditions as outlined by tests manuals
- Attempted to counterbalance but unable to do so in all situations due to constraints of implementing in a natural setting
- Recorded days between testing

Mixed method approach to analysis
- Pearson product-moment correlation coefficients
- T-tests as an estimate for score equivalence
- Chi-square like tests
- ROC Curves
BENEFIT OF ROC CURVES

- Receiver Operating Characteristic
- Used frequently in the medical field
- Curve that visualizes the trade off or relationship between sensitivity and specificity of a test or multiple assessments
- Additionally helps us understand the benefit of using an assessment to diagnose a problem
- Provides the best cut-off to ensure the highest rate of true positives with the lowest rate of false positives.
- The Area Under the Curve (AUC) shows how well the test can differentiate students with disabilities from those without disabilities (or the condition of interest)

RESULTS
## RESEARCH QUESTION I: CORRELATIONS

<table>
<thead>
<tr>
<th>WJ IV TESTS</th>
<th>GORT-5 ORI</th>
<th>GORT-5 Fluency</th>
<th>GORT-5 Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad Reading</td>
<td>.893*</td>
<td>.894*</td>
<td>.791*</td>
</tr>
<tr>
<td>Reading Fluency</td>
<td>.852*</td>
<td>.872*</td>
<td>.725*</td>
</tr>
<tr>
<td>Basic Reading Skills</td>
<td>.869*</td>
<td>.865*</td>
<td>.784*</td>
</tr>
<tr>
<td>Reading Comprehension Ext.</td>
<td>.874*</td>
<td>.808*</td>
<td>.855*</td>
</tr>
</tbody>
</table>

Note: *pp < .001

## RESEARCH QUESTION II: SCORE COMPARABILITY

<table>
<thead>
<tr>
<th>Pair 1* n=118</th>
<th>GORT-5 ORI</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WJ IV Broad Reading</td>
<td>87.2</td>
<td>14.6</td>
<td>1.3</td>
</tr>
<tr>
<td>Pair 2* n=113</td>
<td>GORT-5 ORI</td>
<td>86.7</td>
<td>14.2</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>WJ IV Reading Comp Ext</td>
<td>83.9</td>
<td>15.9</td>
<td>1.5</td>
</tr>
<tr>
<td>Pair 3* n=117</td>
<td>GORT-5 ORI</td>
<td>87.4</td>
<td>14.6</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>WJ IV Reading Fluency</td>
<td>85.0</td>
<td>17.8</td>
<td>1.6</td>
</tr>
</tbody>
</table>
RESEARCH QUESTION III: WILL KNOWING GORT-5 ORI HELP US PREDICT WJ IV PERFORMANCE?

<table>
<thead>
<tr>
<th>GORT-5 ORI</th>
<th>WJ IV BROAD READING</th>
</tr>
</thead>
<tbody>
<tr>
<td>At or below 10th percentile</td>
<td>32.2% (38 students)</td>
</tr>
<tr>
<td>&gt; 10th Percentile</td>
<td>6.8% (8 students)</td>
</tr>
</tbody>
</table>

RECEIVER OPERATOR CHARACTERISTIC (ROC) CURVE: SENSITIVITY AND SPECIFICITY
CONCLUSION

- Very large correlations (.725-.893)
- Average GORT-5 ORI is 3.6 points higher than WJ IV Broad Reading Score.
- Floor on WJ IV is lower and better discriminates students with extreme reading struggles.
- Lowest possible GORT-5 ORI score is 52, whereas WJ IV standard scores minimum is <40
In the current sample, the addition of the GORT-5 as a diagnostic screener would enhance our statistical conclusion validity as practitioners may reduce need for unnecessarily lengthy special education evaluation for students whose Tier 2 / Tier 3 progress monitoring results are inconclusive, yet GORT-5 ORI is > or = 81. However, if the 10th percentile were used as the cut score, eight students (6.8%) would be overlooked when, they, in fact, performed significantly below the age expected range on the WJ IV in Broad Reading.

**CONCLUSION**

**Sensitivity & Specificity**
- AUC value of .947 indicates the GORT-5 reliably distinguishes among students with satisfactory and unsatisfactory reading performance on the WJ IV, whereas values at 0.50 indicate the predictor is no better than chance.
- Thus, for students whose STAR Reading (or other progress monitoring results) are unclear or highly variable, the GORT-5 can be a strong predictor of performance on the WJ IV.
### PRACTICAL CONSIDERATIONS

- When to add additional assessment into MTSS process?
- Who will assess?
- How can we maximize instructional benefit from assessment, in addition to informing referral decision?

### LIMITATIONS

- Regional sample was not a representative of national demographics
  - Results will need to be cross-validated by other studies
- More participants needed without any referral concerns
- Differential classification accuracy has not been examined yet across demographic groups
  - Currently have students examining classification accuracy by programmatic level and disability status.
MORE ABOUT THE WOODCOCK INSTITUTE GRANT PROGRAM

Woodcock Institute

Woodcock Institute
For the Advancement of Neurocognitive Research and Applied Practice
TEXAS WOMAN’S UNIVERSITY

REFERENCES

CONTACT INFORMATION

R. LANAI JENNINGS
ASSISTANT PROFESSOR
MARSHALL UNIVERSITY
JENNINGSKRUT@MARSHALL.EDU