If you are viewing this course as a recorded course after the live webinar, you can use the scroll bar at the bottom of the player window to pause and navigate the course.

This handout is for reference only. It may not include content identical to the powerpoint. Any links included in the handout are current at the time of the live webinar, but are subject to change and may not be current at a later date.
Evidence-Based Practices in Selecting and Using Standardized Tests: Psychometric Adequacy

Dr. Stacey Pavelko, CCC-SLP

Learner Outcomes

1. Describe the 10 psychometric criteria
2. Describe an adequate norming sample
3. Define reliability and validity
Advance Organizer

- How to pick tests
  - Psychometric Characteristics
  - Diagnostic Accuracy
  - Test Content
  - Considerations for Dialect and ELL
  - Cultural/Linguistic Load

SLP Goals for Assessment

1. Children with *Speech-Language Impairment* correctly identified as SLI
2. Children with *Speech-Language Impairment* inappropriately found “not impaired”
3. Children inappropriately identified as *Speech-Language Impairment*
4. Children without impairment found “not impaired”
Test Considerations

Psychometric Considerations

What is a “good” test?

- Psychometrically Adequate
- Diagnostically Accurate
- Appropriate for your client and purpose of testing
  - Purpose of the test
  - Culturally/Linguistically Appropriate
  - Norming Sample
Psychometric Adequacy

- Standardization Sample
- Sample Size
- Systematic Item Analysis
- Mean and Standard Deviation of Test Scores
- Concurrent Validity

McCauley & Swisher (1984)

Psychometric Adequacy

- Predictive Validity
- Test-Retest Reliability
- Inter-Examiner Reliability
- Test Administration Procedures
- Special Qualifications

McCauley & Swisher (1984)
**Standardization Sample**

- Description of the number of sample participants
  - geographic residence
  - socioeconomic status
  - information regarding the “normalcy” of the participants, including any participants who were excluded because they exhibited nonnormal development or language.

McCauley & Swisher (1984)

**Standardization Sample Size**

- at least 100 participants
  - for each age or grade.

McCauley & Swisher (1984)
Systematic Item Analysis

- evidence of quantitative methods to study and control
  - item difficulty
  - item validity


Mean and Standard Deviation of Test Scores

- means and standard deviations
  - total raw scores of all relevant subgroups

Reliability
the degree to which an assessment tool produces stable and consistent results.

Validity
A test is valid if it measures what it is supposed to measure.

Reliability and validity are independent of each other. A measurement maybe valid but not reliable, or reliable but not valid.
Concurrent Validity

categorization of children as impaired or normal agreed closely with other valid methods such as scores on other validated tests


Predictive Validity

Evidence that it could be used to predict later performance of the same language or literacy behavior.

**Test-Retest Reliability**

Test-retest reliability coefficient of 0.90 or higher.


**Inter-Examiner Reliability**

Inter-examiner reliability coefficient of 0.90 or higher.

Administration Procedures

Sufficient detail to allow the user to duplicate the administration and scoring procedures used during standardization.


Qualifications

The manual described any special qualifications required of the test administrator or scorer.

Improvements in Reliability & Validity?
McCabe & Swisher (1984)
• 30 Tests of Language and Articulation

<table>
<thead>
<tr>
<th>Number of Criteria Met</th>
<th>Number of Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

Improvements in Reliability & Validity?
McCabe & Strand (2008)
• 6 Tests of Motor Speech Disorders in Children

<table>
<thead>
<tr>
<th>Number of Criteria Met</th>
<th>Number of Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3 or more</td>
<td>0</td>
</tr>
</tbody>
</table>
Improvements in Reliability & Validity?

Kirk & Vigeland (2014)

- 6 Tests of Phonological Error Patterns

<table>
<thead>
<tr>
<th>Number of Criteria Met</th>
<th>Number of Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4 or more</td>
<td>0</td>
</tr>
</tbody>
</table>

By becoming more aware of the strengths and limitations of these tests, clinicians will be able to select the instrument that is most appropriate for their specific needs

Kirk & Vigeland (2014)
Resources


What’s Next?

The Science…Diagnostic Accuracy