What is valid? What is fair?

Technical requirements for the assessment of ELL students

H. Gary Cook, Ph.D.
Value-added Research Center
Overview

• What should we expect to see in an ELL Assessment Technical Manual?
• What are the characteristics of student growth on ELL Assessments:
  - As it relates to AMAOs
  - As it relates to expectations of student growth
ELL Assessment Technical Manual

• Purpose of the assessment
• Test Design
• Test Technical Characteristics
• Summary of the Year’s Test Administration
Purpose of the assessment

• Purpose of the assessment *(Boilerplate)*
  - Program compliance
  - Student placement
  - Student proficiency
  - Student progress

• Appropriate inferences made by the assessment
Test Design

- **Test Specifications** *(Boilerplate)*
  - Standards on which test is based
  - Assessed domains and grades
  - Item formats and organization
  - Numbers and types of items by domain and grade

- **Field and Operational Test Design** *(Annual Update)*
  - How are items field tested? (standard-alone, embedded designs)
  - What is the operation test design? Refresh rate?
Test Technical Characteristics

- **Scaling & equating design** *(Boilerplate)*
  - Scaling methodology
  - Horizontal and vertical equating design
  - Adopted scale and its characteristics

- **Setting performance standards** *(Boilerplate)*
  - Standard setting methodology
  - Standard setting process
  - Proficiency level cut scores
Test Technical Characteristics

• Reliability *(Annual update)*
  - Test reliability
  - Rater reliability/ generalizability
  - Measurement error

• Validity—often this is the slimmest part of technical manuals. It shouldn’t be!!!
Test Technical Characteristics-Validity

- Construct validity evidence
  - test & item intercorrelations (*Annual update or Boilerplate*)
  - underlying trait structure of tests and scores
  - relationship between ELL and academic content assessments
  - growth profiles
Test Technical Characteristics-Validity

• Criterion-related validity evidence *(Annual update or Boilerplate)*
  - Concurrent and Discriminant evidence

• Content validity evidence *(Annual update or Boilerplate)*
  - alignment studies
  - content expert and bias review process
Test Technical Characteristics-Validity

• Consequential validity evidence *(Annual update or Boilerplate)*
  - Focus groups, surveys
  - Observations of classroom behavior
  - Survey’s of Enacted Curriculum-like analyses

• Studies by other sources on test’s validity *(Annual update or Boilerplate)*
Summary of This Year’s Test Administration

(Annual update)

- Test Form Overview & Issues
- Test Administration Overview & Issues
- Scoring and reporting Overview & Issues
Student Growth on ELL Assessments

• Annual Measurable Achievement Objectives (AMAO)
  – AMAO 1-Progress
  – AMAO 2-Proficiency

• Understanding the nature of student growth on ELL assessments will help set meaningful AMAOs

• The following is a 3-state sample of student progress over 3 years using ACCESS for ELLs data.
Percent Gaining One or More Proficiency Levels by Grade Band and Proficiency Level
Annual Scale Score Growth by Grade Band and Level

ACCESS Scale Score Growth Between SY 2005-2006 Across 3 States

<table>
<thead>
<tr>
<th>Initial Proficiency Level</th>
<th>Grade Band K-2</th>
<th>Grade Band 3-5</th>
<th>Grade Band 6-8</th>
<th>Grade Band 9-12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>54.60</td>
<td>34.47</td>
<td>28.38</td>
<td>15.43</td>
</tr>
<tr>
<td>2.00</td>
<td>31.62</td>
<td>21.43</td>
<td>17.29</td>
<td>8.94</td>
</tr>
<tr>
<td>3.00</td>
<td>21.77</td>
<td>14.43</td>
<td>11.77</td>
<td>6.43</td>
</tr>
<tr>
<td>4.00</td>
<td>13.88</td>
<td>10.28</td>
<td>7.18</td>
<td>3.99</td>
</tr>
<tr>
<td>5.00</td>
<td>7.69</td>
<td>4.00</td>
<td>10.18</td>
<td>1.31</td>
</tr>
</tbody>
</table>
Student Growth on ELL Assessments

- From observed data on ACCESS for ELLs, we observe the following
  - Younger students grow faster than older students.
  - Lower proficiency levels grow faster than higher proficiency levels.
  - There is an interaction between student age and student proficiency level.

- The above features are observed on other ELL assessments as well.
The Shape of ELL Assessment Student Growth

Estimated ACCESS Composite Scale Score Growth Profile by Grade Cluster

Years

ACCESS Composite Score

Grades 1-2
Grades 3-5
Grades 6-8
Grades 9-12
The Shape of Student Growth at Grade Band 3-5 by Level

Estimated ACCESS Composite Scale Score Growth Profile for the 3-5 Grade Cluster by Starting Proficiency Level
Things to Think About

• What contributes to the observed student growth characteristics? The assessment itself? The nature of child language acquisition?
• What factors affect the shape of student growth?
• What are the policy implications if group characteristics affect the shape of student language growth on these assessments?