Tests of Adult Basic Education

Workshop Objectives

- Overview of TABE
- Top Best Practices
- Interpreting TABE results
- Discussion of how to accommodate students with disabilities
- Discussion of basic test management
- Linking to instruction
- Bridging the gap to TABE 11/12
- Overview of TABE 11/12
- Questions
History of TABE

- TABE 1 & 2 1967
- TABE 3 & 4 1974-1975
- TABE 5 & 6 1983
- TABE 7 & 8 1994
- TABE 9 & 10 2003-2004
  - TABE Online 2006
- TASC 2014
- TABE 11 & 12 2016

Versions of TABE

- Two Alternate Forms – 9 & 10
- Two versions for each form
  - Complete Battery (3 hours, 37 minutes)
    - Complete norm- and criterion-referenced information
    - Detailed diagnostic information
  - Survey (2 hours, 8 minutes)
    - Shorter testing time
    - Easy screening and placement
Flexible Options: Tabe Online

- Web based version of Tabe 9&10
- Rapid Registration of student in the system
- Test scored immediately and automatically
- Auto Locator for placing students in test level
- Completely secure test administration and data storage
- Timer and book marking features
- Reports allow for aggregation of data
- Improved accommodation functionality – font color/size

Flexible Options: Tabe PC

- Entire test or any part completed on computer
- Test scored immediately and automatically
- Auto Locator for placing students in test level
- Option for randomizing distractors/ correct answer to improve test security
- Timer and book marking features
- Option for integrating data with TestMate Tabe testing and upload
- Same reporting formats as TestMate Tabe
Frequency of TABE assessments?

• Time prior to post-testing
  • 50-60 hours of instruction (40 min) recommended for ABE students
  • 30-59 hours of instruction recommended for ASE students
    • same level, different form
  • 120 hours of instruction recommended
    • same level, same form

Best Practices in Administering TABE

• Inform student about the test experience.
  • Purpose of testing and use of results
  • When and where the test will be given
  • What subject matter will be tested
  • Types of test questions
  • Timing and test length
  • Practice questions
• Create and maintain appropriate testing environment.
  • Trained administrators and proctors
  • Appropriate spacing/seating of students
Variables That Can Affect Test Results

**Student**
- Fatigue
- Motivation level
- Physical discomfort
- Test anxiety

**Environmental**
- Light levels
- Temperature
- Noise level
- Ventilation
- Other distractions

Best Practices in Administering TABE

- Use “Locator” to determine appropriate test level.
  - May not be used instead of test level
  - Is not an absolute prediction of ability
  - Has limited number of items
  - Avoid administering a level “too easy” or “too difficult”

- Use determined test level in a “standardized” manner.
  - Use proper timing
  - Read script verbatim
  - Follow test directions explicitly
  - Ensure comparable testing situation for a
How to Help Examinees

Do explain:
- Won’t know all the answers
- Time frame
- No preparation needed
- Current skills
- Breaks
- No trick questions

Do not say:
- Easy test
- Extra time is allowed

TABE Accommodations

- Making Test Accessible for Adult Students with Special Needs
  - TABE 9 & 10 - Large Print-(available in Complete and Survey editions)
  - TABE 9-Braille-(available in Complete and Survey Editions).
  - Use of TABE Online or TABE PC.

- Must self-identify
- Must provide approved documentation
- Must request accommodation
Security Issues

- Handling Testing Materials
  - Tracking with TABE Administrators, Students
  - Keeping Materials in Locked Storage
- No cell phones or electronic devices
- Number tests and answer sheets
- Plan for contingencies
- Active monitoring
  - 1 proctor for every 15 students being tested
  - Direct observation for entire test period

Score Range vs. Content Range

<table>
<thead>
<tr>
<th>Skill Level</th>
<th>Grade Equivalent Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>0.0–4.9+</td>
</tr>
<tr>
<td>E</td>
<td>0.0–6.9+</td>
</tr>
<tr>
<td>M</td>
<td>0.0–9.9+</td>
</tr>
<tr>
<td>D</td>
<td>0.7*–12.9+</td>
</tr>
<tr>
<td>A</td>
<td>1.1*–12.9+</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Skill Level</th>
<th>Grade Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>0–1.9</td>
</tr>
<tr>
<td>E</td>
<td>2.0–3.9</td>
</tr>
<tr>
<td>M</td>
<td>4.0–5.9</td>
</tr>
<tr>
<td>D</td>
<td>6.0–8.9</td>
</tr>
<tr>
<td>A</td>
<td>9.0–12.9</td>
</tr>
</tbody>
</table>
Turning the Focus to Instruction

Item Analysis Data

1. Individual Diagnostic Profile
2. TestMate for TABE software or TABE PC
3. TABE Online Group Item Analysis Report
4. TABE Teachers Guides Instructional Binders
Best Practices for Administering TABE

- Follow test publisher’s recommended testing procedures
- Be concerned about test security
- Follow NRS and state guidelines regarding the timing and frequency of testing
- Use the TABE to drive instruction
**Tabe Current Status**

- TABE 9/10 is approved at least through 2017
  - Most asked questions, alignment to HSE and 11/12
- TABE Online is helping to field test new items now
  - Testlets are the bridge between today and the new NRS
- More customers still use paper version than computer based
  - Winter field testing includes paper version
- TABE CLAS-E approved in 20 states
  - CLAS-E Online is pending the release of the new NRS EFLs

**NRS Changes**

- Public Comment period for WIOA draft regulations is closed
  - [www.nrsweb.org](http://www.nrsweb.org)
- Rules for all parts of NRS testing/reporting and AEFLA funding
- Oct. 1, 2016 is the first purposed application date for publishers
**TABE Plans for 2015 and Beyond**

- New FREE Formative tests released in October
  - Aligned to College and Career Readiness Standards
  - Included new Technology Enabled items
- Provides exposure to new CCSS/CCR items before HSE tests
- Does not effect reportable TABE scores
- Content areas include Reading, Math and Language
- Testlets are optional and can be used at anytime after initial TABE test

**TABE Plans for 2015 and Beyond**

**TABE 11/12**

- Field Testing late December/January
  - Sign up to receive email updates
- College and Career Readiness Standards
- New length of the test
- Changes to Math sections
- Improved Locator design
- Reading, Math and Language tests only
- Spanish version planned
- Alignment to all 3 HSE exams
TABE Plans for 2015 and Beyond

TABE PC
- New Platform planned for TABE 11/12
- Several options being discussed

TestMate TABE
- New Scanning option for TABE 11/12
- Uploading to a central internet database

TABE CLAS-E
- New NRS Functional Levels and CCR Standards
- Computerized Testing option
- Approved at least through 2017

Questions? - Thank you!

Mike Johnson
National Adult Assessment Manager
630-995-6712
mike.johnson@ctb.com
## Individual Profile: Johnson, Mike

**Report Criteria**

- **ID:** 513160
- **Test Name:** Tabe 9 Online Complete Battery
- **Test Finish Date:** 08-09-2007
- **Report Date:** 08-24-2007 06:39:27 PM
- **State:** Wisconsin
- **District:** MATC
- **School:** Downtown
- **Class:** Monday ABE
- **Test Scheduler:** Rose West

### Test Results

<table>
<thead>
<tr>
<th>Content Area</th>
<th>Level</th>
<th>Number of Questions</th>
<th>Total</th>
<th>Correct</th>
<th>Attempted</th>
<th>Scale Score</th>
<th>Grade Equivalent</th>
<th>National Percentile</th>
<th>Normal Curve Equivalent</th>
<th>National Stanine</th>
<th>% Objectives Mastery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applied Mathematics</td>
<td>A</td>
<td></td>
<td>50</td>
<td>12</td>
<td>12</td>
<td>423</td>
<td>3.3</td>
<td>17</td>
<td>30</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>Language</td>
<td>A</td>
<td></td>
<td>55</td>
<td>12</td>
<td>12</td>
<td>337</td>
<td>1.4</td>
<td>5</td>
<td>14</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Language Mechanics</td>
<td>A</td>
<td></td>
<td>20</td>
<td>12</td>
<td>12</td>
<td>565</td>
<td>9.5</td>
<td>75</td>
<td>64</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Math Computation</td>
<td>A</td>
<td></td>
<td>40</td>
<td>12</td>
<td>12</td>
<td>486</td>
<td>5.2</td>
<td>42</td>
<td>46</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Reading</td>
<td>A</td>
<td></td>
<td>50</td>
<td>12</td>
<td>12</td>
<td>325</td>
<td>1.6</td>
<td>5</td>
<td>16</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>Spelling</td>
<td>A</td>
<td></td>
<td>20</td>
<td>12</td>
<td>12</td>
<td>531</td>
<td>8.6</td>
<td>63</td>
<td>57</td>
<td>6</td>
<td>33</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>A</td>
<td></td>
<td>20</td>
<td>12</td>
<td>12</td>
<td>497</td>
<td>5.2</td>
<td>39</td>
<td>44</td>
<td>4</td>
<td>67</td>
</tr>
<tr>
<td><strong>Total Battery</strong></td>
<td></td>
<td></td>
<td>195</td>
<td>48</td>
<td>48</td>
<td>372</td>
<td>2.2</td>
<td>5</td>
<td>16</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>Total Mathematics</strong></td>
<td></td>
<td></td>
<td>90</td>
<td>24</td>
<td>24</td>
<td>454</td>
<td>4.4</td>
<td>26</td>
<td>37</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

### Predictive Analysis Recommendation

<table>
<thead>
<tr>
<th>Content Area</th>
<th>Predictive GED Score</th>
<th>Recommended Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>280</td>
<td>Instruct</td>
</tr>
<tr>
<td>Math</td>
<td>310</td>
<td>Instruct</td>
</tr>
<tr>
<td>Reading</td>
<td>220</td>
<td>Instruct</td>
</tr>
<tr>
<td>Science</td>
<td>230</td>
<td>Instruct</td>
</tr>
<tr>
<td>Social Studies</td>
<td>230</td>
<td>Instruct</td>
</tr>
<tr>
<td>Writing</td>
<td>290</td>
<td>Instruct</td>
</tr>
</tbody>
</table>

### NRS Levels

<table>
<thead>
<tr>
<th>Content Area</th>
<th>NRS Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>Level 1</td>
</tr>
<tr>
<td>Reading</td>
<td>Level 1</td>
</tr>
<tr>
<td>Total Mathematics</td>
<td>Level 3</td>
</tr>
</tbody>
</table>

### Performance on Objectives

<table>
<thead>
<tr>
<th>Objective</th>
<th>Number of Questions</th>
<th>Percent Correct</th>
<th>Mastery Level</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Applied Mathematics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computation in Context</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>7</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>Estimation</td>
<td>4</td>
<td>3</td>
<td>75</td>
</tr>
<tr>
<td>Geometry and Spatial Sense</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Measurement</td>
<td>6</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>Number and Number Operations</td>
<td>6</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>Patterns, Functions, Algebra</td>
<td>8</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>Problem Solving and Reasoning</td>
<td>5</td>
<td>4</td>
<td>80</td>
</tr>
<tr>
<td>Statistics and Probability</td>
<td>4</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td></td>
<td>30.1</td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capitalization</td>
<td>5</td>
<td>2</td>
<td>40</td>
</tr>
<tr>
<td>Paragraph Development</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Punctuation</td>
<td>6</td>
<td>1</td>
<td>17</td>
</tr>
</tbody>
</table>
Why TABE 9 & 10 Can Be Used Interchangeably In Your Assessment Program

The National Research Council considers assessments to be *alternate* forms if they
- measure the same constructs,
- are intended for the same purposes,
- and are administered using the same directions.

If two forms of an assessment meet these three conditions they can be used interchangeably in an assessment program. TABE meets these three conditions.

Two forms of a test do not have to be *parallel* forms to be considered alternate forms. Parallel forms have equal number correct means, equal number correct standard deviations, and equal correlations with other measures for any given population. This means that if someone took two parallel forms of a test, without any change in ability level, they would obtain the same number correct score. There would be no need to convert the number correct scores to scale scores or standard scores. Parallel forms are rare and generally do not exist in the educational testing market.

The more common scenario is for two forms of a test to be *equivalent* forms. Equivalent forms do not have equal number correct means or standard deviations, but the differences in the number correct statistics are compensated for by the conversion of the number correct scores to scale scores and derived scores, such as national percentiles and normal curve equivalents. Equivalent forms also have form-specific norms tables, which provide the number correct to scale score conversions.

Therefore, if a student takes two forms of an assessment such as TABE, it is very likely that he/she will obtain different number correct scores even if his/her ability level remains constant. This is due to the fact that one form of an assessment is almost always going to be slightly more difficult, or slightly easier, than another form. This is true for TABE. The table below compares summary statistics for TABE 9 & 10 Level M Complete Battery Language.

<table>
<thead>
<tr>
<th>Descriptive Statistic</th>
<th>TABE 9</th>
<th>TABE 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of items</td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td>Mean number correct</td>
<td>32.53</td>
<td>34.03</td>
</tr>
<tr>
<td>Mean p-value</td>
<td>.59</td>
<td>.62</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>10.46</td>
<td>10.22</td>
</tr>
<tr>
<td>Standard Error of Measurement (SEM)</td>
<td>3.14</td>
<td>3.07</td>
</tr>
<tr>
<td>KR-20 (reliability)</td>
<td>.91</td>
<td>.91</td>
</tr>
</tbody>
</table>

As table 1 shows, Form 9 is slightly more difficult than Form 10. The mean number correct is 1.5 points lower for Form 9 than Form 10, and the mean p-value (item difficulty value) for Form 9 is slightly lower than Form 10. However, the other data show how equivalent TABE 9 & 10 are—not parallel, but close. The number correct score standard deviations are very similar, the standard error of measurement (SEM) coefficients are very similar, and the KR-20 correlation coefficients are identical.

The following example illustrates how TABE 9 & 10 provide equivalent scores. Two students, Bill and Maria, are in the same Language class and have identical grades and the teacher feels they have equal ability. To confirm this, she gave Bill TABE Form 9 and Maria TABE Form 10 of the Level M Language test. Bill answered 36 items correctly and Maria answered 38 items correctly. At first the teacher thought Maria performed better on the test than Bill, but when she used her norms book to get their grade equivalents and national percentiles, she found the results presented in Table 2.

<table>
<thead>
<tr>
<th>Student</th>
<th>TABE Form</th>
<th># Correct</th>
<th>Scale Score</th>
<th>SEM</th>
<th>Grade Equiv</th>
<th>Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bill</td>
<td>9</td>
<td>36</td>
<td>513</td>
<td>17</td>
<td>5.2</td>
<td>51</td>
</tr>
<tr>
<td>Maria</td>
<td>10</td>
<td>38</td>
<td>514</td>
<td>18</td>
<td>5.3</td>
<td>51</td>
</tr>
</tbody>
</table>

Despite answering two fewer items correctly, Bill has virtually identical results in terms of scale scores, grade equivalents, and percentile scores. The reason their scale scores and GE scores are slightly different is because when two different forms of a test are scaled using number correct scaling there is simply not a scale score for every possible number correct score. For example, the scale for Form 9 Level M Language goes from 260 to 807. Since there are only 55 items on the test there are many potential scale score points that will not have a number correct score associated with it.
Locator tests are given to help determine which level of an assessment to administer to obtain the most accurate information about a student’s academic strengths and weaknesses. Locator tests, such as those for TABE 9&10, are built to measure a wide range of ability with a limited number of items for each content area. As a result, the information from a locator test about a student should be viewed as only a very rough estimate of the student’s functional level, not as an absolute prediction. As is true for all tests (and in accordance with Standard 13.7 of the Standards for Educational and Psychological Testing: AERA, APA, and NCME, 1999), any decisions about a student should not be made on the basis of a single locator test score, but should include other relevant information about the student.

That being said, it is often the case with adult students that little is known about the student’s ability level when a test such as TABE needs to be administered, so locator tests are heavily relied upon to make decisions about which level of the test to administer. Because locator tests have a limited number of items, they do not provide results that are as reliable as the main assessments, nor can the same kind of generalizations about a student’s probability of success in academic coursework be made based on the results. Therefore, locator tests should never be used in place of a main assessment such as the TABE Survey or TABE Complete Battery.

Moreover, the standard error of measurement (SEM) should be taken into account when using results from a locator test. SEM is an attribute of all tests because tests sample from a content domain, just like the results from a Gallup Poll always contain sampling error. Sampling error in Gallup Poll results is directly related to the size of the sample—the larger the sample, the lower the sampling error. The same is true for a test—SEM for a test will be lower if a larger sample of items is given. If a student’s score on a locator test is right at a cut-score boundary, SEM alone could lead to a student being identified as having more, or less, ability than he/she actually has.

For example, the recommended cut-scores for the Language Locator Test are

- 6 items correct or below administer Level E
- 7-8 items correct administer Level M
- 9-10 items correct administer Level D
- 11-12 items correct administer Level A
The SEM for the TABE Language Locator Test is 1.42, so a student could be identified as being ready for Level D if they got 9 items correct when their actual functional level meant they should have taken Level M (i.e., 9-1.42=7.58).

The recommended cut-scores and SEM values are shown in Table 1.

Table 1: Recommended TABE Locator Test Cut-Scores

<table>
<thead>
<tr>
<th>Reading</th>
<th>Mathematics</th>
<th>Language</th>
<th>TABE level to administer</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 and below*</td>
<td>4-6**</td>
<td>6 and below</td>
<td>E</td>
</tr>
<tr>
<td>7-8</td>
<td>7-8</td>
<td>7-8</td>
<td>M</td>
</tr>
<tr>
<td>9-10</td>
<td>9-11</td>
<td>9-10</td>
<td>D</td>
</tr>
<tr>
<td>11-12</td>
<td>12-16</td>
<td>11-12</td>
<td>A</td>
</tr>
<tr>
<td>SEM = 1.26</td>
<td>SEM = 1.54</td>
<td>SEM = 1.42</td>
<td></td>
</tr>
</tbody>
</table>

In order to avoid administering a TABE level that is too difficult for the student a good rule of thumb would be to administer a lower level if the student scores at the lower bound of the recommended cut-scores (e.g., if a student got 9 Mathematics Locator Test items correct, administer Level M rather than Level D—if they got 10 or 11 correct, administer Level D). Because TABE is vertically scaled across its four levels it theoretically does not matter if a student takes an adjacent level—their scale score would be the same—but taking a level of TABE that is appropriate for the student’s ability level will provide more accurate diagnostic information and will be a less frustrating experience for the student.