Online Resources for Selecting Foreign Language Assessments

March 27, 2009
Interagency Language Roundtable

Margaret E. Malone
Victoria C. Nier
Megan J. Montee

U.S. Department of Education grant #P017A050033
Purpose

- Provide an overview of online resources available to assist with foreign language test selection and use.
- Discuss opportunities and challenges in promoting language assessment literacy.
- Explore the development of an online tutorial as an example of promoting assessment literacy.
- Present interactive examples from the tutorial to illustrate real-life testing situations.
Overview

I. Language assessment literacy

II. Foreign Language Assessment Directory

III. Online tutorial: Understanding Assessment
   a. Development
   b. Content
   c. Next steps

IV. Conclusions
Language Assessment Literacy

- **Assessment literacy** = what instructors need to know about assessment (Boyles, 2005; Stiggins, 1997; Stoynoff and Chapelle, 2005)

- **Language assessment literacy** = what stakeholders need to know about language and assessment in order to conduct reliable, valid and fair assessment of all students

- **Stakeholders** = all participants including test developers and instructors
Jessica works for a department of the federal government. She is in charge of administering an oral proficiency test to prospective employees. The results of the test may influence potential hiring decisions as well as salary levels.

All applicants are over 18, most have a college education and some are native speakers of English while others are native speakers of other languages.
How do the following terms relate to the test that Jessica should choose?

- **Practicality:**
- **Reliability:**
- **Validity:**
- **Impact:**
Key Concepts

- **Practicality**: The test should be reliable and valid and meet the cost demands of the department.

- **Reliability**: The test should show giving the same rating regardless of administrator or day.

- **Validity**: The test should measure a range of oral proficiency levels and be appropriate for adults.

- **Impact**: The test should be reliable and valid, because it may affect hiring decisions.
Discussion (2)

- What language testing challenges do you face?
- What are your needs?
Challenges in Language Testing

- Increased focus on language assessment and outcomes
  - *No Child Left Behind*
  - Program evaluation

- Divide between language testing experts and language educators
  - Foreign Language Assessment Directory
  - *Survey of Assessment Uses and Needs*
  - Input from ILR testing committee- include different scenarios for different audiences
Challenges in Language Testing

- Instructors must understand language assessment in order to participate in the language learning culture (Shepard, 2000).
- Many resources are context-specific (e.g., elementary or high school; university; business, etc.)
My university is doing an evaluation of all its foreign language programs and wants data about the proficiency levels of our majors upon graduation. I need to find a test. Help!

- What resources are available?
Foreign Language Assessment Directory

http://www.cal.org/CALWebDB/FLAD/

- Purpose: Provide educators with information about available foreign language assessments
  - Information on nearly 200 tests in over 90 languages other than English
The Foreign Language Assessment Directory (FLAD) is a free, searchable database with information on more than 200 assessments in over 90 languages other than English. FLAD contains information about assessments currently used in elementary, middle, secondary, and post-secondary school programs around the United States.

Please note, this resource contains information about various assessments. It does not contain actual assessments or sample assessments. This information was provided by test developers or administrators. It does not necessarily reflect the views of CAL staff members and the quality of these tests has not been evaluated by CAL.

### Search instructions

You may search the database for specific foreign language assessment instruments that fit the criteria below, or you may browse through the entire list by clicking the "Search" button without specifying any search criteria.

You may choose all, several, or one category to conduct a search. **Using fewer categories will usually generate a longer list of results.**

<table>
<thead>
<tr>
<th>Name of Test:</th>
<th>All Tests...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language:</td>
<td>All Languages...</td>
</tr>
<tr>
<td>Grade Level(s):</td>
<td>All Grade Levels...</td>
</tr>
<tr>
<td>Proficiency Level(s):</td>
<td>All Proficiency Levels...</td>
</tr>
<tr>
<td>Intended Test Use:</td>
<td>All Intended Test Uses...</td>
</tr>
<tr>
<td>Skill Tested:</td>
<td>All Skills Tested...</td>
</tr>
</tbody>
</table>

Search | Reset

**Submit a new test**

Test developers and administrators may use this form to let CAL know about new tests to be listed in the directory.

### Resources and Links

Be sure to check out the internet and publications resources on CAL’s website. For more information on this directory and related projects, go to the FLAD project page.
Foreign Language Assessment Directory

- **FLAD** search criteria
  - Name of test
  - Language
  - Grade/age level(s)
  - Proficiency level(s)
  - Intended test use
  - Skill tested
# Foreign Language Assessment Directory (FLAD)

**Search Results:**

Your search for: **Chinese (Mandarin)** 9-12 All Proficiency Levels... All Intended Test Uses... All Skills Tested... returned 10 record(s).

To sort by language, intended use or grade level, click the appropriate header. For more information about a specific test, click on the name of the test.

<table>
<thead>
<tr>
<th>Name of Test</th>
<th>Languages</th>
<th>Intended Test Use</th>
<th>Grade Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTFL Oral Proficiency Interview (OPI)</td>
<td>Afrikaans, Albanian, Amharic, Arabic, Armenian, Bulgarian, Cebuano, Chavacano, Chinese (Cantonese), (...)</td>
<td>Placement</td>
<td>10, 11, 12, College, Adult</td>
</tr>
<tr>
<td>Advanced Placement Examinations [AP]</td>
<td>Chinese (Mandarin), French, German, Italian, Japanese, Latin, Spanish</td>
<td>Credit by Examination, Placement</td>
<td>9, 10, 11, 12</td>
</tr>
<tr>
<td>Australian Language Certificate - Beginner's Level</td>
<td>Chinese (Mandarin), French, German, Indonesian, Italian, Japanese</td>
<td>Diagnostic, Proficiency</td>
<td>8, 9, 10</td>
</tr>
<tr>
<td>Cambridge Step by Step Assessment to Language Dominance (SSALD)</td>
<td>Chinese (Mandarin), Haitian Creole, Korean, Portuguese, Spanish</td>
<td>Language Dominance</td>
<td>K, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12</td>
</tr>
<tr>
<td>Chinese Proficiency Test (CPT)</td>
<td>Chinese (Mandarin)</td>
<td>Proficiency</td>
<td>9, 10, 11, 12, College, Adult</td>
</tr>
<tr>
<td>Chinese Speaking Test (CST)</td>
<td>Chinese (Mandarin)</td>
<td>Proficiency</td>
<td>9, 10, 11, 12, College, Adult</td>
</tr>
<tr>
<td>Foreign Language Achievement Tests (FLATS)</td>
<td>Afrikaans, Albanian, Arabic, Armenian, Aymara, Bulgarian, Cakchiquel, Cambodian, Cebuano, (...)</td>
<td>Credit by Examination</td>
<td>12, College, Adult</td>
</tr>
<tr>
<td>Preliminary Chinese Proficiency Test (Pre-CTP)</td>
<td>Chinese (Mandarin)</td>
<td>Proficiency</td>
<td>9, 10, 11, 12, College, Adult</td>
</tr>
<tr>
<td>SAT Subject Tests</td>
<td>Chinese (Mandarin), French, German, Hebrew, Italian, Japanese, Korean, Latin, Spanish</td>
<td>Admission, Placement</td>
<td>9, 10, 11, 12</td>
</tr>
<tr>
<td>Standards-Based Measurement of Proficiency (STAMP)</td>
<td>Chinese (Mandarin), French, Italian, Japanese, Spanish</td>
<td>Admission, Fulfilling a Requirement, Placement, Proficiency, Program Evaluation, Research</td>
<td>8, 9, 10, 11, 12, College</td>
</tr>
<tr>
<td>Name of Test</td>
<td>Chinese Speaking Test (CST)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Languages</td>
<td>Chinese (Mandarin)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Description</td>
<td>The CST is a simulated oral proficiency interview that measures an examinee’s speaking ability in Chinese. It is intended for students who perform at a Novice-High to Superior level on the ACTFL scale. This tape-mediated proficiency-based performance assessment requires students to listen to directions to speaking tasks from a master tape while following along in a test booklet. Student responses to each task are recorded on a separate response tape that is later evaluated by a trained rater who scores the performance according to the proficiency guidelines developed by the American Council on the Teaching of Foreign Languages.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade Level(s)</td>
<td>9, 10, 11, 12, College, Adult</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proficiency Level(s)</td>
<td>ACTFL Intermediate, ACTFL Advanced, ACTFL Superior</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intended Test Use</td>
<td>Proficiency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skills Tested</td>
<td>Speaking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Author/Publisher</td>
<td>University of Michigan/Center for Applied Linguistics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publication Date</td>
<td>1986, revised 1994</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Cost</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Length</td>
<td>25 to 50 minutes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Materials</td>
<td>Test booklet, master test tape (audio), examinee response tape (audio); test manual and examinee handbook</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Format</td>
<td>Simulated Oral Proficiency Interview</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scoring Method</td>
<td>Holistic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scoring Process</td>
<td>Central trained rater scored</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Results Reported</td>
<td>Proficiency Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administered By</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administered When</td>
<td>On request</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administered How</td>
<td>Individually, Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administered Where</td>
<td>On request</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contact Information</td>
<td>+ Show Contact Information</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Test Entry Information

- Selected information provided by the FLAD
  - Test description
  - Grade/age level(s)
  - Proficiency level(s)
  - Intended test use
  - Skills tested
  - Scoring information
  - Cost
  - Administration information (length, resources needed)
  - Publisher contact information
New Feature: Moderated User Review

What is your position/role?
Other

For how many students did you use this test?
-- Select one --

For which age group of students did you use this test?
-- Select one --

If you chose Pre-K, K-3, or 4-6, with which type of program do you work?
-- Select one --

For which language did you use this test?

For which purpose did you use this test?
-- Select one --

Please rate your level of agreement with the following statements:

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>This test provided me with the information that I needed.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>The instructions for administration were clear.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>This test took a reasonable amount of time to administer and score.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>This test was cost-effective.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>The score reporting was understandable.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>This test was appropriate for my students.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>My students thought this test was fair.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>I would use this test again.</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Comments:
How many of you have ever had to select a language test?
- What are some factors that influenced the test you chose?
- Was the test a good fit for your purposes and the population you were testing?

How many of you have had to use a language test that someone else selected?
- Was the test a good fit for your purposes and the population you were testing?
- What are some factors that made it a good fit or not?
Scenario

- My agency needs to find a test to determine proficiency for hiring. I found two tests. Which one should I use? Help!
Understanding Assessment

- free, web-based tutorial
Purposes: Online Tutorial

- Purposes:
  - Support the *Foreign Language Assessment Directory* by:
    - providing basic information on language assessment; and
    - assisting language educators in selecting assessments from the *FLAD*. 
Tutorial Audience

- Target Audience
  - Foreign language educators
    - K-12
    - Post-secondary
    - Government
    - Business
    - Non-profits
    - Other
Tutorial Development Process

1) Conduct background research
2) Draft online tutorial
3) Review online tutorial
   - Cognitive labs
   - Modified focus groups: ILR Testing Committee
4) Revise online tutorial based on feedback
5) Make tutorial available on CAL’s website (November, 2008)
Focus groups with foreign language educators
  • 6 sessions; 41 participants
  • PK-12, university and government foreign language professionals

Live assessment workshops

Previous CAL projects
  • *Survey of Assessment Uses and Needs*
  • Rater training
What do language educators need to know?

- Lessons learned:
  - Need for more language assessment training
  - Need for connections between assessment and classroom instruction
    - Test selection
    - Score reporting and interpretation
  - Mismatch between tests’ intended purposes and ways they are being used
Tutorial Outline

- Home page
- Module 1: Introduction
- Module 2: Practicality
- Module 3: Reliability
- Module 4: Validity
- Module 5: Impact
- Module 6: Putting It Together
- Module 7: Resources
What is Understanding Assessment?

Understanding Assessment: A Guide for Foreign Language Educators is a tutorial on language testing for foreign language educators. It is designed to help new users of the Center for Applied Linguistics' Foreign Language Assessment Directory (FLAD) and it is also an excellent resource for anyone who wants to learn more about language testing.

How should I use this resource?

- Use the buttons on the left to go through the modules in order.
- This tutorial is divided into 7 modules, most of which have an introduction page and 3 subsections.

1. Introduction
   - Key concepts
   - Test selection
   - Example
2. Practicality
3. Reliability
4. Validity
5. Impact
6. Putting It Together
7. Resources
Module 1: Introduction to Test Selection
(screen 1 of 4)

Introductory questions

1. Where do I start when trying to choose a test?
2. What do I need to consider in selecting a test for my students?
3. How will I know that a test works for me?

An introduction to the principles of test selection

Module 1 introduces the basics of test selection and outlines some essential vocabulary, including practicality, reliability, validity, and impact. The module is an introduction to the tutorial and to key concepts in language testing.

Objectives

- Identify factors in test selection.
- Define basic assessment concepts, including practicality, reliability, validity, and impact.
Module 1: Introduction to Test Selection

Key Concepts

When selecting a test, you need to think about practicality, reliability, validity, and impact.

Practicality
Practical elements of test selection involve thinking about resources such as time, funds, support and expertise, and how they relate to your students and your program. You should choose a test based on your resources. All in all, it is important to initiate a testing program that you are able to maintain for a number of years.

Reliability
Reliability refers to the consistency of test results. Will a student get almost the same score on a test taken Monday and again on Tuesday? Will one teacher score the assessment the same way as another? If so, then the test and the scoring are reliable. Reliability is closely tied to validity. A test can be reliable without being valid. However, a test cannot be valid unless it is reliable.
Now that you have an introduction to some basic testing concepts, it's time to put that knowledge to use.

Below are some scenarios that ask you to use your language testing knowledge in a real-life situation. Pick one:

1. **Elementary education**
2. **Secondary and post-secondary education**
3. **Government**
4. **Business**
5. **Non-profit**
Jessica works for a department of the federal government. She is in charge of administering an oral proficiency test to prospective employees. The results of the test may influence potential hiring decisions as well as salary levels. All applicants are over 18, most have a college education and some are native speakers of English while others are native speakers of other languages.

Get out a sheet of paper and brainstorm one example each for how the following terms influence the language test that Jessica administers:

<table>
<thead>
<tr>
<th>term</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>validity</td>
<td></td>
</tr>
<tr>
<td>reliability</td>
<td></td>
</tr>
<tr>
<td>practicality</td>
<td></td>
</tr>
<tr>
<td>impact</td>
<td></td>
</tr>
</tbody>
</table>

The test should...

**Validity:** measure oral language abilities, measure a range of proficiency levels, be appropriate for adults

**Reliability:** give the same score no matter who takes it or what form is taken

**Practicality:** be reliable and valid while meeting the cost demands of the department

**Impact:** be reliable and valid because its high-stakes nature may affect hiring decisions and salary ranges
Module 2: Practicality (Screen 1 of 4)

**Introductory questions**

1. What funds are available for testing?
2. How much do my colleagues know about language testing?
3. How much time is there for assessment?

**Considering the practical issues**

Module 2 addresses practical issues relevant to test selection, including resources, classroom characteristics, and methods of administration and scoring. You will gain an understanding of resources and challenges that factor into planning and implementing a successful testing program.

**Objectives**

- Consider how available resources of time, money, and expertise relate to decisions about testing.
- Explain processes of administration and scoring.

In the end, you must be realistic about the resources available to put a test to use. This will likely help you gain the support of your colleagues and avoid wasted effort. Such concerns include whether available resources will cover the demands of test administration and scoring in terms of time, money, expertise, and personnel.
Module 2: Practicality (screen 2 of 4)
Methods of Testing

Language tests offer you many choices in test administration, test format, materials, scoring method, and test items. Which options best fit your students, your resources, and your reasons for testing?

Let’s begin by exploring some of the aspects of testing methods. Read the list below of factors to consider when thinking about test methods. Then click here or on the expanded checklist for the Test Methods Checklist.

1. Test administration
2. Input
3. Scoring process
4. Scoring method
5. Response type
## Test Methods Checklist

### Considerations for Test Administration

1. **Giving the test.** Is it...
   - [ ] Individually-administered?
   - [ ] Group-administered?

This is important because it is related to resources such as time, staff, and expertise needed.

2. **Directions, examples and questions and prompts.** Are they...
   - [ ] In English?
   - [ ] In the target language?
   - [ ] Spoken?
   - [ ] Written?

This is important because it is related to validity and practicality. For example can the student understand the directions if they’re in the target language? Are you sure you’re testing what you think you’re testing rather than their understanding of the directions?

3. **Student answers and responses.** Are they...
   - [ ] In English?
   - [ ] In the target language?
   - [ ] Spoken?
   - [ ] Written?

This is important because it is related to resources such as time, to scoring, and to students’ strengths in various modalities.

### Considerations for Scoring Tests

4. **Method of scoring.** Does it involve...
   - [ ] Counting the number correct?
   - [ ] A holistic approach (one overall score)?
   - [ ] An analytic approach (separate scores for separate aspects of the performance)?

This is important because it’s related to what you’re testing (e.g., it’s harder to count number correct with writing!), resources (e.g., do you have time to score each writing piece analytically?), and to how you intend to use the test (e.g., if you want to provide specific feedback to students, then an analytic rubric might be useful).
Module 2: Practicality

Testing Resources

When determining which test is best for you, you should think about staff members’

- Time
  - test administration
  - scoring
- # of students
- # of staff
- Staff expertise
  - knowledge of students
  - knowledge of test
- Resources
  - Other
    - proctors
    - raters
    - rubrics
    - security
  - Funding
  - Materials
    - language lab
    - special equipment
    - computers
What resources do you have at your disposal for testing?
If the scale is reliable it tells you the same weight every time you step on it as long as your weight has not actually changed. However, if the scale is not working properly, this number may not be your actual weight. If that is the case, this is an example of a scale that is reliable, or consistent, but not valid. For the scale to be valid and reliable, not only does it need to tell you the same weight every time you step on the scale, but it also has to measure your actual weight.

Switching back to testing, the situation is essentially the same. A test can be reliable, meaning that the test-takers will get the same score no matter when or where they take it, within reason of course. But that doesn’t mean that it is valid or measuring what it is supposed to measure. A test can be reliable without being valid. However, a test cannot be valid unless it is reliable.
Module 3: Reliability (screen 3 of 4)
Reliability and Test Selection

**Reliability of student performance**
One student should be able to take a test on Monday and then again on Tuesday and get very similar results each day.

**Reliability of scoring**
The results should be the same no matter who scores the test and when they score it.

**Reliability between different forms of the same test**
Different forms of a test have slightly different questions in a slightly different order. However, the content and difficulty level are the same and a student should perform comparably on both tests.

Before you select a test, you’ll want to think about your program, your needs, and the resources you have. Is this a high-stakes test for the student? If so, it is especially important that the test give consistent results each time. Do students often contest their results? Then it may be especially important to find a test whose reliability information has been published in the test manual or professional journal articles. Do you have very few staff for scoring or even double-scoring the same test? Then you may want to choose a test with objective responses that can be scored quickly and reliably by a person or machine.
Module 4: Validity (screen 1 of 4)

Introduction questions

1. What do I want to know about my students?
2. What types of tests are available?
3. Which language skills do I want to test?
4. How can I use the test results?

An introduction to the principles of test selection

Module 4 focuses on purposes for testing and possible uses for test results. You will take an inventory of assessment needs and identify test types appropriate for different situations.

Objectives

- Identify various purposes for testing.
- Understand how the results of tests can be used.
A basic decision that many face is whether to test achievement or proficiency. **Achievement tests** generally measure specific content covered in a course or program, while **proficiency tests** measure general language ability and are not tied to the content of a particular course or program.
Pramukta works for his city school system. He is searching for a test to use as part of a program evaluation for the foreign language classes in the public high schools.

What should Pramukta think about as he searches for a test?

<table>
<thead>
<tr>
<th>Questions</th>
<th>Possible Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What information does Pramukta want to know?</td>
<td></td>
</tr>
<tr>
<td>2. What decisions might be based on this information?</td>
<td></td>
</tr>
</tbody>
</table>
Pramukta is searching for a test to use as part of a program evaluation. What should Pramukta think about as he searches for a test?

<table>
<thead>
<tr>
<th>Questions</th>
<th>Possible Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What information does Pramukta want to know?</td>
<td></td>
</tr>
<tr>
<td>2. What decisions might be based on this information?</td>
<td></td>
</tr>
</tbody>
</table>
## Questions

1. What information does Pramukta want to know?
   - Proficiency levels
   - Achievement of targeted skills

2. What decisions might be based on this information?
   - Revisions to the program in the future
   - Feedback to stakeholders
   - Funding for the future
Module 5: Impact (screen 1 of 4)

Introductory questions

1. How can I make sure tests are fair for each student?
2. How can I use test results to inform instruction?
3. How can I prepare my students for required tests?
4. How can I use tests positively with my students?

Teaching and testing

Module 5 explores the possible effects of tests and focuses on the relationship between assessment and instruction. You will consider the idea of washback and think about how instruction and assessment can implement one another in the classroom.

Objectives

- Compare the possible effects of tests on various stakeholders.
- Define negative and positive washback and plan ways to promote positive washback.
- Incorporate assessment effectively into the instructional process.

Impact

- Washback and instruction
- Stakeholders
- Example

Tests that accurately measure the skills you want to know about and that are used appropriately can have positive results for classroom instruction. A good way to promote positive washback is by keeping assessment in mind as you begin planning for instruction.
- Would be very high-stakes for the test takers. Needs to be reliable.

- Costs of administering and scoring the test? Does someone at the agency need to be trained to administer and score the test?

- Time constraints. Test would need to be taken and scored in a short time frame for the hiring process.

- Validity is important. Need to make sure that the test measures skills that are actually needed on the job.
Module 6: Putting It Together (screen 1 of 4)

Introductory questions

1. Do I have the necessary resources for the test I select?
2. Will this test give me the information I need about my students?
3. What type of test do I need most?
4. How can I convince my colleagues that a particular test is right for our students?

Putting it all together

Throughout this tutorial, you’ve learned about practicality, reliability, validity and impact. Module 6 furthers your understanding of how to select the best test for your students.

Objectives

- Explain connections between validity and test taker characteristics, purposes for testing and effects of testing.
- Define testing needs clearly.
- Consider the roles of practicality, reliability, validity and impact in selecting a particular test.

Putting It Together

- Needs assessment
- Test selection
- Example
<table>
<thead>
<tr>
<th>Class Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target language</td>
</tr>
<tr>
<td>Grade level and age range of students</td>
</tr>
<tr>
<td>Proficiency level(s)</td>
</tr>
<tr>
<td>Number of students in class or program</td>
</tr>
<tr>
<td>Amount of time in class</td>
</tr>
<tr>
<td>Program description</td>
</tr>
<tr>
<td>Classroom management</td>
</tr>
</tbody>
</table>
# Test Selection Checklist

**Directions:**
Use the checklist below to be sure you have made a principled selection or to serve as a springboard for dialogue.

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. ☐ ☐ I have the necessary resources to administer the assessment.

2. ☐ ☐ I have the necessary resources to score the assessment.

3. ☐ ☐ The assessment’s intended purpose matches my purpose for assessing.

4. ☐ ☐ The assessment is appropriate for the language skills I want to assess.

5. ☐ ☐ The assessment is appropriate for the age and proficiency level of my student population.

6. ☐ ☐ The assessment results will be used as intended by the developers.

7. ☐ ☐ I have thought through the possible impact of the test and I think it will have the effects I want for students, teachers, the school and the program.
Consider the questions in the chart below:

<table>
<thead>
<tr>
<th>Needs Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who are you testing?</td>
</tr>
<tr>
<td>What skill(s) are you testing?</td>
</tr>
<tr>
<td>What do you want to find out by testing?</td>
</tr>
<tr>
<td>What will you do with the information you find out from testing?</td>
</tr>
<tr>
<td>How will testing impact your population? Your teaching? Your program?</td>
</tr>
</tbody>
</table>
Module 7: Resources (screen 1 of 5)

Introduction questions

1. Where can I find more information?
2. How can I learn more about a specific test?
3. What testing resources are available?

Test selection resources

Module 7 introduces resources you can use to gain more information about language assessments. This module will also help you consider principles for good testing practice.

Objectives

- Locate resources on specific language assessments
- Locate general resources on language assessment

Resources

- Online resources
- Print resources
- Key assessment terms
- Forms

Focus on........ Resources

There are various print and electronic resources which you can consult for more information on assessment. These can be useful if you want more help selecting tests or if you just want to learn more information about language testing in general.
Next Steps

Continual revision of resources

Continual use of Moderated User Review
Conclusions

The tutorial provides users with:

- basic information necessary for selecting and implementing language assessment;
- a chance to define their own testing needs; and
- resources about assessment.

Importance of developing evidence-based resources for training and professional development
Questions
Thank you!

Margaret Malone
mmalone@cal.org

Victoria C. Nier
vnier@cal.org

Megan J. Montee
mmontee@cal.org
References


