

**Testing – Standardized and  
Otherwise – and the Massachusetts  
ABE Pilot Test:**

**A Curriculum Unit for ABE Classes**

Written by Jenny Lee Utech (Massachusetts Worker  
Education Roundtable), with Ruth Byrne and Susan Prior  
(Methuen Adult Learning Center),  
for the Massachusetts Department of Education's  
Adult and Community Learning Services, March 2005

## Testing in General, Standardized Testing, and the Massachusetts ABE Pilot Test: A Curriculum Unit for ABE Classes

Written by Jenny Lee Utech (Massachusetts Worker Education Roundtable), with Ruth Byrne and Susan Prior (Methuen Adult Learning Center), for the Massachusetts Department of Education's Adult and Community Learning Services, March 2005.

### Introduction

#### Why use this curriculum?

We hope that this curriculum will help you and your students understand the reasons behind the upcoming pilot testing of the ABE tests for reading and math, to prepare for and take the pilot test (and other standardized tests), and give important feedback during the pilot testing process. But we also see this curriculum and the pilot testing process as opportunities to help students explore testing and their experiences with it, and develop and refine strategies and life skills they can use in many situations.

Each lesson in the unit is designed to accomplish specific objectives that will help you and your students answer these questions:

- Why are students being tested? What is the **purpose of tests** (in general)? And what is the **purpose of standardized testing**? What are some **pros** of standardized testing? What are some **cons**?
- How can we empower students in the test-taking process? What test-taking and test-preparation **strategies** can we help students develop?
- How can we introduce the **computer tutorial** in a meaningful way?
- What is the role of **pilot testing**? How does pilot testing help develop a fair and consistent test?
- How can we collect **feedback** on the pilot test from teachers and students?

#### How to use this curriculum

We hope that you will use as much of this curriculum as possible. We understand that class time is limited and that you are being asked to pilot the new test in addition to everything else you must do in your classes from week to week. We hope that you will be able, at a minimum, to use enough of this unit so that you and your students can address the questions above adequately.

We have designed lessons and activities that satisfy a range of student skill levels, and that include a combination of brainstorming and discussion, reading, and writing. While lessons provide important language work (vocabulary review, reading comprehension questions, writing exercises), the most important work in each lesson is to elicit and build on students' experiences with testing, help students understand

why the pilot testing is happening and how they can prepare for it, and help students develop strategies for approaching testing in general. If you don't have time for all of the reading and vocabulary work, we hope that you will have time to use the discussion questions in each lesson, and the handouts on pilot testing and strategies.

Teacher notes after each lesson list key ideas that the lesson is designed to cover, as well as activity tips and variations. You should adapt lessons and handouts to meet the needs and levels of your students. While most handouts are at fairly low reading levels, they are designed to be used at all levels to convey key information effectively and promote critical thinking.

Approximate reading levels for this curriculum's stories are provided in the teacher notes following story activities. If the stories or other handouts are at too high a reading level for your class, you can adapt and use them in many ways. For example, you can read a story out loud for students and ask them to tell it back to you orally while you write up key sentences. You can divide students in to small groups and ask each group to read part of a story, or enlarge text on the photocopier, cut it up and ask each student to work with one sentence of the story. Or, you can simply read the story out loud and discuss it orally, using it as a catalyst but not for reading or writing practice.

The time it takes your class to complete any one activity will vary according to class levels and how you choose to do the activity (whole class, small groups, pairs). You may choose to condense a lesson into 30 minutes, or use all the suggested activities to create a 1 or 2-hour lesson.

Lessons build on each other and are designed to be used sequentially. If your time is very limited, we suggest that you focus on the following lessons and activities:

- Either Lesson 1, step 1 (discussion about students' experiences with testing), or Lesson 2, parts 1 and 2 (reading about experiences with tests and discussion about students' experiences with testing).
- Lesson 3, parts 1 and 2 (defining standardized testing, identifying pros and cons)
- Lesson 4, part 1 (identifying strategies for testing-preparation and test-taking)
- Lesson 5 (examining the definition of pilot testing)

## Curriculum outline

**Lesson 1. Experiences with Tests** – Students discuss their experiences with taking tests (including standardized testing) and write about their experiences.

**Lesson 2. Stories about Tests** – Students read stories about other adult students' experiences with taking tests, and relate them to their own experiences.

**Lesson 3. Standardized Tests** – Students examine and discuss the definition of standardized testing, and identify the purposes, pros and cons of standardized testing.

**Lesson 4. Test-Taking Strategies** – Students identify test-preparation and test-taking strategies, and create their own lists of test-preparation and test-taking strategies to use now and in the future. After the pilot test, students can evaluate and revise these strategies.

**Lesson 5. Pilot Testing** – Students examine and discuss the definition of pilot testing, discuss the new ABE test that will be piloted in Massachusetts, and discuss how they will participate in the pilot.

**Lesson 6. Pilot Test Follow Up** – Students discuss and evaluate their experiences with the pilot test, and evaluate and revise their test-preparation and test-taking strategies.

**Many thanks** to Jane Schwerdtfeger (Mass. DOE/ ACLS) for guiding the conceptualization of this curriculum and providing insightful feedback on lesson drafts.

## Lesson 1. Experiences with Tests

### Objective

- Discuss experiences with taking tests (standardized tests and all other kinds).

### Materials

- Flip chart paper, magic markers and masking tape.

### Activities

1. Lead a whole-class discussion using the questions below. For some classes, it may work well to post the questions on flip chart paper and record students' answers. For other classes, you may want to ask students to discuss the questions in small groups first, and then share their ideas with the whole class:
  - What is your **experience** with taking tests? (What was your experience as a child? As an adult / now?) What kinds of tests have you taken? Why?
  - How did you **feel** when you took those tests (or that one particular test)? (Or, how does taking a test in general make you feel? How do you feel when you are taking a test?) Why? (See Teacher Notes.)
  - **Why** do we take tests? (What are the purposes of taking a test?)
2. Ask students to write a story about a test experience (standardized or any other, adulthood or childhood). You can use one or more of the following questions as writing prompts:
  - Write about a time you took a test. Why did you take the test? How did you feel before the test? During? After? How did you do on the test? Why?

### Teacher notes

**Key ideas.** For the brainstorm of students' test-taking experiences, encourage students to think about and name different types of tests they have taken. These might include: tests they took in school while growing up (anything from spelling or math tests to standardized tests like the SSAT, SAT, MCAS, or Regents Exams, or standardized tests in their countries; tests they've had in adult education programs and in your class (teacher-made tests and quizzes, classroom review or other activities, or standardized tests for adult education like TABE, BEST Plus, REEP, GED); driver's license test; citizenship test; work-related tests here or in students' native countries (job-placement tests, typing or other skill-related tests, "personality" tests).

Students will have a chance to talk about standardized tests in more depth in the lessons that follow. For your discussion in this lesson, ask students to name experiences with any tests they've taken in general, standardized or not (a quick brainstorm) and ask about their experiences with these tests.

Students may **feel** differently about tests they or their children taken: proud, anxious, confident, fearful, happy, resigned, smart, angry or slighted, relieved, bored, eager to show what they can do, confused, neutral, lost, challenged in a positive way, resentful, or disappointed. If students have negative feelings about taking tests, they may be reluctant to admit them. In some cases, it may help to point out that you (the teacher) have felt anxious or scared about taking a test, and that many people do.

Some **purposes** of taking tests include: tests help to evaluate what skills students have learned, and what skills they still need to work on. Tests help the teacher to evaluate how effective her / his teaching has been, and to revise or change instruction and class content where needed. Some tests have higher-stakes purposes, for example determining who gets a job, who places into college or other education programs, or who graduates from high school or earns a GED. Other tests have lower-stakes purposes, for example checking comprehension of specific content learned in a particular lesson.

**Activity tips and variations.** If you choose to record students' ideas during the brainstorm, you can type them up and use them in future classes for language work. For example, ask students to read back their own comments, respond to other students' comments in writing, or fill in missing words. For step 1 above, you can also do "Language Experience Approach" (LEA) writing with the whole class as you move through the discussion questions.

Instead of writing stories, students can write original sentences using the words or phrases the class has generated about their test-taking experiences (proud, anxious, confident). You can also ask students to complete sentences, for example, "When I take a test, I feel \_\_\_ because \_\_\_." Or, "The time I took the \_\_\_ test, I felt \_\_\_." Or, "I took a \_\_\_ test so that I could / in order to \_\_\_." You can write key words from your discussion on cards for sight word reading practice and vocabulary pair work. You can type up students' stories or sentences and use them for language work in future classes.

## Lesson 2. Stories About Tests

### Objective

- Read about people's experiences with tests and relate them to our own experiences.

### Materials

- Flip chart paper, magic markers and masking tape.
- Handouts 1a-1d. Stories about Tests
- Handouts 2a-2d. Yes / No and Questions about Stories

### Activities

#### Part 1 - Reading stories about tests

1. Choose one or more of the stories about tests (Handouts 1a-1d). Tell students that the class will read a story (or stories) about people's experiences with tests. Review any vocabulary from the story(s) if necessary.
2. Read the story(s) together with students. Let students read silently for themselves first, and let them ask any questions about words or phrases they don't understand. Then practice reading the story with students. (See Teacher Notes.)
3. Ask students some quick questions orally to make sure they have understood story content. (For example, for Magaly's Story, "What was Magaly's experience in school growing up?" "How did she prepare for tests?" "How did she feel about taking tests?")
4. Ask students to complete Handouts 2a-2d (one or both parts) in class or at home. You can also create other language practice activities. (See Teacher Notes.)

#### Part 2 - Discussion

1. If you have not done Lesson 1, lead a discussion about students' experiences with tests using the questions below (from Lesson 1) to guide your discussion. For some classes, it may work well to post the questions on flip chart paper and record students' answers. For other classes, you may want to ask students to discuss the questions in small groups first, and then share their ideas with the whole class:
  - What was [story character's] experience with taking tests? Why?
  - What is your **experience** with taking tests? (What was your experience as a child? As an adult / now?) What kinds of tests have you taken? Why?

- How did you **feel** when you took those tests (or that one particular test)? (Or, how does taking a test in general make you feel? How do you feel when you are taking a test?) Why?
  - **Why** do we take tests? (What are the purposes of taking a test?)
2. If you have not done Lesson 1, you can ask students to write a story about a test experience (standardized or other, adulthood or childhood). You can use one or more of the following questions as writing prompts:
- Write about a time you took a test. Why did you take the test? How did you feel before the test? During? After? How did you do on the test? Why?

### Teacher notes

**Key ideas.** While the stories in Handouts 1a-1d provide valuable language practice (especially for lower-level readers), they are mainly intended as catalysts, to help students start thinking and talking about their own experiences with tests. For higher-level readers, the stories can be springboards for higher-level language work (responding to the story in writing, writing your own story). Lower-level readers will need more time to work through the story texts. See Lesson 1 Teacher Notes for key ideas for your Part 2 discussion.

**Activity tips and variations.** You may want to choose at least two stories, one that depicts a positive experience and one that depicts a negative experience. Depending on your class, you may want to start this unit with one or more of the catalyst stories (instead of starting out with the discussion activity in Lesson 1). If you do Lesson 1 first, then these stories can serve as further examples of people's experiences with tests, and may help students to write their own stories as well (if they haven't already).

The approximate readability levels for Handouts 1a-1d are 4.0, 4.4, 4.8 and 3.0, respectively (Flesch Kincaid). The first story is excerpted and adapted from an interview with a Guatemalan woman by Jenny Utech, the second and third were written by Jenny Utech, and the fourth is adapted from part of a student story in Rothschild, B and Willoughby, S., eds. (1999). *Proclamation of Independence: A Book of Life Stories Told by Women Students in the Second Start Adult Literacy Program*. Oakland, CA: Second Start Adult Literacy Program.

To practice reading the story(s) with students, you can read out loud for the whole class, ask the whole class to read out loud with you, have students read to each other in pairs, or do whatever works for your class. In addition to, or instead of Handouts 2a-2d (especially for higher-level readers / writers), you can provide further language practice by asking students to write their own questions for the story(s), share them and make a class list of questions, and then write answers to some or all of the

*Testing – Standardized and Otherwise – and the Massachusetts ABE Pilot Test: A Curriculum Unit for ABE Classes.* By Jenny Lee Utech (Massachusetts Worker Education Roundtable), with Ruth Byrne and Susan Prior (Methuen Adult Learning Center), for the Massachusetts Department of Education ACLS, March 2005.

questions. You can have students write letters to the story character(s). You can create a cloze exercise, or cut the story into strips, pair up students and ask them to order the strips. You can ask students to complete sentences, for example, “When I was growing up, I liked / didn’t like taking tests because...” Now, I like / don’t like taking tests because...”). You can create present / past tense exercises using sentences or verbs from the stories.

If Handouts 1a-1d are at too high a reading level for your class, you can adapt and use them in many ways. For example, you can read a story out loud for students and ask them to tell it back to you orally while you write up key sentences. You can divide students in to small groups and ask each group to read part of a story, or enlarge text on the photocopier, cut it up and ask each student to work with one sentence of the story. Or, you can simply read the story out loud and discuss it orally, using it as a catalyst but not for reading or writing practice.

For activity tips for Part 2 (discussion), see the Teacher Notes for Lesson 1.

## Lesson 3. Standardized Tests

### Objectives

- Examine and discuss the definition of standardized testing.
- Identify the purposes, pros and cons of standardized testing.

### Materials

- Flip chart paper, magic markers and masking tape.
- Definition of standardized tests from Handout 3 written on flip chart paper or board (or copied onto an Overhead sheet).
- Handout 3. What are standardized tests?
- Handout 4a. Why use standardized tests?
- Handout 4b. Why use the GED?
- Handout 5a. Word Splash – Why do we take tests?
- Handout 5b. Word List – Why do we take tests?
- Handout 5c. Story – Why do we take tests?
- Handout 5d. Comprehension Questions – Why do we take tests?
- Handout 5e. Jigsaw

### Activities

#### Part 1 – What are standardized tests?

1. Post the following definition of standardized tests (from Handout 3) and give students Handout 3. Review the definition with your class. As you explain the definition, refer back students' ideas from the previous activity, and give your own examples:

If a test is **standardized**, it means that:

- Everyone gets the **same directions** for taking the test.
- Everyone takes the test under the **same conditions** (same amount of time, in a quiet room).
- Everyone gets the **same test content** (or same type of content).
- Everyone's test is **scored the same way**.

The **goal** of a standardized test is to make the test experience the same for everyone who takes it, so that no one has an advantage over anyone else.

2. Ask students these questions about any standardized tests they've taken (especially any they've taken more than once, for example the TABE or BEST Plus). (Or, you can ask what experiences students' children, siblings, or friends have had taking the MCAS.) This will help them connect their experiences to the definition:

- What standardized tests have you taken?
- How were directions given and what were they?
- What were the test-taking conditions (amount of time, help given or not, etc.)?
- Why can't the teacher help you answer the questions during a standardized test?
- What was the test content? How did the content compare each time you took the test?
- How was your test scored? What did it tell you about what you did well on and what you need to work on?

## Part 2 – Purposes, pros and cons of standardized tests

1. Post and ask these questions and record students' ideas on flip chart paper:
  - **Why** do we have to take tests? (Tests in general)
  - **Why** do we have to take **standardized** tests?
  - What are some **pros (advantages)** of standardized tests?
  - What are some **cons (disadvantages)** of standardized tests?
2. After students have shared their ideas, review the class's brainstormed lists and add any key ideas you think are missing.
3. Choose one standardized test that your students have some experience with (based on your class discussions), or ask the class to choose. Explain that the class will discuss the purpose, pros and cons of this particular standardized test.
4. Divide the class into small groups and give everyone Handout 4. Explain that small groups will work to name the purpose, pros and cons of the standardized test chosen, and record their ideas on the handout.
5. When groups have finished, ask each group to share their ideas with the whole class.
6. Summarize the main points from your class discussions. You can give everyone Handout 4b as an example of purposes, pros and cons, and review it together.

## Part 3 – Story – “Why Do We Take Tests?”

1. Post the Word Splash (Handout 5a) or Word List (Handout 5b) on flip chart paper, the board or Overhead Projector, and give students the handout you've chosen (5a or 5b). Ask students to talk about what they think these words mean, or what their experiences with these words are. Students can write sentences with some or all of the words. (See Teacher Notes.)

2. Give students the story *Why Do We Take Tests?* (Handouts 5c). Tell students beforehand that the class will read a story about one person's experience with tests, and that they will find the words they discussed above in this story.
3. Read the story together with students. Let students read silently for themselves first, and let them ask any questions about words or phrases they don't understand. Then practice reading the story with students. (See Lesson 2 Teacher Notes for ideas on how to read with students.)
4. Ask students some quick questions orally to make sure they have understood story content, for example, "How did Carmen feel about going back to school?" "How did she feel the first time the teacher gave a test?" "What reasons did the teacher give for why tests are important?"
5. Discuss how this story connects to students' experiences using these questions to guide your discussion:
  - How does this story compare to your experience? (How did you feel when you went back to school? How did you feel the first time you had to take a test in class? How do you feel about taking tests now? If you feel better about taking tests now, what has helped you? Or if not, why not?)
  - What do you think about the teacher's explanation of why tests are important? (What is your opinion of the teacher's explanation?)
  - If you were in Carmen's place, how would you feel at the end of the story?
6. Ask students to complete the Comprehension Questions (Handout 5d) and check their work.
7. Use Handout 5e to provide additional language practice with story content. You can also create other language practice activities. (See Teacher Notes.)

### Teacher notes

**Key ideas.** In your **Part 1** discussion (step 2), be sure to address **why teachers can't help students** during standardized tests. Many students may be concerned about this. Also, you as the teacher may find it difficult not to be able to help students during a standardized test. The teacher must play a different role during the administration of a standardized test – that of tester, not just teacher. To make the test experience the same for everyone who takes it, teachers must stick to their tester role (and do it consistently from test to test) so as not to create advantages for some students that others don't receive.

The teacher not being able to help may be a **comparison** students make between taking standardized tests and other tests. Other comparisons might be the test format (different from teacher-made tests), or the amount of time people had to complete the test. For some classes, discussing what experiences students' family members or friends have had with the MCAS can help students start to think about standardized testing.

In your **Part 2** discussion (step 1), a main **purpose of tests in general** is to tell if a person has the knowledge or skills needed for a particular purpose (for a diploma, a certificate, the next level class, going on to college, or a certain job). Tests also help to evaluate what skills students have learned, and what skills they still need to work on. Tests help the teacher to evaluate how effective her / his teaching has been, and to revise or change instruction and class content where needed. Some tests have higher-stakes purposes, for example determining who gets a job, who places into college or other education program, or who graduates from high school or earns a GED. Other tests have lower-stakes purposes, for example checking comprehension of specific content learned in a particular lesson.

The main **purpose of standardized tests** is to provide a **level playing field**; in other words, to be **fair**. Since all students are tested on the same content, under the same conditions and with the same scoring, standardized tests won't be influenced by favoritism, differences in test content or personal biases. Standardized tests are often used for higher-stakes purposes, for example to decide who gets a high school diploma, places into a higher-level education program, or gets a referral to a career center or job-training program.

For some classes, you may want to ask students to look up the definition of "standard" in the dictionary and discuss its meaning as it applies to standardized testing.

For more information on standardized tests see the article on page 69, "The Most Frequently *Unasked* Questions about Testing" by Stephen G. Sireci, in *Defending Standardized Testing*. (2005) Mahwah, NJ: Lawrence Erlbaum Associates Publishers. It is included in the handouts for Lesson 3.

Some **pros (advantages)** of standardized tests: standardized tests are the same for everyone; they are an efficient (easy) way to collect information about what people know and can do, and compare it; they are relatively objective and can't be influenced by the people giving or scoring the test; they can show what knowledge and skills people have; since everyone know what the test will be like, they can practice for it; standardized tests offer people an opportunity to move up (in school or a job). Students may name other pros as well.

Some **cons (disadvantages)** of standardized tests: test format may be confusing or difficult if students are not familiar with it; the test may not measure everything

*Testing – Standardized and Otherwise – and the Massachusetts ABE Pilot Test: A Curriculum Unit for ABE Classes.* By Jenny Lee Utech (Massachusetts Worker Education Roundtable), with Ruth Byrne and Susan Prior (Methuen Adult Learning Center), for the Massachusetts Department of Education ACLS, March 2005.

students know; some standardized tests favor people who are good at “pencil and paper” activities; sometimes, people have a limited amount of time to finish the test; standardized tests often have higher-stakes purposes (getting a diploma or job, moving up in an education program) and so they make people more nervous than they would be for an in-class quiz. Students may name other cons as well.

**Examples of purpose, pros and cons of the GED** (examples of what small groups might come up with): **Why do we have this standardized test?** Many adults who didn’t get a high school diploma want to show that they have the same level of knowledge and skills as high school graduates, to be able to have the same opportunities (placing into college or a job). The GED is used as a way for people to show that they have these knowledge and skills. Another way to do this would be for adults to go back to high school and take regular classroom tests, but this would be logistically difficult, time consuming and not very practical. (These examples are also in Handout 4b.) Students may name other purposes as well.

**What are some pros (advantages) of the GED?** Because it’s standardized, you can study and practice for the GED (on your own and in class, with study guides and practice tests available at bookstores or from the teacher). You need a GED to enter into college or college-level programs, to get some kinds of jobs or to enter the military. Everyone takes the same test (it’s fair, a level playing field). (These examples are also in Handout 4b.) Students may name other pros as well.

**What are some cons (disadvantages) of the GED?** Questions and format can be confusing if you aren’t familiar with them. The GED might not measure everything you know. It can take a long time (months or years) to become ready to take the GED. (These examples are also in Handout 4b.) Students may name other cons as well.

**Note:** Some of the Teacher Notes ideas above have been adapted from the following articles: “A Basic Primer for Understanding Standardized Tests” by April Zenisky, Lisa Keller and Stephen G. Sireci, found on page 60, and “How to Do Your Best on Standardized Tests: Some Suggestions for Adult Learners” by Ronald K. Hambleton and Stephen Jirka, found on page 51. These articles are reprinted from *Adventures in Assessment*, Volume 16. If you wish, they can also be found on the web at <http://www.sabes.org/resources/adventures/index.htm> They are also included in the handouts for Lesson 3. These articles are excellent resources for understanding what standardized tests are, and for test-preparation and test-taking strategies.

**Activity tips and variations.** In Part 1, step 1, before you post the definition, be sure to review any vocabulary (directions, conditions, content) that you think your class might not be familiar with. In Part 2, steps 3-5, you can have all small groups work on the purposes, pros and cons for the same test, or you can ask students what standardized tests they want to discuss and group them accordingly. To save time, you can do Part 2 as a whole-class discussion only (as you did in Part 1). To save even more time, you can ask students to name examples of purposes, pros and cons for a

*Testing – Standardized and Otherwise – and the Massachusetts ABE Pilot Test: A Curriculum Unit for ABE Classes.* By Jenny Lee Utech (Massachusetts Worker Education Roundtable), with Ruth Byrne and Susan Prior (Methuen Adult Learning Center), for the Massachusetts Department of Education ACLS, March 2005.

particular test or tests as you move through the discussion questions for Part 2, step 1 (and skip Part 2). Or you can simply go over Handout 4b with the class and ask them to add other examples (instead of doing steps 3-5).

In Part 2, for some classes it may work better to give students Handout 4b first and discuss it (before you have students develop their own lists of purposes, pros and cons of different standardized tests in steps 3-5. Going over handout 4b may help students to name purposes, pros and cons for other tests they've taken. If you are pressed for time, you may need to skip small group work (steps 3-5), and give students Handout 4b only (leave out Handout 4a).

In Part 3, use the Word Splash for higher-level classes only. You can handle the Word Splash in different ways. For example, you can divide students into small groups or pairs, ask them to brainstorm ideas and sentences for as many of the words as possible, write them down and share them with the class. You can do the Word Splash as an oral activity only. The approximate readability level for Handout 5c, *Why Do We Take Tests?* is 5.9 (Flesch/Kincaid). This story was written by Ruth Byrne and Susan Prior.

You can use the Jigsaw Activity (Handout 5e) to introduce the story, or to provide language practice and review story content after students have read the story. To use the Jigsaw, divide students in to pairs or small groups. Copy and cut the Jigsaw into pieces, enough for each pair or small group, and ask pairs or small groups to put the story pieces in the correct order. For lower-level readers, you can introduce the story by giving pairs or small groups one piece each and ask them to be responsible for reading, understanding and presenting that piece only. You can also create a cloze exercise to give students a chance to review story vocabulary.

## Lesson 4. Test-Taking Strategies

### Objectives

- Identify test-preparation and test-taking strategies.
- Create our own lists of test-preparation and test-taking strategies to use now and in the future.
- Evaluate and revise our strategies.

### Materials

- Flip chart paper, magic markers and masking tape.
- Statement about test-taking strategies (see step 1 below) written on flip chart paper or board.
- Handout 6a. Strategies Checklist
- Handout 6b. Strategies Checklist (version 2)
- Handout 6c. Strategies Revisited

### Activities

#### Part 1 – Test-taking difficulties and strategies

1. Post this statement and ask students, what do you think about this statement? Let students share their ideas.
  - “We are not born knowing how to take tests. We have to learn how to take them. We have to develop **test-taking skills** and **strategies**.”
2. Post and ask this question and record students’ ideas on flip chart paper so that you develop a class list. (Or, have students discuss the question in small groups, then share their ideas with the whole class, and develop a whole-class list):
  - What makes it **difficult** for you to take tests? (What makes it difficult for you to take / do well on tests?)
  - What makes it easy for you to take tests?
3. Now post and ask these questions and record students’ ideas on flip chart paper so that you develop another class list (Or have students discuss in small groups and report back). You can ask students to give ideas for strategies to tackle each “difficulty” listed during step 2, or simply ask for strategies in general:
  - What can you and others (the teacher, class, your family, your kids) do to prepare (or help you prepare) for taking a standardized test? And what can you do during the test? What **strategies** can you use? What have you done in the

past? (Or, What can you and others do about the difficulties we listed (in step 2)?)

- What strategies do you use after you take a test? (How do you deal with post-test worries, or reflect on how you performed and what you would do differently?)
4. After students have shared their ideas, add in any important strategies that you think are missing from the class list.

## Part 2 - Strategies checklists

1. Divide students into pairs. Give everyone Handout 6a or 6b (see Teacher Notes). Explain that everyone will work with their partner to develop their own checklist of test-preparation and test-taking strategies. Pairs will develop their checklists using the guidelines on the handout:
  - In column 1, write a list of your own test-preparation and test-taking **strategies** that you think will work for you.
  - You can use the ideas your class discussed, plus your own ideas.
  - (Or Version 2: Add your own strategies to the list of strategies in column 1.)
  - When you are finished with your list of strategies, share it with your partner.
  - Then, for each strategy on your list, discuss with your partner **how** you are going to accomplish this strategy (column 2), and **whose / what help** you need in order to do that (yourself, teacher, spouse, kids, friends) (column 3).
  - After you have discussed your ideas for **how** and **what help** for each strategy, write them in columns 2 and 3.
2. After pairs have finished, ask each student to share her / his list with the class and invite other students to ask questions or comment. (Or photocopy everyone's lists and share them with the class.)
3. Discuss how students can use their checklists as a tool now (to prepare for the pilot test) and in the future for tackling other tests.
4. After students have taken the pilot test, ask them to revisit and revise their checklists in class (see Teacher notes). You can use Handout 6c to help students revise their lists.

## Teacher notes

**Key ideas.** Depending on your class, students may have a little or a lot of experience with standardized tests. They may handle standardized tests well and feel confident about them, or not. They may have already developed lots of strategies for taking tests, or not. The goal here is for everyone to share their ideas about what makes

taking standardized tests difficult and what strategies students can use to do as well as possible on them. By asking students to list difficulties, strategies, and ways to accomplish the strategies, you are helping students develop or refine life skills that will help them both in and outside of class.

Things that make taking standardized tests **difficult** for students might include:

- You had negative experiences with testing in the past.
- Your first language is not English.
- You don't have time to study, prepare or come to class regularly (because of family, job or health issues).
- You are stressed because of family, job or health troubles and you can't concentrate during class or during the test.
- You are tired or sick.
- You are not familiar with the test format (how to take the test).
- You run out of time and can't finish the test (you spend too much time on certain sections, or don't pace yourself.)
- You don't understand certain words or problems on the test and you get stuck on those. Students may name other difficulties as well.

**Strategies** for test-preparation and test-taking might include:

- Talk with your family, teachers or classmates about your negative test-taking experiences and how you can overcome them.
- Work with your family to find time to study at home, come to class or get enough sleep.
- Ask your teacher for more practice tests and lessons.
- Listen carefully to test directions. Scan the test before you start answering questions.
- Don't leave test questions unanswered. Make educated guesses if you are not sure of the answer.
- Students may name other strategies as well.

**Post-test strategies** might include:

- Reflect on how you performed and make a note (mental or written) of what you would do the same or differently before and during the next test you take.
- Revisit the test-preparation and test-taking strategies you developed in class and revise or change them as needed.
- Make a plan with your teacher, family, etc. for tackling future tests.
- Then put the test out of your mind and relax.

When students are creating their **strategies checklists**, encourage them to think of specific things they can do to accomplish the strategies. For example, a **strategy** might be "take more practice tests". **How** might be "in class", "get an extra practice book for home", and "make time to do practice tests at home." **Who / whose help** might be "ask my spouse / kids to clean up after supper 2 nights a week so I can practice then,"

“ask my spouse to get the kids ready for school 2 mornings a week so I can practice then,” or “go to my friend’s apartment, where it’s quiet, to do practice tests on Wednesday nights.”

**Note:** See Handout 6b, and the article “How to Do Your Best on Standardized Tests: Some Suggestions for Adult Learners” by Ronald K. Hambleton and Stephen Jirka, for a list of key strategies, on page 51. This article, reprinted from *Adventures in Assessment*, Volume 16, is an excellent resource for understanding standardized tests, and test-preparation and test-taking strategies. It is included in your handouts under Lesson 3, or you can find it on the web at <http://www.sabes.org/resources/adventures/index.htm>

**Note:** This curriculum unit, and Lesson 4 in particular, will help you and your class work toward the Massachusetts ABE Curriculum Frameworks’ Habits of Mind and Guiding Principles, as they are discussed in the Common Chapters for the Frameworks. The Habits of Mind of persistence and reflection are very important to developing and refining lifelong learning. They are discussed on pages 7-9 in the Common Chapters.

In addition, there are two Guiding Principles mentioned on page 12 in the Common Chapters that you can draw from:

"Staff and students of adult learning programs base instruction and assessment on inquiry involving investigation, critical thinking, problem solving, and reflection."

"Staff and students at adult learning programs see assessment as a tool to provide evidence of progress and assist students in becoming self-directed, lifelong learners."

The web address for the Common Chapters and other Massachusetts ABE Curriculum Frameworks is: <http://www.doe.mass.edu/acls/frameworks>

**Activity tips and variations.** You can do this activity in many different ways. For your Part 1 discussions, as you are listing difficulties during step 2, you can leave some space after each difficulty. Then during step 3 you can ask students to name strategies for each difficulty and write them in. Or, you can post another sheet of flip chart paper next to your difficulties list and write strategies on that. Then ask students which strategies could help with which difficulties and connect them with lines. Or, you may not want to ask for specific strategies for the difficulties the class listed, but for strategies in general, and leave the specifics to the individual checklists students will develop during pair work.

For Part 2 (strategies checklists), students can refer to the brainstormed list of strategies and then develop their own personal lists. Or, students can simply take strategies from list brainstormed during Part 1, step 3 (or from the *Adventures in Assessment* article) and write them in column 1. Or, you can do Part 2 as a whole-class activity. Have the whole class create and refine a checklist of strategies based on ideas

*Testing – Standardized and Otherwise – and the Massachusetts ABE Pilot Test: A Curriculum Unit for ABE Classes.* By Jenny Lee Utech (Massachusetts Worker Education Roundtable), with Ruth Byrne and Susan Prior (Methuen Adult Learning Center), for the Massachusetts Department of Education ACLS, March 2005.

that surfaced during Part 1, step 3. Then, for each strategy on the checklist, ask the class how people could accomplish the strategy and whose help they need in order to do it (teacher, spouse, kids, friends.)

Handout 6a lets students create a checklist of strategies “from scratch.” For some classes, you may want to use Handout 6b instead of Handout 6a. If you use Handout 6b, you can give it to students after you have discussed difficulties in Part 1, step 2. Explain that Handout 6b, column 1 lists strategies people can use to deal with the difficulties of preparing for and taking standardized tests. In pairs or as a whole class, students can then add additional strategies to column 1, and work on the “how” and “who” for each strategy.

While it is more effective to have students develop their own checklists (using Handout 6a), Handout 6b may work better for your situation (depending on how much time you have and the level of your class).

Strategizing shouldn't end after students have taken a test. **After students have taken the pilot test**, ask them to revisit their checklists in class. Students can discuss which strategies worked for them and why, and which strategies they weren't able to use and why not. Students can revise or add to their checklists, and discuss how they will continue to work on their strategies. Students can discuss and answer questions about each strategy on their list, for example: How did the strategy work and when did you use it? If you used this strategy again, what would you change (if anything)? You can also have teacher-student or student-student conferences about the checklists. Students can use Handout 6c as a tool for revisiting and revising their checklists.

Students can take their work on strategies even further by presenting to other classes, writing about strategies for a student newspaper or *The Change Agent*, or incorporating their work on test-taking strategies into a class project on student leadership.

## Lesson 5. Pilot Testing

### Objectives

- Examine and discuss the definition of pilot testing.
- Discuss the new ABE test that will be piloted in Massachusetts, and how we will participate in the pilot.

### Materials

- Flip chart paper, magic markers and masking tape.
- Definition of pilot testing (from Handout 7) written on flip chart paper or board (or copied onto an Overhead sheet).
- Handout 7. Pilot Tests
- Handout 8a. Pilot Tests – Matching Sentence Halves (front halves).
- Handout 8b. Pilot Tests – Matching Sentence Halves (back halves). (Copy 8a and 8b onto two colors, one for front halves and one for back halves, and cut up enough for each pair of students.)
- Handout 8c. Pilot Tests – Sentence Halves Matched

### Activities

#### Part 1 - What is a pilot test?

1. Ask students the following questions and let everyone share their ideas:
  - What is a **pilot test**? What does it mean to pilot test something? For example, what does it mean when a TV show gets “pilot-tested”? Or when they pilot test a drug like Celebrex or Viox? (It means they are trying it out to see if it is good, and how it works, and if it works consistently well.)
2. Post the explanation of pilot testing (from Handout 7) and review it with the class:
 

Test makers who are creating a new kind of test need to know if the new test is **fair** and **reliable**. To find out, they give the new test to many people to see how well it works, and if the questions are clear, fair, too easy or too hard. This is called a **pilot test**. Test makers do a pilot test to try out the new test and see how it works. After the pilot, they will look at everyone’s test results and change parts of the test that didn’t work well.
3. Explain that the class will be taking a **pilot test**, along with students around Massachusetts. Give students Handout 7 and review it with the class. (The handout covers what the pilot test is, why we are doing it and when.)

4. Explain what will happen to students' scores will be used to see how well the test worked, so that test-makers can revise it. Students' scores will not "count" as pre- and post-tests (they still have to take the TABE this year).
5. Explain that students will review the information on Handout 7 by matching sentence halves. Divide students into pairs and give each pair a set of front sentence halves and back sentence halves (Handouts 8a and 8b cut into strips). Show people how to match front halves with back halves and let them work in pairs.
6. As pairs finish matching, give them Handout 8c so they can check their work. In lower-level classes, pairs can practice reading matched sentences to each other, or copy sentences they had trouble with.

### Teacher notes

**Key ideas.** As you discuss the pilot test, make sure that students understand that it is an important step in making the final test better for everyone. ABE students in Massachusetts are a key part in the process of developing and shaping the new tests for reading and writing – it could not be done without them. Once the test makers see everyone's test results, they will be able to tell which test questions are at the right level of difficulty, are the clearest (i.e. students give the anticipated answer because the question is worded correctly), and are fair (for women and men, for older and younger adults, for people from the US and from other cultures). The pilot test gives teachers and students the chance to have a voice in how the test is developed, and to raise questions and concerns before the final test is completed.

If test is **reliable**, it gives consistent results each time it is used. To explain "reliable" to students, you can use the example of a bathroom scale. If the scale gives you the same weight every time you step onto it (provided that you are not gaining or losing weight), then it is reliable.

**Activity tips and variations.** If you have time, you can discuss and / or read articles or online information about pilot or field-testing of different kinds (TV show pilots, field-testing of drugs (Celebrex, Viox, cancer drugs, or testing of consumer products in supermarkets or phone surveys.) You can ask students to develop their own test questions to test a particular skill or chunk of knowledge, and then test them out on classmates.

For additional language practice, you can create a cloze exercise using sentences from Handout 7. You can write questions about Handout 7 content onto cards, and have students work in pairs to ask each other the questions and answer them in their own words.

**Note:** See the February 16 letter from Adult Community Learning Services, and the document “Looking Ahead to March: The New ABE Reading and Math Field Testing Process,” for more information on the Massachusetts pilot test and timeline. These are included with the handouts for this lesson.

## Lesson 6. Pilot Test Follow Up

### Objectives

- Discuss and evaluate our experiences with the pilot test.
- Evaluate and revise our test-preparation and test-taking strategies.

### Materials

- Flip chart paper, magic markers and masking tape.
- Handout 9a. Story - The Pilot Test
- Handout 9b. Cloze Exercise - The Pilot Test

### Activities

#### Part 1 - Experiences with the pilot test

1. Lead a whole-class discussion about students' experiences with the pilot test, using the questions below (record students' ideas on flip chart paper if needed):
  - What was your experience taking the pilot test? How did it go?
  - What was difficult about the test (which parts, which questions)?
  - What was easy / good about the test (which parts, which questions)?
  - What would you change about the test (directions, format, content)? Why?
  - How does this test compare to other tests you've taken?
  - How will you prepare for / handle the next round of pilot testing? What help do you need? (You can also discuss this question in step 5 below.)
  - What did you learn that will be useful for you for other tests (TABE, GED)?
2. Have students evaluate and revise the strategies checklists they created in Lesson 4. (See Lesson 4 Teacher Notes).

#### Part 2 - Story - The Pilot Test

3. Post and review key vocabulary from The Pilot Test (Handout 9a), on flip chart paper, the board or Overhead Projector.
4. Give students Handout 9a. Tell students that the class will read a story about one person's experience with the pilot test.

5. Read the story together with students. Let students read silently for themselves first, and let them ask any questions about words or phrases they don't understand. Then practice reading the story with students. (See Lesson 2 Teacher Notes for ideas on how to read with students.)
6. Ask students some quick questions orally to make sure they have understood story content, for example, "At the beginning of the story, how did Carmen feel about coming to school?" "How did she react to the teacher's news that they would take a new test?" "What did the teacher tell the class about the new test?" "What was Carmen's experience with the pilot test?"
7. Discuss how this story connects to students' experiences using these questions to guide your discussion:
  - What things made Carmen feel good about the pilot? What things worried her?
  - How does this story compare to your experience with the pilot test? (How did you feel when you first heard about the pilot test? How did you feel when you took the pilot test?)
  - What do you think about the teacher's explanation of why the pilot test is important? (What is your opinion of the teacher's explanation?)
  - What other questions would you ask the teacher?
  - If you were in Carmen's place, how would you feel at the end of the story?
8. Use the Cloze Exercise (Handouts 9b) to provide additional language practice with story content. You can also create other language practice activities. (See Teacher Notes.)

### Teacher notes

**Key ideas.** It's important for students to understand that the test is not in its final form, and that student feedback and opinions can be helpful in shaping the final test. Teachers and students can give feedback about the test to Adult and Community Learning Services (Massachusetts DOE) and the Center for Educational Assessment, UMASS-Amherst by emailing [abepilottest@doe.mass.edu](mailto:abepilottest@doe.mass.edu).

**Activity tips and variations.** For ideas on how to have students evaluate and revise their strategies checklists, see Lesson 4 Teacher Notes. Students can use Handout 6c to revisit and revise their lists.

In Part 2, The approximate readability level for Handout 9a, The Pilot Test, is 7.6 (Flesch Kincaid). This story was written by Ruth Byrne and Susan Prior. The cloze exercise (Handout 9b) provides additional vocabulary practice.

### **Assessment for this curriculum unit**

An evaluation activity of some kind is built into every lesson (reading comprehension questions, small groups writing lists of pros and cons, matching sentences, cloze exercises, developing and evaluating the check list, etc.), to help students demonstrate that they have met the lesson's objectives. Another way to evaluate what students have learned might be teacher - conferences with a checklist or interview questions about students' test-taking strategies or other unit content. Or, students could develop a portfolio that could include their story about a test-taking experience, handouts they have completed, their strategies checklists and revised checklists, pilot test scores, or other work.

## **Handout 1a. Magaly's Story**

When I was growing up in my country, I was a good student. In high school I was president of my class. We were poor, but we would organize events to raise money for school parties, or organize soccer or basketball teams. I loved doing that.

I organized study groups with some other girls, and we would help each other study for tests. Before every test, we would get together to study at one of our houses. We would ask ourselves, what are the most difficult parts of this subject? And we would study those. I loved the challenge of getting ready for a test!

After a test, I would feel so satisfied when everyone in our group did well.

I always did well on tests. I hated to get less than 90% on a test. That was an insult for me! But that almost never happened.

## Handout 1b. Will's Story

When I was growing up I had trouble with spelling and writing. If I was just doing homework, I could manage. And I was good at expressing myself and talking about whatever subject we were learning. But when it came to writing under pressure I would freeze up.

So I always hated taking tests. I would struggle to write down what was in my head, but I could never put what I knew down on the test paper. My head was full of ideas, but my test paper wasn't. It was only full of spelling mistakes. Whenever I got my tests back they had red marks all over them, and my scores were always low. I still hate taking tests today.

## Handout 1c. Mark's Story

When I was 16 years old I dropped out of school. I was bored with my classes, and school seemed like a waste of time to me when I could be out earning money instead. So I started working in the hospital's transport department when I was 16.

I got married at 20, and soon I had three children. Once I had them to take care of, time just passed by too quickly. I had my job at the hospital and I kept busy raising my kids.

By the time I turned 35, I knew I wanted to go back to school and get my high school diploma. If I didn't, I would just be stuck in my job. And I wanted to help my kids with their schoolwork more. My oldest son will have to get ready for the MCAS soon.

So now I'm taking GED classes and getting ready to take the GED tests. I took one test already. When I took it I was so nervous I felt sick to my stomach! I did O.K., and I am determined to do even better the next time around.

## Handout 1d. Anne's Story

My family never encouraged me to do well in school. My mother didn't finish school, so she couldn't help me. I didn't get a good education, and that hurts your job opportunities.

Last year there was this position, a Dietary Aide at the hospital, and I really wanted it. But they wanted to give me a test. I was so anxious about it! When you hear that word "test," everything just goes out of your brain.

I was so nervous. My hands were sweaty. But my teacher encouraged me. She always told me, "You can do it." During the test I was shaking, but I heard her voice telling me, "Girl, you can do it." I heard her telling me I was smart.

And I went out there and I got a 98% on that test! I did it. I got it. I'm smart!

Adapted and excerpted from Rothschild, B and Willoughby, S., eds. (1999). *Proclamation of Independence: A Book of Life Stories Told by Women Students in the Second Start Adult Literacy Program*. Oakland, CA: Second Start Adult Literacy Program.

## Handout 2a. Magaly's Story - Yes or No

### Magaly's Experience

When Magaly was growing up, she was a good student.	Yes	No
Magaly was president of her class.	Yes	No
Magaly studied for tests by herself.	Yes	No
Magaly studied for tests with her group of friends.	Yes	No
Magaly felt satisfied when everyone in her group did well.	Yes	No
Magaly always did well on tests.	Yes	No
Magaly hated to get less than 70% on a test.	Yes	No
Magaly always got less than 70% on a test.	Yes	No

### My Experience

When I was growing up, I was a good student.	Yes	No
I organized activities at school.	Yes	No
I studied for tests by myself.	Yes	No
I studied for tests with friends.	Yes	No
I always did well on tests.	Yes	No
I hated to get less than 90% on a test.	Yes	No

## Handout 2a, part 2. Questions about Magaly's Story

### Magaly's Experience

When she was growing up in her country, what kind of student was Magaly?

What did Magaly like to organize at school?

How did Magaly study for tests?

How did Magaly feel after taking a test?

How did Magaly usually do on tests?

### My Experience

When you were growing up, what kind of student were you?

What kinds of activities did you like at school? Did you organize activities? What activities?

How did you study for tests?

How did you feel after taking a test?

How did you usually do on tests?

## Handout 2b. Will's Story - Yes or No

### Will's Experience

When Will was growing up he had trouble with spelling and writing.      Yes    No

If Will was doing homework, he had trouble.      Yes    No

Will was not good at expressing himself and talking about different subjects.    Yes    No

Will was good at writing under pressure.      Yes    No

Will always hated taking tests.      Yes    No

Will struggled to write things down on tests.      Yes    No

Will's tests were full of spelling mistakes and red marks.      Yes    No

Will likes taking tests today.      Yes    No

### My Experience

When I was growing up I had trouble with spelling and writing.      Yes    No

If I was doing homework, I had trouble.      Yes    No

I am good at expressing myself and talking about different subjects.      Yes    No

I am good at writing under pressure.      Yes    No

When I was growing up, I always hated taking tests.      Yes    No

My tests were full of mistakes.      Yes    No

I like taking tests today.      Yes    No

## **Handout 2b, part 2. Questions about Will's Story**

### **Will's Experience**

When Will was growing up, what did he have trouble with in school?

When Will was growing up, what was he good at in school?

Why did Will hate taking tests? What was his experience taking tests?

How did Will do on tests?

How does Will feel about taking tests today?

### **My Experience**

When you were growing up, what did you have trouble with in school?

When you were growing up, what were you good at in school?

How did you feel about taking tests in school?

How did you do on tests?

How do you feel about taking tests today?

## Handout 2c. Mark's Story – Yes or No

### Mark's Experience

Mark was 16 when he dropped out of school.	Yes	No
Mark was bored with his classes.	Yes	No
Mark started working in the hospital when he was 20.	Yes	No
Mark got married at 20 and had three children.	Yes	No
Mark went back to school and got his high school diploma.	Yes	No
Mark is taking GED classes and getting ready for the GED tests.	Yes	No
Mark wants to help his kids with the MCAS.	Yes	No
Mark was so nervous that he did badly on the first GED test.	Yes	No

### My Experience

I dropped out of school.	Yes	No
I was bored with my classes.	Yes	No
I got my high school diploma.	Yes	No
I started working when I left school.	Yes	No
I got married and had children.	Yes	No
I went back to school as an adult.	Yes	No
I want to get my high school diploma in the future.	Yes	No
I want to help my kids with their homework or the MCAS.	Yes	No

## Handout 2c, part 2. Questions about Mark's Story

### Mark's Experience

Why did Mark drop out of school?

What did Mark do after he dropped out of school?

What happened after Mark got married?

Why does Mark want to get his GED now?

How does Mark feel about taking the GED tests?

### My Experience

When did you leave school? Did you graduate? If not, why not?

What did you do after you left school?

Why are you back in school now?

How do you feel about taking tests?

Has taking tests gotten easier or harder for you over time? Why?

## Handout 2d. Anne's Story – Yes or No

### Anne's Experience

Anne's family never encouraged her to do well in school.	Yes	No
Anne's mother finished school.	Yes	No
Anne thinks that your education hurts your job opportunities.	Yes	No
Anne wanted a position as a Dietary Aide at the hospital.	Yes	No
Anne really wanted to take the test to be a Dietary Aide.	Yes	No
Anne was so nervous about the test that her hands were sweaty.	Yes	No
Anne's teacher told her, "You can do it."	Yes	No
Anne got a 90% on the Dietary Aide test.	Yes	No

### My Experience

My family encouraged me to do well in school.	Yes	No
My parents finished school.	Yes	No
I think that your education affects your job opportunities.	Yes	No
I have taken a test to get a job.	Yes	No
I get so nervous about tests that my hands get sweaty.	Yes	No
My teacher encourages me.	Yes	No

## Handout 2d, part 2. Questions about Anne's Story

### Anne's Experience

What does Anne think about the education she had growing up?

According to Anne, how does your education help or hurt your job opportunities?

How did Anne feel about taking the Dietary Aide test?

What helped Anne while she was taking the Dietary Aide test?

How did Anne do on the Dietary Aide test, and how did she feel about it?

### My Experience

What do you think about the education you got growing up?

According to you, how does your education help or hurt your job opportunities?

Have you ever taken a test to get a new job? How did you feel about the test? How did you do?

What other kinds of tests have you taken?

How did you feel about taking those tests?

What happened after you took those tests? What were the results? What did you do next?

## Handout 3. What are standardized tests?

If a test is **standardized**, it means that:

- Everyone gets the **same directions** for taking the test.
- Everyone takes the test under the **same conditions** (the same amount of time, in a quiet room).
- Everyone gets the **same test content** (or same type of content).
- Everyone's test is **scored the same way**.

The **goal** of a standardized test is to make the test experience the same for everyone who takes it, so that no one has an advantage over anyone else.

## Handout 4a. Why use standardized tests?

*Write the name of the standardized test you have chosen. Then discuss the three questions below with your group. After you have discussed your ideas, write them down.*

**Test name:** \_\_\_\_\_

**Why do we have this standardized test?**

---



---



---



---

What are some <b>pros (advantages)</b> of this test?	What are some <b>cons (disadvantages)</b> of this test?

## Handout 4b. Why use the GED?

Write the name of the standardized test you have chosen. Then discuss the three questions below with your group. After you have discussed your ideas, write them down.

**Test name:** GED

### Why do we have the GED?

Many adults who didn't get a high school diploma want to show that they have the same level of knowledge and skills as high school graduates, to be able to have the same opportunities (placing into college or a job).

The GED is used as a way for people to show that they have these knowledge and skills.

Taking the GED is easier than going back to high school and taking regular classroom tests (which would not be very practical).

What are some <b>pros (advantages)</b> of the GED?	What are some <b>cons (disadvantages)</b> of the GED?
<p>Because it's standardized, you can study and practice for the GED – on your own and in class, with study guides and practice tests available at bookstores or from the teacher.</p> <p>You need a GED to enter into college or college-level programs, to get some kinds of jobs or to enter the military.</p> <p>Everyone takes the same test (it's fair, a level playing field).</p>	<p>Questions and format can be confusing if you aren't familiar with them.</p> <p>The GED might not measure everything you know.</p> <p>It can take a long time (months or years) to become ready to take the GED.</p>

## Handout 5a. Word Splash – Why do we take tests?

**INFORM**

**INSTRUCTION**

***TESTS***

**ANXIETY**

**SWEAT**

**STANDARDIZED**

**UNDERSTAND**

**LESSONS**

**PROGRESS**

**ASSESSMENT**

## Handout 5b. Word List – Why do we take tests?

anxiety

assessment

inform

instruction

lessons

nervous

progress

standardized

stress

sweat

tests

tool

understand

assess

reliable

## Handout 5c. Story – Why do we take tests?

Carmen Valdez had just enrolled in an adult education program in her hometown. She was really nervous about going back to school because she had always struggled with schoolwork in the past. Her teacher, Miss Brown, assured her that there was nothing to be worried or concerned about, for the school was going to help her through her fears. This made Carmen feel more at ease and relaxed about attending adult education classes.

Carmen's first few days in school went very well, until Miss Brown announced that the class was going to take a test. Carmen felt a panic attack coming, making her feel quite anxious. "Yikes!" Carmen exclaimed to herself. "The teacher is giving us a test! Miss Brown had promised that this place was going to help me. Why is she giving us a test already?" Carmen wanted to run out of the room, but then Miss Brown began to tell the class why the school tests or assesses the students so soon after they come to class.

The teacher explained that by looking at the results of students' tests, she would be able to teach better and the class would be able to learn more. She said that test results inform both the teacher and the students about what the class needs to know. For example, if the tests showed that students had trouble finding the main idea in a reading passage, the teacher would prepare lessons to help the students until they were successful at this skill.

Miss Brown also explained that test scores provide both the teacher and the students with a way to measure progress and improvement. She told the class that she compares test scores from the beginning of the school year to the middle and end of

## Handout 5c, page 2. Story – Why do we take tests?

the year to help her learn if students are making progress throughout the year. “Wow!” Carmen thought. “I always thought that tests only showed what I couldn’t do. I never realized that my mistakes could be used to help me improve. I’m so glad that the teacher helped me to think of testing in a positive way. Next time I take a test, I won’t be so nervous.”

Now Carmen understands that assessment is a way to guide classroom instruction. She doesn’t shake, sweat or perspire anymore when Miss Brown announces, “Today you’re going to take a standardized test.” Her stress level has decreased because she realizes that her school is using the test results to measure what she knows and to help her learn more and improve her skills.

## Handout 5d. Comprehension Questions – Why do we take tests?

1. How did Carmen feel when she first enrolled in an adult education program?
  - a) fine
  - b) feverish
  - c) nervous
  - d) calm
  
2. What is the main idea of the passage?
  - a) Carmen is an anxious person.
  - b) Miss Brown is a good teacher.
  - c) Tests make everyone nervous.
  - d) Test results help both teachers and students.
  
3. The information given from the results of a test \_\_\_\_\_.
  - a) informs the student
  - b) informs the teacher
  - c) both a and b
  - d) doesn't inform anyone
  
4. Why did taking a test make Carmen nervous?
  - a) She didn't understand how the results could help her.
  - b) She was good at test-taking.
  - c) Her mother always got anxious, so she did, too.
  - d) She was worried about a doctor's appointment the same day.
  
5. Which phrase most closely describes Carmen by the end of the story?
  - a) Carmen is more relaxed because she understands the importance of testing.
  - b) Carmen is less relaxed because she understands the importance of testing.
  - c) Carmen realizes that test-taking does not have a purpose.
  - d) Carmen has learned how to find the main idea in a reading passage.

## Handout 5d, page 2. Comprehension Questions – Why do we take tests?

6. According to the passage, which statement is true?
- a) Tests are meant to show failure.
  - b) Tests are used only to measure progress.
  - c) Tests are a waste of time.
  - d) Tests inform instruction for the teacher to help her students.
7. **Assessment** in the classroom is another word for:
- a) understanding why students get nervous
  - b) decreasing anxiety
  - c) measuring what students know and how much they have learned
  - d) stressing what students have learned in life
8. Which comparison best describes Carmen's emotions from the beginning to the ending of the story?
- a) anxious to calm
  - b) calm to anxious
  - c) anxious to nervous
  - d) shaky to concerned

**Why Do We Take Tests?**  
**Handout 5d, page 3. Comprehension Questions Answer Key**

1. c
2. d
3. b
4. c
5. a
6. a
7. d
8. c
9. c
10. a

## Handout 5e. Jigsaw

Carmen Valdez had just enrolled in an adult education program in her hometown. She was really nervous about going back to school because she had always struggled with schoolwork in the past. Her teacher, Miss Brown, assured her that there was nothing to be worried or concerned about, for the school was going to help her through her fears. This made Carmen feel more at ease and relaxed about attending adult education classes.

-----

Carmen's first few days in school went very well, until Miss Brown announced that the class was going to take a test. Carmen felt a panic attack coming, making her feel quite anxious. "Yikes!" Carmen exclaimed to herself. "The teacher is giving us a test! Miss Brown had promised that this place was going to help me. Why is she giving us a test already?" Carmen wanted to run out of the room, but then Miss Brown began to tell the class why the school tests or assesses the students so soon after they come to class.

-----

The teacher explained that by looking at the results of students' tests, she would be able to teach better and the class would be able to learn more. She said that test results inform both the teacher and the students about what the class needs to know. For example, if the tests showed that students had trouble finding the main idea in a reading passage, the teacher would prepare lessons to help the students until they were successful at this skill.

-----

## Handout 5e, page 2. Jigsaw

Miss Brown also explained that test scores provide both the teacher and the students with a way to measure progress and improvement. She told the class that she compares test scores from the beginning of the school year to the middle and end of the year to help her learn if students are making progress throughout the year. “Wow!” Carmen thought. “I always thought that tests only showed what I couldn’t do. I never realized that my mistakes could be used to help me improve. I’m so glad that the teacher helped me to think of testing in a positive way. Next time I take a test, I won’t be so nervous.”

## Lesson 3 Handout, Article #1

[About SABES](#) [Contact Us](#) [Site Map](#)

	<a href="#">What's New</a>	<a href="#">SABES Resources</a>	<a href="#">SABES Calendar</a>	Top of Form
	<a href="#">Teacher's License</a>	<a href="#">Other Resources</a>	<a href="#">Other Sites</a>	<input type="text"/> <input type="button" value="Search"/> <input type="text" value="sp1001a0c1"/>
				Bottom of Form

# How to Do Your Best on Standardized Tests: Some Suggestions for Adult Learners

Ronald K. Hambleton and Stephen Jirka

UMASS Amherst



[Home](#)

[Volume 16 :](#)

[Spring 2004](#)

[CONTENTS](#)

[How to Do Your Best](#)

[on Standardized Tests:](#)

[Some Suggestions for](#)

[Adult Learners](#)

[Using the REEP](#)

[Assessment for ESOL](#)

[and ABE Classroom](#)

[Instruction](#)

[Integrating Goal](#)

[Setting into](#)

[Instructional Practice](#)

[A Basic Primer for](#)

[Understanding](#)

[Standardized Tests](#)

[and Using Test Scores](#)

[Using Data for](#)

[Program Improvement](#)

[ACLS, SABES,](#)

[UMASS: Perfect](#)

[Together!](#)

## PART ONE: WHAT ARE "STANDARDIZED TESTS"?

Educational tests, sometimes called "standardized tests," seem to be everywhere. In Massachusetts, the Department of Education administers the Massachusetts Comprehensive Assessment System tests (better known as MCAS) in English Language Arts, Mathematics, and Science to learners in the public school system. The Educational Testing Service administers the Scholastic Assessment Test (the SAT) to learners who are considering going to college. The American Council on Education administers the Tests of General Educational Development (GED). You can't get a high school diploma, go to college, join the military, get a professional license or certificate, or get a job without passing a test. You can't even get a driver's license without passing a test. With so many standardized tests around, adult learners would be wise to learn how to do their best on them, and to help their children do well on them, too.

Standardized tests are particular kinds of tests, different from the final examination a high school teacher might design for her math course, or the writing exercise an ESOL teacher might design to see how well his learners are doing. When talking about tests, "standardized" simply means that everyone who takes the test is given the same amount of time and sees the same or very similar test questions. "Standardized" also means scoring is done very carefully so that test scores do not depend upon who happens to be doing the

scoring. Why are standardized tests so widely used? Because, by and large, they have been shown to be (1) an efficient way to collect information about what people know and can do, (2) objective in the sense that test scores do not depend to any great extent on who happens to score the answers, (3) valid in that they often provide relevant and useful data for making decisions about mastery of a body of knowledge and skills and potential for success, and (4) convenient and cost-effective because they can be administered to many people at the same time.

Governments, the armed services, industry, universities and colleges, credentialing agencies, and many other groups use standardized tests because they are convinced by the evidence that such tests offer the best basis for making decisions about who has the necessary knowledge and skills for some particular purpose (like going on to college or being hired for a job.) Human beings make tests and human beings administer them, and all human beings have biases. Bias can sometimes creep into standardized tests, but it can usually be spotted and the problem fixed, or, if the problem of bias cannot be solved, the test can be eliminated.

Some people believe that standardized tests are used too often and that there are better ways to measure ability and readiness. For purposes of discussion, let's consider the task of determining whether adult learners have the same knowledge and skills as high school graduates. This is an important task in the United States, because high school diplomas are an entry to higher education, the military, and lots of jobs. Many adults who did not obtain a high school diploma during their teens later want to demonstrate that they have about the same level of knowledge and skills as high school graduates and thereby gain the same opportunities. Today we have the Tests of General Educational Development (the five tests that make up the GED), which are used around the country as a way for people to demonstrate they have knowledge and skills equivalent to those of high school graduates. One alternative to passing the GED would be for adult learners to return to high school and take regular school tests along with state graduation tests, but with a million persons desiring GED certificates each year, this would surely be impractical. The external diploma programs offered by many adult basic education programs are an excellent alternative, but they require a great deal of individual conferencing.

As a standardized test, then, the GED certainly has its place. It provides many thousands of adult learners in this country with a second chance. Teachers are familiar with the material covered by GED tests, so they can design test preparation instruction effectively. And the GED is widely accepted as a high school equivalent: community colleges, universities, the military, skilled trades, and

employers who require a high school diploma welcome those who demonstrate proficiency through the successful passage of the GED tests. Clearly, the GED tests and others like them have an important role to play in this country.

We believe that some of the problems surrounding the standardized tests used in adult basic education programs, such as the GED and the TABE, are not with the tests themselves, but with learners' test-taking anxiety and lack of test-taking skills. These two factors are interrelated; knowing more about standardized tests and how to take them can boost a learner's self-confidence and reduce her test-taking anxiety. However, people are not born with test-taking skills, and sometimes learners from other countries have had very little exposure to American-style tests with multiple-choice items and separate answer sheets, or with the computer-administered tests that are becoming popular.

## **PART TWO: DOING YOUR BEST ON STANDARDIZED TESTS**

At this point, we would like to offer six very practical suggestions to help adult learners perform to their best ability on standardized tests.

### **1. Get positive about taking tests!**

Adult learners need to think positively about themselves, the learning they are doing, and the tests they will be taking to assess their learning. While standardized tests can be daunting, they also offer adult learners a way to move up, to provide a role model for their children, to get a better job, or to go to college. All too often, adults without a basic education see themselves as victims. A positive attitude can boost confidence and improve test performance.

Researchers have found that test performance is, in part, psychological. When learners receive positive messages about their ability to learn and to succeed academically, they are less likely to conform to stereotypes that they believe others have of them, and they perform significantly better on tests. So, adult learners and their teachers must be positive!

Adult learners need to see testing as an opportunity to demonstrate their ability, not evidence that they are victims of a system that cares little about them. Doctors often tell their patients to be positive, because research has shown that patients who remain positive live longer and avoid illnesses better than those who do not. The same is true for adult learners when taking tests—be positive and you'll perform at a higher level.

## **2. Clear the brain for learning and testing!**

Many adult learners lead stressful lives. Stress comes from family, from the job, from personal health concerns, from the times we live in, and so on. But if adult learners want to improve their lives and those of their family members, they need to find time to concentrate on learning. Adult learners need to have some quiet time each week to study, and regularity and consistency make learning easier. They must see this "learning time" as something they deserve. The study place should be quiet to allow for concentration—perhaps the local library on a Saturday morning, or a quiet place at home in the early morning or late night if necessary, and should be dedicated to studying, with books, paper, and pen readily available. Learners need to stay organized because this time is precious, and they owe it to themselves to make the most of it.

Adult learners also need some quiet time right before taking a test. An hour or two to clear their heads of life's stresses, away from family, away from the job: time to think about the challenges associated with the upcoming test. An adult learner who arrives late for a test, huffing and puffing, upset about a family- or job-related problem is not emotionally ready for the challenges of a test. If failure follows, the test is often blamed, but the real problem might be that the adult learner was not psychologically ready to perform to her capabilities. If prior test-taking experience resulted in failure, the adult learner should strive to put that behind her and focus on the present test and her efforts to perform well on it.

## **3. Prepare for the test "strategically"!**

We were talking with a colleague the other day who told us about an adult learner who persisted in studying for one section of a GED test that he thought he was weak in. He had failed the test several times previously, yet this one section was only 10% of the test. This learner would have been much wiser to consider the content coverage of the test (which was information readily available to him) and to plan his study time accordingly.

There are two key strategies for preparing to take a standardized test.

The first strategy is to become familiar with the format of the test: What sorts of questions are asked, how is information conveyed, and how are answers logged in? This knowledge will reduce the level of surprise and confusion that robs the test taker of time she could be using to answer questions. The second strategy is to research the content coverage of a test and then to apply the study time the learner has available on the content that will count the most.

With most standardized tests such as the TABE and GED, the format and content information is readily available. Let's take the GED as

an example. It is based on a high school curriculum and performance standards that are used throughout the country. The five tests are in a multiple-choice format (except for one essay), and have been developed by experts familiar with secondary and adult education. The Language Arts Test emphasizes organizing text and the mechanics of writing. The Mathematics Test includes computational problems and real world problems and applications. The test will give you any formulas you will need to use.

Calculators are used with one of the sections. Some math answers are multiple choice, but many are marked on little "bubble charts." The Social Studies Test draws content from United States and World History, Government, Economics, and Geography. That test contains at least one excerpt from a major historical document, such as the Declaration of Independence. The Reading Test will have the adult learner read and interpret many different forms and varieties of literature, such as fiction, nonfiction, prose, poetry, and drama from different cultures and time periods, as well as use business-related documents. The Science Test has the test taker interpret and use scientific information in the form of text or graphics, and material from the life sciences or physical sciences. Adult learners might be asked to interpret experimental results or explain how results from a classic study apply to the everyday world. Even more detail on specific GED tests is readily available from the GED testing service and in bookstores.

Adult learners who want to study strategically can use information like that provided above to orient themselves to tests and focus their study time for maximum results. They can easily find out what content is covered by a particular test and how much importance will be given to various topics; for example, geometry makes up a small part of the GED math test. With this kind of information, learners can focus their study time on the most important topics, and when those topics have been mastered, they can move to the less-important ones. In addition, knowing what the most important content areas are can help learners find the right study aids.

#### **4. Become familiar with test-taking techniques!**

Going into a test with a good knowledge of basic test-taking techniques will help a learner to do his best. Much has been written on good techniques; here is a sampling of the most often repeated advice:

- Listen carefully to directions.  
One of the most critical rules for adult learners is to listen carefully to the test directions: How much time is available? How will the test be scored? What advice, if any, is given about when to randomly guess on multiple-choice test

questions? Does the test administrator have any special instructions? Knowing available time allows adult learners to apportion their time so that they don't need to rush to finish at the end. Knowing about scoring also helps with time use: if 50% of the score will be assigned to essays, then test takers should devote 50% of their test time to writing the essays. And as for whether to guess on multiple-choice test questions, the answer depends on how the test items are scored. If there is no penalty for wrong answers, learners would be smart to answer all questions, so when time is about to run out, they should randomly guess at any remaining answers prior to handing in their answer sheets. On the other hand, if there is a small penalty for wrong answers, learners should be encouraged to answer if they can eliminate at least one of the answer choices. Otherwise, guessing has no particular advantage. Concerning special instructions, adult learners must remember to listen carefully: the instructions might include information about the most important questions on the test, whether or not calculators can be used, the desirable length of essay questions, and so on.

- **Scan the test before starting to answer questions.**  
Adult learners must remember to scan the test first to get an idea of length and difficulty. If the test is made up of multiple-choice questions, they should work on the questions in order and not spend too much time on any one question. Skipping around the test and doing a question here and there is not a good strategy because valuable time is wasted and might lead to errors in marking the answer sheet. If essay questions are part of the test, however, it makes sense to scan these questions and do the easier ones first.
- **Understand a question before answering it.**  
With multiple-choice questions, adult learners must read the questions carefully prior to answering. One of the most common mistakes is not answering the question that is actually being asked. Negative words in the "question stem" can be especially confusing. Sometimes words are highlighted in the question stem and these too are important clues. When in doubt, adult learners should eliminate choices that they know to be wrong, and then choose an answer, at random if necessary, from the remaining choices. Their partial knowledge will be rewarded with such a test-taking strategy.
- **Review the choices.**  
Here are a few additional tips for multiple-choice questions:

(1) Read the question stem, try to think of an answer, and then look for it among the available answer choices. If that doesn't work, at least eliminate the choices that appear to be wrong prior to guessing an answer. (2) If the answer choices are numbers or dates, middle choices are often correct. Note also that longer answers and/or more general answers among the answer choices are more likely to be correct. (3) Sometimes test takers are given a choice among essay questions. Adult learners should be encouraged to watch for this option. Sadly, many test takers fail to heed directions such as, "Answer one of the three questions below" and try to answer all three instead, thus scoring lower than they could have.

- **Be flexible in approaching essay questions.**  
With short answer and essay questions, adult learners should be encouraged to try to write at least something, even if it's just a few sentences. Often partial marks are assigned, so even a partial answer will generate some points. Before starting to write their essay, adult learners should try to prepare an outline. Paraphrasing the question itself is often a great way to start an essay. Clear writing, along with good grammar and spelling, are typically important in the way essays are scored. Adult learners should therefore remember to review their written answers for the use of good sentence structure, grammar, and spelling.
- **Review your work.**  
It's important to remember to review your answers and essays. We all tend to breathe a sigh of relief when the last question has been completed, but adult learners who leave a test with time still available are missing an opportunity to improve their scores. The test is not over until the time is up, or at least until every answer has been checked and essays have been reviewed for grammar and spelling.
- **Stay as calm as you can.**  
Above all, adult learners should stay calm and simply do the best job they can with the time available. Staying calm will make you more efficient while you are answering.

### **5. Take a practice test-or even better, take several practice tests!**

No one learns to fly a plane, drive a car, swim, or play golf just by reading how-to books. Practice makes perfect, as the saying goes, and testing is no exception. There are lots of practice tests available for the GED; in fact, bookstores are full of books containing practice tests for most national standardized tests. However, adult learners

need to take these tests under test-like conditions, and that means with the time limit that will be in place when the test counts. They need to be exposed to some of the natural anxiety that arises when seeing firsthand the test and test question formats. They need to practice their pacing, practice reading the questions and answering them carefully, practice making judgments about when and how to guess, and so on. Of course, these practice tests can be scored, so both weak and strong knowledge and skill areas can be identified. In a sense, every test, whether it is intended for practice or not, provides experience that can help one perform better on future tests. Adult learners can mull over their performance and how they might do better the next time-by being better rested, being more prepared on the content area, making improved use of available time, and so on.

### **6. Read, read, read!**

Studies have shown that vocabulary is one of the most important factors in doing well on standardized tests. Every time a test taker encounters a word he doesn't know, he is less likely to understand a reading passage or a question. It sounds overly simple, but the fact is that vocabulary development is critical to success in all subject areas. The best way to build vocabulary is by reading, reading, and then more reading. Reading shows words in context-that is, how they are really used in sentences to make meaning-and that's the best way to learn them. Adult learners should read in their spare time, read on the bus to work, and read before going to bed...and should try to read for understanding.

### **Summary**

In this article we have tried to give a good overview of standardized testing and provide practical suggestions for helping adult learners demonstrate their knowledge and skills on these tests. Our hope is that when learners are equipped with basic knowledge about these tests and proven test-taking approaches, they will be able to demonstrate what they are truly capable of.

*Ronald K. Hambleton holds the title of Distinguished University Professor and is Chairperson of the Research and Evaluation Methods Program and Co-Director of the Center for Educational Assessment at the University of Massachusetts Amherst. Professor Hambleton has been teaching graduate-level courses at UMass since 1969.*

*Stephen Jirka is a doctoral candidate in the Research and Evaluation Methods Program at UMass Amherst. His current*

*research includes external validation of test scores and standard setting methods.*

Originally published in *Adventures in Assessment*, Volume 16 (Spring 2004),

SABES/World Education, Boston, MA, Copyright 2004.

Funding support for the publication of this document on the Web provided in part by the Ohio State Literacy Resource Center as part of the [LINCS Assessment Special Collection](#).

[Volume 16 Table of Contents](#)

| [SABES Home Page](#) | [SABES Resources](#) | [SABES Calendar](#) | [Other Web Sites](#) |

## Lesson 3 Handout, Article #2

[About SABES](#) [Contact Us](#) [Site Map](#)

	<a href="#">What's New</a>	<a href="#">SABES Resources</a>	<a href="#">SABES Calendar</a>	Top of Form
	<a href="#">Teacher's License</a>	<a href="#">Other Resources</a>	<a href="#">Other Sites</a>	<input type="text"/> <input type="button" value="Search"/> <input type="text" value="sp1001a0c1"/>
				Bottom of Form

# A Basic Primer for Understanding Standardized Tests and Using Test Scores

April Zenisky, Lisa Keller, Stephen G. Sireci

UMass Amherst



[Home](#)

[Volume 16 :](#)

[Spring 2004](#)

[CONTENTS](#)

[How to Do Your Best on Standardized Tests: Some](#)

[Suggestions for Adult Learners](#)

[Using the REEP](#)

[Assessment for ESOL and ABE Classroom](#)

[Instruction](#)

[Integrating Goal](#)

[Setting into](#)

[Instructional Practice](#)

[A Basic Primer for](#)

[Understanding Standardized Tests](#)

[and Using Test Scores](#)

[Using Data for](#)

[Program Improvement](#)

[ACLS, SABES,](#)

[UMASS: Perfect](#)

[Together!](#)

## Introduction

It's nearly impossible to live in American society today without having to take some kind of standardized test. You have to pass a test to get a driver's license, get American citizenship, receive a General Educational Development (GED) certificate, get into college, and be considered for certain kinds of jobs. Here in Massachusetts, our children's teachers have to pass state tests to be licensed, and the children themselves have to pass the Grade 10 version of the Massachusetts Comprehensive Assessment System test, or MCAS, to graduate from high school.

Why do we have to take all these tests? Basically, because there is widespread agreement (but not complete agreement) that tests can tell if a person has the knowledge or skills needed for a diploma, a certificate, a school class level, or a job. But it's not just any test we're talking about here - it's standardized tests. Standardized tests are used because people feel that if you're going to judge someone's abilities, you'd better use a means that's reliable and fair, and standardized tests are designed to be reliable and fair - though people might disagree about whether they succeed in those goals.

We will not debate that issue here; the purpose of this article is to equip readers with a basic understanding of what goes into standardized test making and what test scores purport to show about

learners' skills and abilities. We welcome any constructive use of this knowledge, whether it be better instruction or better policies, but all constructive uses start with accurate knowledge.

### **Meeting the *reliability* and *validity* criteria**

Federal policies now require that states prove that ABE funds result in learner gains in reading, writing, language acquisition, and math. In addition, they require that states measure these gains with valid and reliable tests. After months of reviewing many standardized assessments and their respective alignment with the Curriculum Frameworks, Massachusetts policymakers and education professionals have agreed to use the TABE for ABE Reading, Writing, and Math; the BEST for ESOL Speaking and Listening; and the REEP for ESOL Writing. Scores in each of these tests are meant to represent what students know or can do in those areas. What does it mean when we say these tests are reliable and valid? Let's take up each of these concepts in turn.

### **Reliability**

The consistency of scores across different administrations or scorers is known as reliability. It is crucial that test scores be adequately reliable in representing a person's knowledge and skills. Some level of error is always a factor in testing (more on this later) and test scores. If a person takes the same test on different days, we expect the results to be slightly different, but the more error there is in the test's make-up, the more different the two test scores are likely to be. If the two test scores are very different, it is reasonable to conclude that the difference is due to test error and that the scores do not really reflect what the test taker knows and is able to do.

Inconsistencies in scoring tests might also undercut reliability. Some tests are composed of multiple-choice questions, while others require that the test taker construct a response, such as an essay. Scoring a multiple-choice question is straightforward, because there is one right answer; the answer provided is either correct or incorrect. Therefore, regardless of who scores the test, the score on that question will be the same. Essay-type questions, however, require human judgment and are therefore more difficult to score. If two people read the same essay, it's likely that each person will give the essay a slightly different score. However, if the two scores given by the two scorers, or "raters," are very different, then the score on that essay is not very consistent, or reliable.

The measure of consistency between scorers is called inter-rater reliability. The closer the scores assigned to an essay by different raters, the higher the inter-rater reliability of that test. While it might

seem impossible to get different raters to assign exactly the same score, it is possible to train raters so that they all score in a very similar way. If this goal is accomplished, there can be more confidence that the score assigned to the essay reflects the ability of the student.

---

### **INTERPRETING TEST SCORES RULE 1**

*Reliability of a given test is associated with the consistency and stability of test scores. Look for test reliability statistics to be higher than 0.80. When reliability is high, teachers and students can have greater confidence that students' scores on that test are an accurate reflection of how they would do if the test was administered over and over, or if different people scored the test.*

---

## **Validity**

How do we know whether a test measures the ability we are interested in? Even if a test is perfectly reliable and virtually error-free, how do we know if it is measuring the abilities we want it to and not something else? This is the central concern of validity, and ultimately involves the kinds of judgments that can be based on test scores.

Let's consider a math test consisting only of word problems. The test score could appropriately be used to indicate the student's ability to solve math problems that require reading; that would be a valid use of the test score. However, using the test score as a representation of the student's math ability in general would not be valid.

People who develop tests analyze them in several ways to determine the appropriate (i.e., valid) use of test scores. Let's review some of the issues considered in determining the valid use of test scores:

- Do the questions on the test represent the entire subject matter about which conclusions are to be drawn? For instance, if a test is designed to measure general arithmetic ability, there should be questions about addition, subtraction, multiplication, and division. If there are no questions about division, the test does not measure the entire content of arithmetic, so the test score cannot be said to reflect general arithmetic ability.
- Is the student required to demonstrate the skill that the test is intended to measure? Tests should be directly targeted to the skills measured and that skill should affect test performance. For example, a test designed to measure writing proficiency should ask test takers to write something, and better writers should be shown to receive higher scores.
- Are the test scores consistent with other indicators of the same knowledge and skills? Suppose a student takes a test designed to measure writing ability. If the student does well on writing assignments in class, then he or she should also do well on the writing test, so long as the type of writing on the test is consistent with that done in class. On the other hand, students who do not perform well on writing assignments in class should not do as well on the test. The validity of using that test score as an indication of the person's ability is questionable if there is inconsistency between the score and classroom performance.

## Using test scores

By itself, a test score is just a number. Elsewhere in this issue, you'll discover how teachers are finding ways to apply elements of goal setting and assessment to classroom practice; our purpose here, however, is to provide readers with a basic understanding of standardized tests and scoring. When teachers, students, and others who use test scores are looking at a test score for a particular student, there are an additional few pieces of information they can use to make that number mean something. In the next few pages, some of these pieces of information are explained to help you understand what test scores do and do not mean.

---

## **INTERPRETING TEST SCORES RULE 2:**

*The questions on the test should be representative of the skill or knowledge being measured, and the test should give the student reasonable opportunity to demonstrate what he or she knows about the subject. For each individual student, the test score should provide a result that is similar to the results of other indicators of performance that measure the same thing.*

---

### **Test score scales**

A *score scale* is the range of possible scores on the test. Score scales come in all shapes and sizes. On the TABE, for example, different students might get scores as divergent as 212 and 657. In contrast, on the REEP, scores range only from 0 to 6. A student who takes the BEST, depending on his or her ability to comprehend and speak English, will score from 0 to 65 or higher. Is a 212 on the TABE a "better" score than a 5.4 on the REEP? Even though 212 is a bigger number, these two scores come from tests that are very different and are designed to test very different things. For this reason, comparing scores across different tests is generally not a good idea.

Because scores from the REEP, the TABE, and the BEST are all on different score scales, the number a person gets as a score on one of those tests has meaning for that test only. It might be confusing to have different score scales, but the people who develop tests do this on purpose to make sure that users do not interpret scores on a particular test according to some other standard or yardstick. For example, in the United States the score scale of 0 to 100 is commonly used in many classrooms, but people who make

standardized tests often avoid that score scale because many people would assume that such scores mean the same thing they do in the classroom. Sometimes test developers work really hard to create a unique score scale: e.g., on one test used in the United States for admission to medical school, scores are graded from J (the lowest score) to T (the highest score)!

---

**INTERPRETING TEST  
SCORES RULE 3:**

*Find out what the range  
of possible scores is  
for the test you are using.  
Knowing how high and  
how low scores can be  
for a particular test is  
important to understand-  
ing students' scores.*

---

## **Error in test scores**

As we explained at the beginning of this article, some error is always a factor in test score interpretation. In fact, tests simply cannot provide information that is 100% accurate. This might sound surprising, but this is true for many reasons; for example:

- The extent to which a student has learned the breadth and depth of a subject will influence how she or he performs on a test. On a reading test, for example, a student might do well with questions about word meaning and finding the main idea of a passage but have had less practice distinguishing fact from opinion. The experience (or lack thereof) that a test-taker brings to the test represents a source of error in terms of using the test score to generalize about the student's reading ability.
- Sometimes a student taking a test is just plain unlucky. If a student is tired, hungry, nervous, or too warm, he or she might do worse on the test than if the circumstances were different.
- A test might have questions that seem tricky or confusing. If a student is not clear about the meaning of a question, he or she will have trouble finding the correct answer.
- As we mentioned earlier, mistakes may be made in scoring a

test. When students are not given credit for correct answers or are given credit for incorrect answers, score accuracy suffers.

## **Standard error of measurement**

The score a person gets on the test is meant to indicate how well that person knows the information being tested. One way of looking at a test score is to think of it as consisting of two parts. One part represents the real but unknowable true ability of a person. This part is unknowable because it is never possible to get inside someone's head and have a perfect measure of their ability in the area of interest. The other part of a test score represents the error, all the things that make the test a less-than-perfect snapshot of someone's knowledge at one moment in time. Unlike the way we can manufacture a yardstick that is exactly three feet long to measure length, even the best tests can provide scores that are only approximations of the true ability.

Unfortunately, it is impossible to break these two pieces of a test score (the true ability and error) apart. But it is important to understand that any test score contains a certain amount of error, and as we've illustrated the error might be due to things that are going on with the test taker or things that involve how the test is created or scored. Errors in test scores cannot be completely eliminated, but fortunately there are techniques that can be used to provide some idea about how much the score is affected by error.

For example, testing specialists can calculate the standard error of measurement, which can be thought of as the range of scores obtained by the same person taking the same test at different times. The standard error of measurement is a "best guess" about how close the test is to measuring a person's knowledge or skill with 100% accuracy. The standard error of measurement is a statistical estimate of how far off the true score the test score is likely to be.

Let's take the TABE as an example. Suppose a student takes the TABE Reading Test, Level 7E and gets a score of 447. First of all, that score isn't very low or very high. The next piece of information that will be helpful in understanding this TABE score is the standard error of measurement. The statistics of test development have shown that the standard error of measurement associated with 447 is 17 points, which means that the student's true score is probably between 430 and 464. This score range was calculated by adding and subtracting 17, the standard error of measurement, from the score of 447. The standard error of measurement gives us a good idea of score accuracy.

In the last example the true score was described as probably falling within 17 points of the score the student got on the test; for a score of 630 on the same test, the standard error of measurement is a much bigger number: 64. In this case, the student's true score falls between 566 and 694. There is probably a very big difference in TABE reading knowledge between a 566 score and a 694, so it would be harder to interpret a student's knowledge within such a large range. The size of the standard error of measurement is in large part dependent on the reliability of the test, which was explained previously.

---

**INTERPRETING TEST  
SCORES RULE 4:**

*For standardized tests,  
look in the technical  
manual and find out  
the standard error of  
measurement. If it seems  
like a small number  
relative to the test score  
scale, you can be more  
confident in the accuracy  
of the test score than  
if it is a big number  
relative to the test score  
scale.*

---

## Conclusions

Concepts like reliability, validity, test score scales, and standard error of measurement give meaning to numbers that on their own might not mean much. Of course, the score that someone gets on a test is just one piece of information that tells what he or she knows and is able to do in one very specific and carefully defined subject area. While tests and test scores are important, and it is important to try your best on any test you take, it is also important to remember that any one test score is just that: one test score. The sidebar rules for interpreting test scores given in this article might help you use test scores in meaningful ways.

Are all tests as good as they should be? Do all tests provide useful information? Unfortunately the answer is "no," but researchers at UMass, working in collaboration with the Massachusetts Department of Education, Adult and Community Learning Services, are striving to create tests for ABE students that produce scores that are reliable and can help us make valid decisions about students and programs.

Our efforts are focused on making sure the numbers that are test scores - whether from the REEP, the BEST, the TABE, or any new tests that will be developed - are as meaningful and dependable as possible.

*April L. Zenisky is Senior Research Fellow and Project Manager at the Center for Educational Assessment at UMass Amherst. Her research interests include computer-based testing methods and designs, applications of item response theory, and innovations in test item formats.*

*Lisa A. Keller is Assistant Professor in the Research and Evaluation Methods Program and Assistant Director of the Center for Educational Assessment at UMass Amherst. Her research interests include Bayesian statistics, computerized adaptive testing and item response theory.*

*Stephen G. Sireci is Associate Professor in the Research and Evaluation Methods Program and Co-Director of the Center for Educational Assessment in the School of Education at UMass Amherst. He is known for his research in evaluating test fairness, particularly issues related to content validity, test bias, cross-lingual assessment, standard setting, and sensitivity review.*

Originally published in *Adventures in Assessment*, Volume 16 (Spring 2004),

SABES/World Education, Boston, MA, Copyright 2004.

Funding support for the publication of this document on the Web provided in part by the Ohio State Literacy Resource Center as part of the [LINCS Assessment Special Collection](#).

**[Volume 16 Table of Contents](#)**

| [SABES Home Page](#) | [SABES Resources](#) | [SABES Calendar](#) | [Other Web Sites](#) |

## The Most Frequently *Unasked* Questions About Testing<sup>1</sup>

Stephen G. Sireci  
*University of Massachusetts Amherst*

Many people ask questions about tests, but they frequently ask the wrong ones. Many people criticize tests, but their criticisms often display their lack of knowledge about the strengths and limitations of tests. Currently, tests are a critical part of educational reform movements throughout the world. However, the testing enterprise is complex and involves scientific and psychological concepts that are not readily accessible to lay audiences. This complexity, in part, may foster a mistrust of educational testing. Most people do not like to take tests, and for many people, test results consistently provide bad news. For these reasons, educational tests are popular targets of criticism.

I am a psychometrician, which means I help build and evaluate tests. My specialty area is educational testing, which includes tests that children take in school, college admissions tests, and tests adults take to get licensed to work in a profession. Many people may find it hard to believe that I justify my work by saying it contributes to the educational process. I believe tests, if properly developed, scored, and interpreted, can improve instruction and help people to achieve their goals. However, I realize that many people do not see things as I do.

One day, as I was driving home from work feeling good about my psychometric self; I saw a bumper sticker with the acronym MCAS stricken out and the caption "These tests hurt kids!" To give you some background, MCAS stands for the Massachusetts Comprehensive Assessment System, which is a set of tests administered to public school students in Massachusetts. I teach at the University of Massachusetts and assist the state in MCAS research from time to time. As I stared at the bumper sticker I had the terrible thought that these people were out to get me. Why do they think these tests hurt kids, and why would they go to the trouble of printing bumper stickers to voice their outrage? Soon after, I became outraged when I learned that the union to which I belong, and to which I pay \$600 annual dues, spent \$600,000 on anti-MCAS commercials.

---

<sup>1</sup> Sireci, S. G. (2005). The most frequently unasked questions about testing. In R. Phelps (Ed.), *Defending standardized testing* (pp. 111-121). Mahwah, NJ: Lawrence Erlbaum.

From these experiences, I decided to try and find out what the criticisms were all about. I began to ask people with these bumper stickers (and t-shirts!) why they thought these tests hurt kids. Some reasons seemed well founded (e.g., administering these tests takes up valuable instructional time) and some did not (e.g., these tests are just used to make minority students look bad). I also followed debates in the popular press and the news. Soon, one thing became very clear. There were fundamental problems with many of the criticisms because the general public has very little knowledge about the qualities of a good test. It appears that many of the criticisms of tests are based on half-truths, at best. I admit there are good tests and there are bad tests. I also know that sometimes, good tests are put to bad uses. However, distinguishing between appropriate and inappropriate test uses is not easy. To do it properly, one must ask the right questions. Unfortunately, the right questions are frequently unasked.

In this chapter, I focus on questions about tests that are rarely asked, but if answered, provide a strong foundation for understanding test quality. By answering these unasked questions, we empower the general public to properly scrutinize the tests that end up on their children's desks. The purpose of this chapter is to discuss important questions that should be raised when critiquing any test that has important consequences for the individuals who take it or for those who use test scores to make "high-stakes" decisions. It is only after consideration of these issues that we can form intelligent opinions about the quality and appropriateness of tests.

### THE SIX MOST FREQUENTLY UNASKED QUESTIONS

There are at least six questions that are critically important for understanding test quality that are rarely asked. These questions are:

1. What is a *standardized* test?
2. What is the difference between a norm-referenced test and a criterion-referenced test?
3. What is *reliability*?
4. What is *validity*?
5. How are passing scores set on a test?
6. How can I get more information about a test?

These questions may seem basic to some and technical to others. In either case, answers to these questions must be well understood to properly critique a particular test for a particular purpose. Therefore, let us ask and answer these questions in turn.

## What is a Standardized Test?

The term, *standardized test* has quite possibly made more eyes glaze over than any other. Standardized tests have a bad reputation, but it is an undeserved one. People accuse standardized tests of being unfair, biased, and discriminatory. Believe it or not, standardized tests are actually designed to promote test fairness. *Standardized* simply means that the test content is equivalent across administrations and that the conditions under which the test is administered are the same for all test takers. Thus, standardized tests are designed to provide a level playing field. That is, all test takers are given the same test under the same conditions. I am not going to defend all standardized tests, for surely there are problems with some of them. The point here is that just because a test is standardized does not mean that it is “bad,” or “biased,” or that it measures only “unimportant things.” It merely means it is designed and administered using uniform procedures. Standardized tests are used to provide objective information. For example, employment tests are used to avoid unethical hiring practices (e.g., nepotism, ethnic discrimination, etc.).

If an assessment system uses tests that are not standardized, the system is likely to be unfair to many candidates. Those of you who had the pleasure of viewing the old film *Monty Python and the Holy Grail* saw an excellent example of a non-standardized assessment. In this film, a guardian protected a bridge by requiring that three questions be answered before a traveler was permitted to cross. Answering any question incorrectly resulted in being catapulted into the abyss. Thus, this was a high-stakes test. The three questions asked of the first traveler were: What is your quest? What do you seek?, and What is your favorite color? The second traveler was asked the same first two questions, but was then asked for the air speed velocity of a swallow: Clearly, the two assessments were not of equal content or difficulty. The second traveler who plummeted into the abyss had a legitimate claim of test bias.

On the other hand standardized tests are designed to be as similar as possible for all test takers. The logic behind standardization stems from the scientific method. Standardize all conditions and any variation across measurements is due to differences in the characteristic being measured, which in educational testing is some type of knowledge, skill, or other proficiency. To claim that a test is standardized means that it is developed according to carefully designed test specifications, it is administered under uniform conditions for everyone, the scoring of the test is the same for everyone, and different forms of the test are statistically and qualitatively equivalent. Thus, in testing, standardization is tantamount with fairness.

Before leaving our discussion of standardized tests, there are two common misconceptions about standardized tests that need to be dismissed. The first misconception is that standardized tests contain only multiple-choice items. Obviously, uniform content, test administration, and scoring conditions pertain to all item types, not just multiple-choice items. The second misconception is that

standardized tests measure only lower level thinking skills. That is simply not true, and if you do not agree, sign up to take the Graduate Records Exam or the Law School Admissions Test.

At this juncture, it should be clear to what the term *standardized test* refers and to what it does not refer. Thus, we will now ask and answer the second most frequently unasked question.

### **What is the Difference Between Norm-Referenced and Criterion-Referenced Tests?**

The terms *norm-referenced* and *criterion-referenced* are technical and represent one reason why people accuse psychometricians of speaking in an incomprehensible language. These terms refer to very different ways in which meaning is attached to test scores. That is, they refer to different ways in which the tests are *referenced* to something.

In norm-referenced testing, a person's test score is compared to (referenced to) the performance of other people who took the same test. These other people are the "norm group," which typically refers to a carefully selected sample of students who previously took the test.

There are several types of norm groups, the most common being national, local, and state. *National norms* refer to a nationally representative sample of test takers. This sample of students is carefully selected to represent key demographic characteristics of our nation. Local norms usually refer to the entire population of students within a school district. For example, local norms on an eighth-grade test would be used to compare one eighth-grade student's score with all other eighth-grade students in the district who took the same test. State norms are used in the same manner, with students' scores being referenced to all other students across the state who took the same test.

Many scores reported on norm-referenced tests focus on percentiles, which reflect the percentage of students who scored at or below a specific score. For example, if Fiona achieved a national percentile rank of 94, she performed as good as or better than 94% of the national norm group to whom the test was initially administered. Another way of describing her test performance is that she scored in the top 6% of the national norm group. If her local percentile rank were 87, then she scored as good as or better than 87% of the students in her district who took the same test (in the same grade level, of course). Other scores associated with norm-referenced tests include grade equivalent scores, which represent the average score of students in a particular grade at a particular point in time. For example a grade equivalent score of 3.0 represents the average score of third-grade-students at the beginning of the school year who served as the norm group.

The utility of norm-referenced test information lies in ascertaining how well a particular student, classroom, school, or state compares to others. For example, if I

want to know how well my son performed on a reading comprehension test, his national percentile rank score will tell me the percentage of the national norm group that he performed as good as or better than. When evaluating the utility of norm-referenced test information, the recentness of the norms, and the degree to which they are representative of the group they claim to represent (e.g., all fourth graders in the nation) are important factors. If my son's national percentile rank score of 84 were based on norms from 1994, I may be less impressed than if the norms were more recent.

A serious limitation of norm-referenced scores is that in many cases it is less important to know how well a student did relative to others, than it is to know what a student has or has not learned. For this reason, criterion-referenced tests are much more popular today. Rather than reference a student's test score to the performance of other students, criterion-referenced tests compare students' test performance with carefully defined standards of expected performance. Examples of criterion-referenced scores are classifications such as *pass*, *fail*, *needs improvement*; *basic*, *proficient*, and *advanced*. The information derived from criterion-referenced tests allows one to gauge whether a student mastered specific course material. The standards of *mastery* specified in criterion-referenced testing are typically decided upon by groups of subject matter experts (e.g., experienced and well respected teachers and administrators).

To summarize our discussion of norm- and criterion-referenced tests, it should be clear that norm-referenced tests describe a student's test performance in relation to the performance of one or more specific reference groups of students who took the same test. The group to which a student's performance is compared is called the *norm group*. The statement that Carlos scored at the sixtieth percentile on the ITBS math subtest means that Carlos equaled or outperformed 60% of students in the norm group.

Criterion-referenced tests, on the other hand, describe what a student can and cannot do with respect to a well-defined domain of knowledge or skill. Thus, students' scores are interpreted in terms of mastery of the components of this well-defined domain. The statement that Carlos demonstrated mastery of solving algebraic equations, suggests that he attained the knowledge and skill intended in this portion of the school curriculum. In norm-referenced testing, it is critical that the norm group is appropriate for the students tested. In criterion-referenced testing, it is critical that the criterion domain is clearly defined.

### **What is Reliability? What is Validity?**

Reliability and validity are two important concepts that apply to all types of tests. Many people often ask if a test is reliable or if a test is valid, but few people ever ask "what do these concepts mean?" In fact, when people ask me if a test is

“reliableandvalid,” they often say it as a single word, which leaves me wondering if they realize they are two very different concepts.

*Reliability* refers to the degree to which test scores are consistent. For example, if you took a test on a Monday and received a score of 80, and then took the same test on Tuesday and received a score of 50, the scores produced by this test are certainly not reliable. Your bathroom scale is reliable. If you weigh yourself, step off the scale, and then weigh yourself a second time, you should get the same reading each time. Such physical measurements are often very reliable. Psychological measurements, such as measuring a teacher candidate’s readiness for teaching, are a little trickier. A person’s test score could be influenced by the particular sample of questions chosen for the test, how motivated or fatigued the person is on the testing day, distracting test administration conditions, or the previously ingested extra large pastrami sandwich that suddenly causes trouble during the middle of the test. A great deal of statistical theory has been developed to provide indices of the reliability of test scores. These indices typically range from zero to one, with reliabilities of .90 or higher signifying test scores that are likely to be consistent from one test administration to the next. For tests that are used to make pass/fail decisions, the reliability of the passing score is of particular importance.

The reliability of a test score is an important index of test quality. A fundamental aspect of test quality is that the scores derived from the test are reliable. Readers interested in the technical details regarding test score reliability should consult any standard measurement textbook such as Anastasi (1988) or Linn and Gronlund (2000).

Validity is different from reliability. This concept refers to the soundness and appropriateness of the conclusions that are made on the basis of test scores. Examples of questions pertaining to test score validity include “Is this test fair?,” “Is this test measuring what it is supposed to measure?,” and “Is this test useful for its intended purpose?” Validity refers to all aspects of test fairness. It is a comprehensive concept that asks whether the test measures what it intends to measure and whether the test scores are being used appropriately.

The validity of test scores must always be evaluated with respect to the purpose of testing. For example, the Scholastic Achievement Test (SAT) is designed to help college admissions officers make decisions about who should be admitted to their schools. The validity of SAT scores for this purpose has been supported by studies showing the ability of SAT scores to predict future college grades. However, some people question the utility of using the SAT for a different purpose: to determine whether a student athlete should be eligible to play sports in college. Using test scores for purposes other than what they were originally intended for requires additional validity evidence.

Another way of thinking about validity is the degree to which a test measures what it claims to measure. For educational tests, this aspect of test quality is often described

as *content* validity. Content validity refers to the degree to which a test represents the content domains it is designed to measure. When a test is judged to have high content validity, the content of the test is considered to be congruent with the testing purpose and with prevailing notions of the subject matter tested. Given that educational tests are designed to measure specific curricula, the degree to which the tests match curricular objectives is critical. Thus, in educational assessment content validity is a fundamental characteristic of test quality (see Sireci, 1998a, 1998b for further discussion of content validity; for more comprehensive treatment of validity issues in testing see Kane, 1992, Messick, 1989, or Shepard, 1993).

To distinguish between reliability and validity, I often tell the following story. Although many people have trouble losing weight, I can lose 5 pounds in only 3 hours. Furthermore, I can eat whatever I want in this time period. My secret is simple. I weigh myself on my bathroom scale, and then I drive 3 hours to my mother-in-law's house. Upon arrival, I weigh myself on her bathroom scale and, poof!, I'm 5 pounds lighter. I have accomplished this weight loss many times and weighed myself on both scales repeatedly. In all cases, I have found both scales to be highly reliable. Although I hate to admit it, one of these scales is reliable, but probably not valid. It is biased. It systematically underestimates or overestimates my true weight.

Before concluding our discussion of reliability and validity, a few further clarifying words about these concepts are in order. First, reliability and validity do not refer to inherent properties of a test, but rather to properties of test scores, or even more precisely, to the decisions that are made on the basis of test scores. Second, as mentioned earlier, the validity of decisions made on the basis test scores must be evaluated with respect to the purpose of the test. Third, to claim that scores from a test are valid for a particular purpose requires multiple types of evidence—some statistical and some qualitative. Finally, evaluating validity and reliability is an ongoing endeavor. There is not any one study that could validate the use of a test for a particular purpose from now until infinity. Therefore, test validation is comprehensive and continuous.

### **How Are Passing Scores Set on a Test?**

Test takers are ecstatic when they pass a test and upset when they fail. However, unless passing rates for a test are very surprising, the question of how the passing score was determined is rarely asked. Setting a passing score on a test is perhaps the most difficult problem in the entire testing enterprise. No matter which method is used, some test takers who are competent will fail, and some who are not yet competent will pass. The goal in setting a passing score is to minimize these two types of errors.

Before describing defensible ways in which passing scores are set on a test, a few words about how passing scores should not be set are in order. For example, it may be

*Testing – Standardized and Otherwise – and the Massachusetts ABE Pilot Test: A Curriculum Unit for ABE Classes.* By Jenny Lee Utech (Massachusetts Worker Education Roundtable), with Ruth Byrne and Susan Prior (Methuen Adult Learning Center), for the Massachusetts Department of Education ACLS, March 2005.

tempting to decide upon a passing score that would pass or fail a specific percentage of test takers. Using this logic a score that fails the bottom 30% of test takers could be used to screen out the “worst” candidates for licensure in a profession. However, if all test takers were competent, this procedure would fail 30% of those who are competent. On the other hand, if all test takers were not competent, 70% of those who are incompetent would pass! This type of *norm-referenced* passing score is not defensible for high school graduation, licensure, or certification tests because scores are interpreted with respect to being better or worse than others, rather than with respect to the level of competence of a specific test taker.

Passing scores should also not be set based on arbitrary notions of how many items a student should answer correctly. For example, on the PBS program *The News Hour With Jim Lehrer*, John Silber, former Chair of the Massachusetts Board of Education, exclaimed that the passing scores set on the Massachusetts Teacher Tests corresponded to correctly answering 70% to 75% of the questions. He described this passing standard as a C average. It is hard to know whether the 70% to 75% standard really corresponds to what most people would consider to be a C. If the test comprised very difficult questions, 70% correct could represent a level of excellence well above a C level. As an example, an unethical person could develop a test comprising only extremely difficult questions to prove the point that teachers were performing poorly. On the other hand, another unethical person could intentionally create a test comprising very easy items to make it appear that teachers are performing extremely well. In this case, 70% correct may represent a very low standard such as what most people would consider to be a D or F. Thus, establishing passing scores based on arbitrary notions of percent correct is indefensible, and is rarely used in practice. The mistake of equating percentage of items answered correctly with school grades should not be made.

In elementary and high school, our teachers decided what constituted an F or an A and so forth. When the stakes are high, such as in the case of teacher licensure, the judgments of a single person cannot be used to establish a passing score or any other type of test-based standard. The most defensible procedures currently used to set passing scores on tests are done by committee. Experts from a variety of perspectives are brought together to scrutinize all test questions and make judgments regarding the likely performance of competent test takers on each question. These judgments are discussed thoroughly and often statistics summarizing test takers’ performance on the items are used to inform the process. The passing scores that are set using this process are often considered to be preliminary. They are sometimes adjusted using statistical analyses that attempt to validate the standard and on the basis of anticipated social and political consequences. A thorough discussion of methods for setting passing scores is beyond the scope of this chapter. However, regardless of the method used, it is critical that the entire process is carefully documented so that the public can judge the appropriateness of the methodology. In fact, clear documentation of the procedures used for setting passing scores is a fundamental requirement stipulated in professional standards for educational and psychological testing.

*Testing – Standardized and Otherwise – and the Massachusetts ABE Pilot Test: A Curriculum Unit for ABE Classes.* By Jenny Lee Utech (Massachusetts Worker Education Roundtable), with Ruth Byrne and Susan Prior (Methuen Adult Learning Center), for the Massachusetts Department of Education ACLS, March 2005.

## How Can I Get More Information About a Test?

It is unfortunate that this last question made the list of frequently unasked questions because it is perhaps the most important question a test user or test taker can ask. Tests can be confusing. Sometimes the directions are confusing, sometimes the scores are confusing, and sometimes the content is confusing. Sometimes, the whole process is confusing! Professional standards for testing require test developers to provide documentation to interested parties regarding what the test measures, how the test was developed, how to complete the test, and how to interpret test scores (American Educational Research Association (AERA), American Psychological Association, & National Council on Measurement in Education, 1999). Furthermore, these standards require test developers to provide evidence of the reliability and validity of test scores with respect to the testing purpose. Good test developers are proactive and provide a great deal of technical and non-technical information about their tests. This information can be invaluable to those who want to evaluate the merits of a particular test for a particular purpose. To acquire information about a particular test, start by contacting the test publisher. If you do not know who published the test, ask the person or organization that administered the test.

There are also several reference books available that provide critiques of popular tests, so, ask your local librarian. Examples of such books include the *Mental Measurements Yearbook*, *Tests in Print*, and *Test Critiques*. The first two books are published by the Buros Center for Testing ([www.unl.edu/buros](http://www.unl.edu/buros)). The third book is published by Test Corporation of America.

A very important source of information for evaluating tests is the *Standards for Educational and Psychological Testing* (AERA et al., 1999). The original version of these standards dates back more than 50 years. The *Standards* stipulate standards of quality in test development, use, interpretation, and evaluation. Given that there is no “policing” organization for the testing industry, these standards have become the authoritative source for gauging how to ensure quality test development and how to evaluate whether a particular test is suited for a particular purpose. They are also useful for determining if specific testing practices are defensible. Anyone interested in challenging the use of a test for a particular purpose will want to be familiar with these standards, as will anyone involved in testing who strives to ensure quality and fairness.

## CONCLUSION

This chapter may not have provided complete and satisfactory answers to all the important questions that could be raised about tests, but at least now you know more about the right questions to ask. The next time you receive a score from a test you took, knowing the answers to one or more of these questions may help you better

*Testing – Standardized and Otherwise – and the Massachusetts ABE Pilot Test: A Curriculum Unit for ABE Classes.* By Jenny Lee Utech (Massachusetts Worker Education Roundtable), with Ruth Byrne and Susan Prior (Methuen Adult Learning Center), for the Massachusetts Department of Education ACLS, March 2005.

understand the meaning of your performance. Good tests help people make important decisions. Good tests also tend to be accompanied by adequate documentation that provides answers to the important, but often unasked, questions raised above. Understanding fundamental concepts in educational assessment helps us realize the strengths and limitations of such assessments. Educational measurement is an inexact science, but one that is improving steadily. By engaging more people into the test development and test monitoring processes, we can accelerate the process of making our tests more valid and more useful.

## REFERENCES

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.
- Kane, M. T. (1992). An argument based approach to validity. *Psychological Bulletin*, 112, 527-535.
- Linn, R. L., & Gronlund, N. E. (2000). *Measurement and assessment in teaching* (8th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-100). Washington, DC: American Council on Education.
- Shepard, L. A. (1993). Evaluating test validity. *Review of Research in Education*, 19, 405-450.
- Sireci, S. G. (1998a). Gathering and analyzing content validity data. *Educational Assessment*, 5, 299-321.
- Sireci, S. G. (1998b) The construct of content validity. *Social Indicators Research*, 45, 83-117.

## Handout 6a. Strategies Checklist - instructions

Create your own checklist of strategies to prepare for and take standardized tests:

- In column 1, write a list of your own test-preparation and test-taking **strategies** that you think will work for you.
- You can use the ideas your class discussed, plus your own ideas.
- When you are finished with your list of strategies, share it with your partner.
- Then, for each strategy on your list, discuss with your partner **how** you are going to accomplish this strategy (column 2), and **whose / what help** you need in order to do that (yourself, teacher, spouse, kids, friends) (column 3).
- After you have discussed your ideas for **how** and **what help** for each strategy, write them in columns 2 and 3.



## Handout 6b. Strategies Checklist (version 2) - instructions

This is a list of strategies you can use to prepare for and take standardized tests.

- Read the list of strategies in column 1.
- Add your own strategies to the list in column 1.
- Then, for each strategy on your list, discuss with your partner **how** you are going to accomplish this strategy (column 2), and **whose / what help** you need in order to do that (yourself, teacher, spouse, kids, friends) (column 3).
- After you have discussed your ideas for **how** and **what help** for each strategy, write them in columns 2 and 3.

## Handout 6b. Strategies Checklist (version 2)

1. Strategies	2. How are you going to accomplish this strategy?	3. Whose help do you need to accomplish this strategy?
Get positive about taking tests.		
Clear your brain for studying and learning (have quiet time each week to study and learn.)		
Have quiet time right before taking a test.		
Become familiar with the test format. (What kinds of questions are asked? How are answers recorded?)		
Use your study time wisely. (Find out what content is most important and spend the most time studying that content.)		
Practice test-taking techniques, for example: Listen carefully to directions.		
Scan the test before you start to answer questions.		
Understand a question before you answer it.		

## Handout 6b. Strategies Checklist (version 2) page 2

1. Strategies	2. How are you going to accomplish this strategy?	3. Whose help do you need to accomplish this strategy?
Review the answer choices before you choose one.		
Be flexible when answering essay questions.		
Check over your work before time is up.		
Stay calm.		
<b>Other strategies:</b>		

These strategies are adapted from Hambleton, R.K. and S. Jirka (2004). "How to do Your Best on Standardized Tests: Some Suggestions for Adult Learners." In *Adventures in Assessment* 16, <http://sabes.org/resources/adventures/index.htm>

## Handout 6c. Strategies Revisited

Now that you have taken the pilot test, look over the strategies checklist you created in class. For each strategy you listed:

- Write the strategy in column 1.
- Think about how this strategy work for you, or not, and why. Then write your ideas in column 2.”
- Think about how you would revise (change) this strategy for future tests, or not. Write how you would change the strategy in column 3. If you would not change it, write “no change.

Strategy	How did this strategy work for you (or not)? Why?	How would you change this strategy (or not) for future tests?

## Handout 7. Pilot Tests

### What is a pilot test?

Test makers who are creating a new kind of test need to know if the new test is **fair** and **reliable**. To find out, they give the new test to many people to see how well it works, and to see if the questions are clear, fair, too easy or too hard. This is called a **pilot test**. Test makers do a pilot test to try out the new test and see how it works. After the pilot test, they will look at everyone's test results and change parts of the test that did not work well.

### What is the pilot test in Massachusetts?

In Massachusetts, test makers at the University of Massachusetts are creating a new kind of test for students of Adult Basic Education. Eventually, we hope to use this new test instead of the TABE to test students' reading and math skills. We hope that this new test will measure students' learning achievements in reading and math better than the TABE does. By taking the pilot test, students are helping to make the test better.

Students will take the pilot test on computers. Everyone will have the chance to practice on the computer before they take the test.

### How will the pilot test work?

#### In March 2005

- Teachers will introduce the pilot test to students.
- Teachers and students will talk about test taking in class.
- Teachers will get training on how to give the pilot test.
- Teachers will help students practice for the test with the computer tutorial.
- Students around Massachusetts will start taking the pilot test.

#### From March to December 2005

- Students will take 3 different pilot tests.
- Students will still take the TABE during this time.

### What will happen to the tests I take?

The test makers at the University of Massachusetts, Amherst will use your test results to decide which test questions are the most clear and fair, and the best for testing your achievements in reading and writing.

## Handout 8a. Pilot Tests - Matching Sentence Halves (front halves)

Test makers who are creating a new kind of test ...

Cut sentences into strips .....

Test makers do a pilot test ...

.....

After the pilot test ...

After the pilot test ...

In Massachusetts, test makers are creating a new kind of test ...

We hope to use the new test in Massachusetts instead of ...

We hope that the new ABE test will measure ...

Teachers and students will talk about ...

Teachers will help students practice ...

Test takers will use your test results to decide ...

## Handout 8b. Pilot Tests – Matching Sentence Halves (back halves)

... need to know if the new test is fair.

Cut sentences into strips.....

... to try out a new test and see how it works.

.....

... they will look at everyone's test results.

... they will change parts of the test that did not work well.

... for students of Adult Basic Education.

... the TABE to test students' reading and math skills.

... students' achievements in reading and math.

... test taking in class.

... for the new test with the computer tutorial.

... which test questions are fair and clear.

## Handout 8c. Pilot Tests – Sentence Halves Matched

Test makers who are creating a new kind of test ... need to know if the new test is fair.

Test makers do a pilot test ... to try out a new test and see how it works.

After the pilot test ... they will look at everyone's test results.

After the pilot test ... they will change parts of the test that did not work well.

In Massachusetts, test makers are creating a new kind of test ... for students of Adult Basic Education.

We hope to use the new test in Massachusetts instead of ... the TABE to test students' reading and math skills.

We hope that the new ABE test will measure ... students' achievements in reading and math.

Teachers and students will talk about ... test taking in class.

Teachers will help students practice ... for the new test with the computer tutorial.

Test takers will use your test results to decide ... which test questions are fair and clear.

## Handout 9a. Story – The Pilot Test

Carmen Valdez was a student in an adult education program in her hometown. She and her classmates had been studying with Miss Brown for several months. Carmen felt comfortable in her school surroundings. Her class followed a regular routine, and she felt confident that school was helping her to improve in her skills every day. She looked forward to coming to class and learning something new with people whom she trusted.

Carmen had even become accustomed to taking standardized tests periodically so that from time to time, she and her teacher would be able to measure her progress. Carmen considered this to be a real accomplishment because, before coming to the adult education center, she had dreaded taking any kind of test. She thought with pride, “I’ve come a long way since starting here in September!”

As spring was approaching, Miss Brown told the class that she would be giving them a new test. There were a few groans from the students around her, and Carmen had to admit that she wasn’t expecting to take any test that day, especially a new one! It took her by surprise. She had gotten used to the format of the test that was usually given. Suddenly, she felt anxious. “Now that I feel confident with the old test, they have to go and change it! Why would they do that?” she wondered.

Miss Brown explained that the state’s Department of Education, that funds the classes, was working on a new test to replace the old one. Like technology, the old test needed to be updated. In fact, students would be taking the new test on a computer. She also told the class that teachers around the state have been working on some new test questions. Now the state needs to find

## Handout 9a, page 2. Story - The Pilot Test

out which of the new questions accurately measure what students know. To do this, the Department of Education is piloting, or trying out, the new questions in some programs around the state. Miss Brown's class was being asked to pilot the new test.

Carmen asked, "Will the questions be harder?" Miss Brown replied that some of the questions might seem more difficult, but not to worry because the state would look at the results to determine which test questions should remain and which ones should be eliminated. This is part of the process of creating a standardized test.

By the time Carmen completed the pilot test, she had plenty of feedback, or valuable information, to give her teacher. This was a new experience for her, for she had never taken a test on a computer before. She felt valued because she was able to contribute to an effort to benefit all adult learning programs in the state.

## Handout 9b. Cloze Exercise - The Pilot Test (word box)

### WORD

accomplishment  
accurately  
benefit  
computer  
confident  
contribute

eliminated  
feedback  
format  
measure  
mouse  
periodically

pilot  
process  
routine  
standardized  
technology  
test

## Handout 9b. Cloze Exercise – The Pilot Test

**Directions:** Complete the following story using the words from the word box.

The (1)\_\_\_\_\_ Test

Carmen Valdez was a student in an adult education program in her hometown. She and her classmates had been studying with Miss Brown for several months. Carmen felt comfortable in her school surroundings. Her class followed a regular (2)\_\_\_\_\_, and she felt confident that school was helping her to improve in her skills every day. She looked forward to coming to class and learning something new with people whom she trusted. Carmen had even become accustomed to taking (3)\_\_\_\_\_ tests (4)\_\_\_\_\_ so that from time to time, she and her teacher would be able to measure her progress. Carmen considered this to be a real (5)\_\_\_\_\_ because, before coming to the adult education center, she had dreaded taking any kind of test. She thought with pride, “I’ve come a long way since starting here in September!”

As spring was approaching, Miss Brown told the class that she would be giving them a new (6)\_\_\_\_\_. There were a few groans from the students around her, and Carmen had to admit that she wasn’t expecting to take any test that day, especially a new one! It took her by surprise. She had gotten used to the (7)\_\_\_\_\_ of the test that was usually given. Suddenly, she felt anxious. “Now that I feel (8)\_\_\_\_\_ with the old test, they have to go and change it! Why would they do that?” she wondered.

Miss Brown explained that the state’s Department of Education, who funds the classes, was working on a new test to replace the old one. The old test was good enough when it first came out, but, like (9)\_\_\_\_\_, it needed to be updated. In fact, students would be taking the new test on a

(10)\_\_\_\_\_ using a (11)\_\_\_\_\_. She also told the class that educators around the

## **Handout 9b, page 2. Cloze Exercise - The Pilot Test**

state have been working on some new test questions to develop an even better assessment. Now the state needs to find out which of the new questions (12)\_\_\_\_\_ (13)\_\_\_\_\_ what students know. To do this, the Department of Education is piloting, or trying out, the new questions in some programs around the state. Miss Brown's class was being asked to pilot the new test.

Carmen asked, "Will the questions be harder?" Miss Brown replied that some of the questions might seem more difficult, but not to worry because the state would look at the results to determine which test questions should remain and which ones should be (14)\_\_\_\_\_. This is part of the (15)\_\_\_\_\_ of creating a standardized test.

By the time Carmen completed the pilot test, she had plenty of (16)\_\_\_\_\_, or valuable information, to give her teacher. This was a new experience for her, for she had never taken a test on a computer before. She felt valued because she was able to (17)\_\_\_\_\_ to an effort to (18)\_\_\_\_\_ all adult learning programs in the state.