

Chapter 4

Assessment Practices

In this chapter

- Introduction
- The Rationale for Standardized Assessment
- Mandated Tests in New York State
- Educational Functioning Levels
- Pre- and Post-testing
- Developing an Effective Testing Schedule
- Assessment Guidelines
- Qualifications for Testing Staff

Introduction

Assessment refers to the collection of information—using specially designed instruments—regarding a student’s initial skill level and the development of his or her literacy and language skills as a result of instruction. The *NRS Implementation Guidelines* state:

- ✘ “At intake, an individual learner’s educational functioning level is determined within the functional level descriptors, using a uniform, standardized assessment procedure approved by the state. The assessment procedure may be a standardized test or a performance assessment with standardized scoring protocols.”
- ✘ “To determine gain, the learner should be assessed again at least once after a standard instructional period, at the end of the class or at the end of the program year, as determined by state policy.”

As part of its effort to comply with federal National Reporting System (NRS) guidelines, the New York State Education Department (NYSED) mandates that adult education programs use standardized tests to determine students’ initial skill levels, as well as the educational gain they achieve as a result of their participation in a program. Standardized tests are evaluative instruments that feature uniform questions, answers, and administration procedures. Standardization allows for the comparison of program results from across the state and the nation.

This chapter will clarify the NYSED’s assessment policies. Programs are required to administer state-approved tests in order to report results according to NRS guidelines; therefore, programs should use these policies as a basis for assessment. However, programs are free to perform additional forms of assessment as they see fit in response to the needs of their students.

The Rationale for Standardized Assessment

Standardized tests are used to:

- Determine the student’s skill level at intake. Assessment done during the student intake process helps determine the instructional setting in which a student will be placed. (The educational levels for Adult Basic Education ((ABE)) and English as a Second Language ((ESL)) will be discussed in detail later in this chapter.) The intake assessment provides the basis for establishing educational gain, which is determined by comparing students’ future skill levels to those ascertained during the initial intake.
- Assess educational gain. Students should be tested at regular intervals to determine if their reading, math, or English skills are improving.

- Report results. The NRS guidelines mandate that the only way a program can report educational gain is by demonstrating that a student has moved from one NRS level to the next based on his or her standardized assessment scores.
- Guide instruction. In addition to providing a summary assessment of a student's skill level in the form of a grade equivalent or NRS level, standardized instruments can be used as a diagnostic tool to determine a student's strengths and needs and to design instruction accordingly. Test results can provide students with valuable feedback on their own progress.

Mandated Tests in New York State

The instrument used to assess educational skill levels should correspond to the instruction a student receives. The NYSED-approved tests for adult education programs in New York State are as follows:

Instruction Type	Testing Instrument
Basic Education	TABE Reading (complete battery and survey) 7/8 and 9/10 for all levels
	TABE Math (complete battery and survey) 7/8 and 9/10 for all levels
	READ for BE levels 1 & 2 (levels 1-3 for volunteer based programs)
	TABE Reading (level L) for levels 1 & 2 (levels 1-3 for volunteer based programs)
English as a Second Language	BEST Plus (ESLOA for volunteer-based programs only)

Please note the following policies regarding assessment:

- ✓ When using the Test of Adult Basic Education (TABE), programs may administer either the survey or the complete battery version of the test. NYSED recommends that programs use the complete battery for diagnostic and accountability purposes and the survey for accountability if other methods are in place for diagnostic purposes.
- ✓ When using TABE Math, programs should administer both sub tests (Computation and Applied) and combine the scores to produce one scale score or grade equivalent.
- ✓ Administration of the TABE Language and Spelling are not necessary for NRS purposes. However, agencies may wish to use them for diagnostic purposes.
- ✓ The following assessments previously used in NYS programs can no longer be used for NRS reporting purposes: TABE Reading and Math 5/6, TABE Español, CTBS, GED Practice, John, and SABE.
- ✓ On July 1, 2003 (the first day of fiscal year 2004), the NYSED began its transition to establish BEST Plus as the standard test for ESL students. By July 1, 2004, all programs should be testing all new students using the BEST Plus. Students who have been pre-tested on the NYS PLACE should be post-tested on the same instrument. However, students with NYS PLACE pre-tests who continue attending classes after January 1, 2005 should be re-tested using the BEST Plus. Students re-tested after January 1, 2005 will only have six months to demonstrate gain. For this reason, it is recommended that students be re-tested with the BEST Plus as soon as agency staff members are trained.

Educational Functioning Levels

The NRS defines six functioning levels each for ABE and ESL. During the intake process, programs should use standardized tests to place students in one of these levels. Students should be assessed periodically during the course of their participation in a program to determine whether they have acquired enough literacy or English skills to advance to the next NRS level. The percentage of students who complete one level and move on to the next level is an important NRS measure.

The NRS levels and associated assessments are summarized below. A detailed description of each level can be found in the *NRS Implementation Guidelines*; they are not presented here.

Instruction Type	Test Used	Ranges	
ABE Beginning Literacy	- TABE Reading 7/8 and 9/10 for all levels	Grade Equivalents	0 – 1.9
ABE Beginning Basic Education			2 – 3.9
ABE Intermediate Low	- TABE Math 7/8 and 9/10 for all levels		4 – 5.9
ABE Intermediate High			6 – 8.9
ASE Low	- READ for BE levels 1 & 2* - TABE Reading (level L) for BE levels 1 & 2*		9 – 10.9
ASE High			11 – 12.9
ESL Beginning Literacy	- BEST Plus for all levels - ESLOA for all levels*	BEST Plus Scores	Below 401
ESL Beginning			401 – 438
ESL Intermediate Low			439 – 472
ESL Intermediate High			473 – 506
ESL Low Advanced			507 – 540
ESL High Advanced			Above 540

*Exceptions for Volunteer Based Programs

- ✓ The READ test and the TABE Reading (level L) may be used for NRS levels 1, 2, and 3 for volunteer-based programs only—i.e., they can be used through the 6th grade reading level.
- ✓ Only volunteer-based programs may use the English as a Second Language Oral Assessment (ESLOA) test. (Other programs must receive authorization from the NYSED.) The ESLOA score ranges and corresponding NRS levels are as follows:

ESLOA Test Level	ESLOA Test Score	NRS Level
1	0 – 7	1
1	8 – 10	2
2	1 – 9	2
2	10 – 20	3
3	1 – 11	4
3	12 – 15	5
4	1 – 12	6

Pre- and Post-testing

To monitor progress, students should be tested at regular intervals during the course of their participation in a program. The pre-test, which determines a student's initial skill level, is the first test a program should administer.

- ✓ The pre-test must be administered before the student completes 12 hours of instruction.

The skill level ascertained by the pre-test provides the basis for all future evaluations. Student progress and educational gain is determined by comparing future skills levels to those ascertained during the pre-test. Subsequent tests given during the fiscal year are called post-tests. While students are allowed to take an unlimited amount of post-tests, they may take only one pre-test of any given type. For example, an ABE student is allowed to have only one TABE Reading pre-test score and one TABE Math pre-test score—not two TABE Reading pre-test scores or two TABE Math pre-test scores. However, he or she may have several TABE Reading post-test scores, each representing the result from a test administered at a different time during the year, and each indicating the student’s skill level at that specific point in time.

- ✓ Students enrolled in distance education programs must be tested using the same instruments and following the same procedures as all other students—i.e., distance education students must be physically present when taking standardized assessment tests at a designated testing site.

Assessments should be consistent with a student’s educational goals. For instance, a student who wants to improve his or her English language skills should not be administered the TABE Reading for NRS purposes because that test will not help program staff determine the student’s initial knowledge of the English language or assess the student’s progress.

To determine educational gain for a fiscal year, a student’s earliest and lowest pre-test score will be compared to his or her highest post-test score of the same type. For example, if a student has both a TABE Reading and a TABE Math pre-test score, and the Math score is lower, his or her gain will be determined by comparing the TABE Math pre-test to the highest TABE Math post-test. The following are additional considerations that may affect educational gain calculations:

- ✓ A pre-test score that was obtained more than six months prior to the beginning of a fiscal year is considered “stale”; i.e., it is not a valid indicator of the student’s initial skill level for that fiscal year. For example, results from a pre-test administered before January 1, 2003 are not valid for FY 2004.
- ✓ The TABE Reading (level L) may be used to place students in the ABE Beginning Literacy and Basic Education levels (NRS levels 1 and 2) and to demonstrate gain from level 1 to level 2. However, the student will need to be pre- and post-tested on other levels of the TABE (E, M, or D) to indicate further educational progress. Therefore it is recommended that students who test into Level 2 using the TABE Reading (level L) be retested using the TABE Reading (level E).
- ✓ The READ test may be used to place students at the ABE Beginning Literacy and Beginning Basic Education levels (NRS levels 1 and 2) and to demonstrate gain from level 1 to level 2. However, the student will need to be pre- and post-tested on the TABE (E, M, or D) to indicate further educational progress.
- ✗ For students who have multiple pre- and post-tests, their earliest and lowest pre-test score will be compared to their highest post-test score in order to measure gain.
- ✗ Students whose pre-test places them in the ASE High level (NRS level 6) must obtain a GED or high school diploma in order to complete that level. Receiving a GED is the automatic goal for all students who are placed at that level.
- ✗ Students whose pre-test places them in ESL High Advanced (NRS level 6) must obtain the maximum score on the BEST Plus in order to complete that level.
- ✗ All students pre-testing on NRS level 5 (Low ASE) or below will be counted in the denominator for educational gain. Educational gain can only be demonstrated by pre- and post-testing with the TABE. Passing the GED test will not be accepted for demonstrating gain.

Developing an Effective Testing Schedule

Devising an effective testing schedule is critical to program success. Being able to post-test a high percentage of students is important for two reasons. First of all, post-test scores determine educational gain, which is an important NRS measure. The percentage of students post-tested is used to measure student retention: A low percentage indicates that a program is unable to retain its students long enough for them to be post-tested.

There is no formula or universal testing schedule that can be applied to all adult education programs. In fact, programs are encouraged to customize their testing schedules for each of their various offerings. When developing a testing schedule, programs should consider the following questions:


- *How long do students stay enrolled in our program?* If a program waits too long to post-test its students, it may measure a lower student retention percentage because some of its students may have left the program before they've had a chance to be tested. However, programs should be careful not to test students too close to the beginning of the term—students may not have attained enough skills to demonstrate gain. Nor should they test students too frequently—excessive testing can be both enervating and discouraging.
- *What is the intensity of our classes?* A class that is more intense—i.e., a class that offers more frequent and extensive sessions in a given time period—may be more effective in helping its students achieve educational gain than one that meets less frequently for shorter sessions. Therefore, programs should schedule post-tests accordingly.

✓ The table below lists minimum requirements for post-testing schedules:

Intensity of Program	Testing Done Every
Nine or fewer hours per week	100 instructional hours
Ten or more hours per week	200 instructional hours
Tutorial program	50 student contact hours

Assessment Guidelines

For testing to serve as an effective indicator of a student's educational progress, it must be implemented with care and competence. This manual does not provide step-by-step instructions for administering the TABE and BEST Plus. That information can be found in each test's administration manual or at training sessions scheduled by NYSED and Regional Adult Education Network (RAEN) across the state.

 *Test score reliability.* This term refers to the degree to which test scores are free from errors of measurement. Errors of measurement can occur when programs do not adhere to the following recommendations:

- Assessment tests should correspond to a student's educational goals as well as the instruction he or she receives. For instance, TABE is not an appropriate test to measure the educational gain of an ESL student. Similarly, the BEST Plus is useless for assessing the reading skills of an ABE student. (For more information about evaluating appropriate assessment tests, see the "Mandated Tests in New York State" section of this chapter.)
- A student should be assessed at a test level that corresponds to his or her level of skill. Determining the appropriate test level will depend on the test used. For instance, the computer version of the BEST Plus provides questions based on how the student has answered previous questions: The more questions the student answers correctly, the more difficult the questions become. Programs using the paper-based version of the BEST Plus, however, should make sure that the version of the test they administer matches the skill level of the student taking the test.

- ✓ For TABE Reading and Math, programs must administer a TABE locator to determine the level at which students are to be tested. Using a test that does not match a student's skill level will provide unreliable results. Consequently, students may be placed in the wrong educational setting, and their post-test scores may be lower than their pre-test scores.
 - ✓ At the very minimum, all programs must purchase the E(asy), M(oderate), and D(ifficult) levels of the TABE. Each level corresponds to a different skill range. Therefore, using the E level of the TABE would be inappropriate for students at the High ASE level, whereas the D level would yield unreliable results for a student at the lowest literacy level.
 - ✓ Agencies are expected to provide reliable data for use in program evaluation. The State Education Department will publish specific guidelines regarding the range of scores on each of the TABE subtests that are acceptable. Students who test outside of the acceptable range on the TABE subtests must be retested on a higher or lower level.
 - ✓ Programs are allowed to compare pre-tests and post-test scores using different levels of the TABE as long as the levels represent contiguous skill ranges. For example, a student may be pre-tested on TABE E(asy) and post-tested on TABE M(oderate).
 - ✓ Programs must use different versions of an assessment when pre- and post-testing their students. If a program administers the same version, students might remember items from the previous testing session, which in turn might distort the results. Therefore, if a TABE 7 is administered as a pre-test, then a TABE 8 should be administered as a post-test.
- Programs are allowed to use other tests as part of their comprehensive testing strategy; however, to report results according to NRS guidelines, they are required to administer state-approved tests. (For a list of state-approved tests, see the "Mandated Tests in New York State" section of this chapter.)
 - Standardized instruments must be administered in accordance with the procedures listed in the test administration manual—i.e., the administration of assessment tests must be uniform. This allows for the comparison of test scores from around the state as well as the nation. If uniform procedures are not followed, scores may become unreliable. For example, a variance in an administration procedure might affect a student's performance; consequently, his or her test score might not accurately reflect his or her skill level.
 - Programs that offer both reading and math instruction as part of their ABE curriculum must administer both the TABE Reading and the TABE Math. However, programs that teach only reading should administer only the TABE Reading, and programs that teach only math should administer only the TABE Math.

Qualifications for Testing Staff

Programs are required to identify staff members who are responsible for administering the standardized tests in a systematic manner. In order to be allowed to administer the tests, these staff members must receive training that is conducted by either the test publisher, NYSED, LAC or RAEN.