BEFORE YOU READ . . .

The purpose of Chapter 14 is to discuss how and why standardized tests are used, and how scores on these tests can be interpreted and combined with other assessment techniques to make important educational decisions such as assigning grades. Information regarding standardized testing and grading is listed below.

Standardized tests are uniform in content, administration, and scoring; therefore, they allow for the comparison of results across classrooms, schools, and school districts.

Standardized tests measure aptitude, intelligence, and achievement as well as diagnose learning difficulties.

Standardized tests are reported in terms of percentiles, grade equivalents, and standard scores.

Standardized tests must have content, predictive, and construct validity; and, they must be reliable.

CHAPTER OUTLINE

WHAT ARE STANDARDIZED TESTS AND HOW ARE THEY USED?
A. Selection and Placement
B. Diagnosis
C. Evaluation
D. School Improvement
E. Accountability

WHAT TYPES OF STANDARDIZED TESTS ARE GIVEN?
A. Aptitude Tests
B. Norm-Referenced Achievement Tests
C. Criterion-Referenced Achievement Tests
D. Standard Setting

HOW ARE STANDARDIZED TESTS INTERPRETED?
A. Percentile Scores
B. Grade-Equivalent Scores
C. Standard Scores

WHAT ARE SOME ISSUES CONCERNING STANDARDIZED AND CLASSROOM TESTING?
A. Test Validity
B. Test Reliability
C. Test Bias
D. Computer Test Administration

WHAT TESTING INFORMATION IS AVAILABLE ON THE INTERNET?
AS YOU READ . . .

WHAT ARE STANDARDIZED TESTS AND HOW ARE THEY USED?

1. What is the main difference between standardized and non-standardized tests? How are standardized test results used in student selection, placement, diagnosis, and evaluation?

2. What is a minimum competency test, and how does it hold teachers and schools accountable for what students learn?

WHAT TYPES OF STANDARDIZED TESTS ARE GIVEN?

3. Define aptitude test and achievement test and describe how each is measured.

4. What is the difference between a norm-referenced achievement test and a criterion-referenced achievement test? Give an example of an appropriate use for each type of test in the space provided.
HOW ARE STANDARDIZED TESTS INTERPRETED?

5. Describe how standardized test percentile scores, grade-equivalent scores, and standard scores are derived and interpreted. The diagram below will help with your answer.

WHAT ARE SOME ISSUES CONCERNING STANDARDIZED AND CLASSROOM TESTING?

6. Define validity and reliability. What criticisms of standardized tests relate to issues of validity and reliability?

Validity:

Reliability:

WHAT TESTING INFORMATION IS AVAILABLE ON THE INTERNET?

7. On a computer connected to the Internet, open the web browser, type in the following URL, and locate the ERIC Clearinghouse on Assessment and Evaluation: http://www.ed.gov. (Hint: you will have to follow the link from the first URL to find the ERIC system.). In the space below, list some of the testing information and statistics you can find regarding the grade level of your future students. You may also wish to locate specific information for your state.
SELF-CHECK

DIRECTIONS: Below are questions related to the main ideas presented in the chapter. Correct answers or typical responses can be found at the end of the study guide chapter.

1. What are the advantages and disadvantages of minimum competency testing?

2. Which of the following types of standardized tests is designed to predict future performance?
   A. placement test
   B. achievement test
   C. aptitude test
   D. diagnostic test

3. Currently intelligence is measured as a
   A. ratio of chronological age to mental age.
   B. percentile.
   C. mean of 100 and a standard deviation of 15.
   D. percentage of items correct on an IQ test.

4. Which of the following interpretations would apply to a sixth-grade student who has scored at the mean on a standardized test?
   A. percentile = 90; stanine = 9; z = 20
   B. NCE = 50; z = 0; percentile = 50
   C. GE = 7.2; stanine = 5; NCE = 45
   D. z = 1; NCE = 60; percentile = 50

5. Two classes of math students average 75 on a test, but the students in one class have scores that are much more spread out; this means their results will have a larger
   A. mean.
   B. median.
   C. standard deviation.
   D. normal curve.

6. A seventh-grader has a grade-equivalent score of 9.4 on a standardized test. Which of the following interpretations can be made from these results?
   A. The student is ready for ninth grade work.
   B. The test was too easy and should be renormed.
   C. The student has done as well as an average ninth-grader.
   D. The student scored at the 9.4 percentile.
7. Which of the following terms refers to a measure of consistency in test results over multiple applications?
   A. predictive validity
   B. content validity
   C. construct validity
   D. reliability

8. Match each validity type with its description.

   _____ a measure of the match between the test items and instruction given
   A. construct validity

   _____ a measure of the ability of a test to forecast behaviors
   B. predictive validity

   _____ a measure of the degree to which test scores reflect its intention
   C. content validity

9. What are the major criticisms of standardized tests?

10. What testing resources are available on the Internet?
AFTER YOU READ . . .

PRACTICE EXAM

DIRECTIONS: Each chapter heading is listed below and followed by a series of related questions worth a total of ten points. Respond to each question, check your answers with those found at the end of the chapter, then determine your score. Consider nine points per heading to be mastery.

WHAT ARE STANDARDIZED TESTS AND HOW ARE THEY USED?

Short Answer/Essay
1. (10 points) Describe five uses of standardized tests.

WHAT TYPES OF STANDARDIZED TESTS ARE GIVEN?

Matching
2. (3 points) __________ test designed to predict the ability of students to learn or perform particular types of tasks
A. aptitude tests

__________ test of a student’s knowledge of a particular content area in which her or his scores are compared with others who were tested
B. norm-referenced tests

__________ test of a student’s knowledge of a particular content area in which his or her scores are compared to well specified skills
C. criterion-referenced tests

Short Answer/Essay
3. (2 points) List two types of aptitude tests.
4. (3 points) List three types of norm-referenced achievement tests.

5. (2 points) Describe two ways that criterion-referenced achievement tests differ from norm-referenced achievement tests.

HOW ARE STANDARDIZED TESTS INTERPRETED?

6. (1 point) ___________ are students’ raw scores translated into percentiles, grade equivalents, or normal curve equivalents.

7. (1 point) ___________ are scores reported as the percentage of students in the norming group who scored lower than a particular score.

8. (1 point) ___________ are scores reported as the average scores obtained by students at particular levels of achievement.

9. (1 point) ___________ Which of the following terms refers to the statistical measure of the degree of dispersion in a distribution of scores?
   
   A.  standard deviation  
   B.  raw score  
   C.  percentile score  
   D.  mean

10. (1 point) ___________ What type of curve is produced by a frequency graph of a normal distribution?
   
   A.  a curve that is skewed left  
   B.  a curve that is skewed right  
   C.  a bell-shaped curve  
   D.  a curve that is depressed at the mean

11. (1 point) ___________ IQ scores are usually presented in which of the following manners?
   
   A.  with a mean of 10 and a standard deviation of 1  
   B.  with a mean of 50 and a standard deviation of 5  
   C.  with a mean of 100 and a standard deviation of 15  
   D.  with a mean of 500 and a standard deviation of 50
CHAPTER 14

Educational Psychology: Standardized Tests

12. (2 points)  Explain “stanines” and how they are calculated.

13. (2 points)  Explain normal curve equivalents.

WHAT ARE SOME ISSUES CONCERNING STANDARDIZED AND CLASSROOM TESTING?

14. (1 point)  _________ The term “validity” refers to the results of a test measuring the type of information purported.

15. (1 point)  _________ The term “reliability” refers to the accuracy of the test results.

16. (3 points)  _________ type of validity evidence that relates a test to other similar measures A. content validity

__________ type of validity evidence that relates a test to future performance B. predictive validity

__________ type of validity evidence that relates a test to the objectives of a lesson, unit, or course C. construct validity

17. (4 points)  List four major criticisms of traditional multiple choice standardized tests.

18. (1 point)  Describe a type of authentic assessment.
WHAT TESTING INFORMATION IS AVAILABLE ON THE INTERNET?

Short Answer/Essay

19. (10 points) List five different testing and educational statistics resources available on the Internet.

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________
Now check your answers and total your points. Answers are found at the end of the study guide chapter.

<table>
<thead>
<tr>
<th>SCORING</th>
<th>POINTS NEEDED FOR MASTERY</th>
<th>POINTS RECEIVED</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHAT ARE STANDARDIZED TESTS AND HOW ARE THEY USED?</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>WHAT TYPES OF STANDARDIZED TESTS ARE GIVEN?</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>HOW ARE STANDARDIZED TESTS INTERPRETED?</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>WHAT ARE SOME ISSUES CONCERNING STANDARDIZED AND CLASSROOM TESTING?</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>WHAT TESTING INFORMATION IS AVAILABLE ON THE INTERNET?</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>
SELF-CHECK ANSWERS  Chapter 14

1. Minimum competency tests focus on important skills students are expected to have mastered to qualify for promotion. Advantages of minimum accountability testing include increased pressure on schools and teachers to pay attention to students who might otherwise be overlooked. Another advantage is that accountability encourages schools to search out improved instructional methods and guarantees routine evaluation of any innovation they try. However, the accountability movement has its critics. Many argue that minimum competency testing focuses schools on minimums rather than maximums. Others are concerned that schools will teach only what is tested. Many educators point out that accountability assessments fail to take into account differences among students. A school or classroom may test low because the students are from disadvantaged backgrounds rather than because they were given poor instruction. High-stakes testing can lead schools and districts to adopt policies that artificially inflate scores.

2. C. Aptitude tests are designed to predict future performance. It is meant to predict the ability of students to learn or perform particular types of tasks rather than to measure how much the students have already learned.

3. C. Over the years the chronological age/mental age comparison has been dropped, and IQ is not defined as having a mean of 100 and a standard deviation of 15. Most scores fall near the mean, with small numbers of scores extending well above and below the mean.

4. B. The national mean on the standardized test is 50, and the percentile of the student’s reported score is 50. A z-score mean is set at zero.

5. C. Standard deviation is a statistical measure of the degree of dispersion in a distribution of scores.

6. C. It can be assumed that the student has scored as well on the test as the average ninth-grader, the equivalent of 9.4, which may be interpreted as ninth grade, fourth month.

7. D. Reliability is a measure of the consistency of test scores obtained from the same student at different times.

8. C, B, A. Content validity is a measure of the match between the content of a test and the content of the instruction that preceded it. Predictive validity is a measure of the ability of a test to predict future behavior. Construct validity is the degree to which test scores reflect what the test is intended to measure.

9. Critics argue that standardized tests: 1) give false information about the status of learning in the nation’s schools, 2) are unfair (or biased against) some students, 3) tend to corrupt the process of teaching and learning, and 4) focus time, energy, and attention on simpler skills that are easily tested.

10. The Internet is a dynamic and ever changing medium that allows teachers and students access to vast amounts of information. Most good testing information is found at the US Department of Education website and the Educational Resources in Education (ERIC) website. Other related websites offer interactive test writing modules, Java applets and objects such as confidence-interval calculators, and so on. Take some time to explore the Internet. It will provide a useful resource for your classroom.
PRACTICE TEST ANSWERS      Chapter 14

1. Selection and placement, diagnosis, evaluation, improvement, and accountability

2. A, B, C; aptitude test, norm-referenced achievement test, criterion-referenced achievement test

3. Intelligence tests and multifactor aptitude tests

4. Achievement batteries, diagnostic tests, and subject area achievement tests

5. Criterion-referenced achievement tests differ from norm-referenced achievement tests in that they are constructed around well-defined objectives, measured by assessing the number of objectives or “skills” met, and use a cutoff score.

6. Derived scores

7. Percentile scores

8. Grade equivalent scores

9. A; Standard deviation

10. C; Bell-shaped curve

11. C; With a mean of 100 and a standard deviation of 15

12. Stanine scores are standard scores that have a mean of five and a standard deviation of two. Each stanine is reported as a whole number and represents .5 standard deviations.

13. Normal curve equivalents can range from 1 to 99, with a mean of 50 and a standard deviation of approximately 21. NCE scores are like percentiles except that the intervals between the scores are equal.

14. True; Valid test results are a measure of the student’s knowledge of some content area.

15. True; Reliable results are consistently accurate.

16. C, B, A; Construct validity evidence, predictive validity evidence, content validity evidence

17. Multiple choice standardized tests give false information about a school’s status, are unfair to some groups of students, corrupt the process of teaching and learning, and attend to simple, easily tested skills

18. Portfolios or performance assessments

19. If you have not yet done so, you should be exploring the Internet for educational resources for lesson planning, informational sites, enrichment ideas, maps, satellite images, and so on for use in your own classroom. Sites for testing information include ERIC, National Center for Education Statistics, Council of Chief State School Officers, Buros Institute of Mental Measurements, and National Assessment of Educational Progress. In addition, you should search your state’s department of education website for information about graduation standards, PRAXIS, INTASC, and Standards of Effective Practice for teachers. A teacher-friendly search engine is http://www.google.com. Type in whole phrases or keywords for a complete search.