UNIT 13 STANDARDIZED ACHIEVEMENT TESTS

Structure

13.1 Introduction
13.2 Objectives
13.3 Standardized Achievement Tests
   13.3.1 Functions of Standardized and Teacher-made Tests
   13.3.2 Standardized Test vs. Teacher-made Tests
   13.3.3 Uses of Standardized Achievement Tests
13.4 Administering and Interpreting Standardized Tests
   13.4.1 Administering Standardized Tests
   13.4.2 Types ofScores/Norms for Interpretation
13.5 Standardized Achievement Test Batteries
   13.5.1 Achievement Test Batteries or Survey Batteries
   13.5.2 Indian Achievement Tests
13.6 Academic Aptitude Tests vs. Achievement Tests
   13.6.1 Aptitude-Achievement Discrepancies
13.7 Let Us Sum Up
13.8 Unit-end Exercises
13.9 Answers to Check Your Progress
13.10 Suggested Readings

13.1 INTRODUCTION

In the field of education, interdependence of teaching, learning and testing is recognised. The first step in the teaching-learning process is defining and determining the objectives of learning and the outcomes to be expected from classroom instruction. In the light of these processes, the main reliance is placed on tests which are constructed by teachers. Through these tests you can see as to how well your students have mastered the unit of instruction. But when we want to compare achievement of an individual with a group, class and school or study the student growth over a period of time to know whether the progress is more or less rapid than might be expected, standardized tests are used. Both the standardized and the teacher-made tests are important. So you are expected to know the list of achievement tests made in the country and abroad. In this unit, we discuss as to why traditionally students have been labeled as underachievers or overachievers on the basis of their academic achievement. You will also learn that as to when and how the achievement tests should be used.

13.2 OBJECTIVES

After going through this unit, you should be able to:

- discuss the concept of a standardized achievement test,
- describe as to how standardized tests could be used,
- differentiate between standardized and teacher-made tests,
- compile a list of available achievement tests batteries, and
- differentiate between achievement tests and aptitude tests.
13.3 STANDARDIZED ACHIEVEMENT TESTS

Achievement tests are generally classified as standardized tests and teacher-made tests. Normally in schools teacher-made tests are used. The teacher-made achievement tests assess student achievement in relation to instruction provided in a school. Standardized tests are constructed by test construction specialists, usually with the assistance of curriculum experts, teachers, and school administrators for the purpose of determining a student's level of performance relative to the performance of other students of similar age and grade. These tests often take long time to construct. These tests are called standardized because they are administered and scored according to specific and uniform procedures. In other words, a standardized test administered and scored in a school would be administered and scored in exactly the same manner as anywhere else in the country.

When standardized tests are employed, test results from different students, classes, schools and districts can be more easily and confidently compared than would be the case with teacher-made tests. Imagine the difficulty in comparing teacher-made tests results from Ms. Sharma's class in New Delhi with Ms. Sundram's class in Chennai. Not only would the test items be different the length of the test, the amount of time allowed, the instructions given by the teachers and the scoring criteria would also be different. In short, there would be little or no basis for comparison. The value of standardized test lies particularly in situations in which comparisons can be made; comparisons of one school with other schools, comparison of achievement in different areas by a student or by a school or comparison of achievement with the potentiality for achievement indicated by an aptitude test. The norms provided with standardized tests make such comparisons as school achievement, readily possible with national norms. Age or grade level in different subjects can be compared. Age or grade level on an achievement test may be compared with age or grade level on a measure of scholastic aptitude.

The distinctive features of a standardized test represent important advantages for some purposes and disadvantages for others. Basing the test upon a careful analysis of the common objectives expressed in textbooks, courses of study, and report of committees of professional bodies should guarantee that the thinking of many specialists has entered into the test plan. However, a published test is fixed for a period of years in terms of broad and common objectives. It is not a flexible tool. It cannot be adapted to special current needs, to local emphasis/conditions; or to particular limited units of study.

13.3.1 Functions of Standardized and Teacher-made Tests

In the light of differences between standardized and teacher-made tests, more emphasis is laid on teacher-made tests when we want to:

- test how well students have mastered a unit of instruction,
- determine the extent to which distinctive objectives have been achieved, and
- provide a basis for assigning course marks.
Standardized test should be used when we wish to:

- compare achievement with potentiality for an individual or a group,
- compare achievement of different skills or subject areas,
- make comparisons between different classes and schools, and
- study student's growth over a period of time to see whether progress is appropriate as expected.

13.3.2 Standardized Tests vs. Teacher-made Tests

Standardized achievement tests differ from the tests that you prepare for your own class. The broad differences are as follows:

i) A standardized test is based on the general content and objectives common to many schools in the country whereas a teacher-made test can be adapted to content and objective(s) taught in classroom.

ii) A standardized test deals with large segments of knowledge or skill, whereas a teacher-made test can be prepared in relation to any specific/limited topic.

iii) A standardized test is developed with the help of professional writers, reviewers and editors of test items whereas a teacher-made test usually relies upon the skill of one or two teachers.

iv) Quality of items in a standardized test is high. They are pretested and selected on the basis of difficulty level and discriminating power. Quality of test items is often unknown in a teacher-made test and is typically lower than that of a standardized test.

v) Reliability of a standardized test is usually high. Reliability of teacher-made test is usually unknown but can be high if test items are carefully constructed.

vi) In standardized test the administration and scoring procedures are standardized and specific instructions for its administration are given. In teacher-made test uniform procedure of administration is possible, but is usually flexible.

vii) Interpretation of scores of a standardized test can be compared to normgroups. Test manuals and other guides aid interpretation and use. In a teacher-made test comparisons and interpretations of scores are limited to local class or school situation.

13.3.3 Uses of Standardized Achievement Tests

Standardized tests are used for comparative purposes. These are quite different from the main uses of teacher-made tests which are to determine student's mastery or skill levels, to assign grades, and to provide students and parents with feedback. Now, a question arises as to why the classroom teacher administers a standardized test? May be to compare the performance of students of the current year with the performance of students of the previous year or to compare class A with class B. But the most appropriate answer is more likely that the classroom teacher administers standardized tests because he/she is required to do so. This is the case in many, if not most of the schools
in the country. Part of the reason for this is the current trend toward increasing accountability which includes evaluation of various state funded programmes. Most, if not all, such programmes require that standardized achievement tests be administered as part of the programme evaluation and further funding may depend on the results of these tests. By doing so we would be able to compare students, schools and districts with each other in order to make judgements concerning the effectiveness of programmes across the schools, districts or states. As long as this objective remains, use of standardized tests will be necessary part of teaching. Hence, you should learn to administer and interpret the results of standardized tests which are sometimes used for evaluating the general educational development of students in the basic skills. They are also used for evaluating student progress during the year or over a period of years and for grouping the students for instructional purposes. These tests can also be used for diagnosing relative strengths and weaknesses of students in terms of broad subject or skill areas.

Check Your Progress 1

Notes: a) Write your answers in the space given below.

b) Compare your answers with those given at the end of the unit.

1. Write two differences between standardized tests and teacher-made tests.

........................................................................................................................................
........................................................................................................................................

2. List three uses of standardized tests.

........................................................................................................................................

13.4 ADMINISTERING AND INTERPRETING STANDARDIZED TESTS

To ensure that the standardized tests are able to serve the specific purpose, we should know the procedure of administering and interpreting standardized tests.

13.4.1 Administering Standardized Tests

The best way to guard against any error in the test administration is to instruct everyone to administer the test as per the directions given in the test. Sometimes, classroom teachers individualize the test administration by helping slower students or pushing faster students. This is a violation of standardized testing procedure. The test and its administration and scoring are called ‘standardized’ because everyone gets the same treatment. Therefore we should follow the requirements of every standardized test:
The following are some of the do’s and don’ts about administering standardized tests.

**Do’s**

- Read the manual before test administration.
- Be sure you have been given the correct form for your grade level.
- Adhere strictly to the administration instructions.

**Don’ts**

- Try to minimize the achievement nature of test.
- Deviate from the standardized administration instructions (i.e. do not allow more time, give hints, spell words, define words, etc.).

### 13.4.2 Types of Scores / Norms for Interpretation

In this sub-section, we shall consider the types of scores offered for standardized tests: grade norms or equivalents, age norms, percentile and standard scores. For example, on a language ability test, if a student aged 15 years scores 76 on the total test score, what can we say about this student? We have the following information in the test manual:

- 90 to 100 : Fluency in language
- 80 to 90  : Command over language
- 70 to 80  : Working knowledge of language
- below 70  : Low language competency

These details will help us in having an assessment of the 15 year old student. These details are also available for other age groups. So there is a need for expressing test performance in a form other than the actual score. Let us elaborate this point for the sake of clarity.

**Grade norms/equivalents:** Grade norms are probably most widely used to report test results. The representative behaviour is presented on yearly basis at the beginning of the grade/class and the average performance during the year can be calculated. These are likely to be most often misinterpreted. Let us consider the following example:

Anil, student of IV grade, obtained a Mathematics computation grade equivalent score of 7.6 (seventh grade, sixth month) on the CAT (California Ability Test) that means even though he is only in fourth grade, he can do seventh-grade level Mathematics. Do you agree with this statement? If you do, you have fallen victim to the most common kind of misinterpretation regarding these grade norms. Anil’s score is the score that the publisher estimates would be obtained by seventh grader during six month. It does not mean that he is ready for seventh grade Mathematics. All we know for sure is that a fourth grader who obtains Mathematics grade equivalent to 7.6 is well above average in Mathematics. So grade norms or equivalents have limitations for (i) extreme ability groups, and (ii) may not be equal due to different school programmes or motivational levels.
Age norms equivalents: Age-equivalent scores are very similar to the grade-equivalent scores. Age-equivalent scores are determined in a fashion similar to that described for grade equivalents. Age norms are based on the average scores earned by students at different age levels and are interpreted in terms of age equivalents. That is, samples of 7-, 8-, and 9-years old might be tested and average scores for each age determined. Scores for younger or older students would then be estimated or extrapolated from these scores. Problems similar to those affecting grade equivalents affect age equivalents also. Main problems are as follows:

- Equal differences in scores may not reflect equal differences in achievement. In other words, does growth from age 6 to age 7 present the same amount of growth as that from age 10 to age 11? It may or may not, depending on the trait being measured. Furthermore, growth in most traits slows down or stops during the teens or early twenties. In other words, a year’s growth in reading after age 17 is likely to be very different from a year’s growth in reading at age 7.

- Age equivalents are only meaningful if subjects are taught across all grades. It makes little sense to say that someone has an age equivalent of 16.9 in subtraction.

- Age equivalents may be misinterpreted as standards, rather than averages or norms.

- Growth across subjects may vary greatly, even if age equivalents show equal growth. A year’s increase in language age equivalent does not necessarily mean the same thing as a year’s increase in science age equivalent.

Unlike grade equivalents, age equivalents have not attracted widespread acceptance in the schools. Like grade equivalents, they are most useful in the elementary grades to compare growth across a common core of subjects. The above mentioned shortcomings should always be considered in interpreting age equivalents.

Percentile ranks: With grade and age-equivalent scores we indicate the grade or age group in which a student’s test performance would be considered average. That is, if a student obtains a grade-equivalent score of 4.5, we can say the student did as well on the test as an average fourth-grader during the fifth month of school. At times, however, we may not be interested in making such comparisons. In fact, we are more interested in determining as to how a student’s performance compares with that of students in his or her own grade or of the same age. Percentile ranks enable us to make such comparisons.

Percentile ranks are a substantial improvement over grade and age-equivalent scores in that they do not suffer from the many limitations of the latter. Since comparisons are within grade, it does not matter whether subjects are taught across grades, and since growth is only relative to others in the grade, the problem of growth being unequal at different grade levels is avoided. In addition, percentile ranks are less likely to be considered as standards for performance. However, percentile ranks do have two major shortcomings which are listed below:
i) Percentile ranks are often confused with percentage correct. In using percentile ranks, be sure you are communicating that a percentile rank of 62, for example, is understood to mean that this individual's score was higher than 62 percent of the people who took the test or, conversely, that 62 percent of those taking the test received scores lower than this individual. Commonly, a score at the sixty-second percentile is misinterpreted to mean the student answered only 62 percent of the items correctly. A score at the sixty-second percentile might be equivalent to a B or a C whereas a score of 62 percent would likely be an F.

ii) Equal differences between percentile ranks do not necessarily indicate equal differences in achievement. In a class of 100 students, the difference in achievement between the second percentile and fifth percentile is substantial, whereas the difference between the forty-seventh and fiftieth is negligible assuming a normal distribution. Interpretation of percentile ranks has to consider that units toward the tails of the distribution tend to be spread out while units toward the center tend to be compressed as illustrated in Figure 13.1.

![Normal Curve (Approximate Percentile Ranks Indicated Along Baseline)](image)

**Fig.13.1: Normal Curve (Approximate Percentile Ranks Indicated Along Baseline)**

**Standard scores:** Like percentile ranks, standard scores compare a student’s performance to that of other students at the same grade level. The problem of equal differences between units not representing equal differences in achievement is overcome through the use of standard scores. You will recall that the z-score is the basic type of standard score, and all other standard scores are derived from it. This is an important consideration to keep in mind since many test publishers create new types of standard scores with various means and standard deviations when they publish new tests (e.g., Developmental Standard Scores on the Iowa Tests of Basic Skills, or Expanded Standard Scores on the Comprehensive Tests of Basic Skills). You need not be overwhelmed by such scores since conceptually they are identical to z-scores.

Though different from z-scores, there are a special type of standard scores called stanines. Stanines are ranges or bands within which fixed percentages of scores fall. They are determined by dividing the normal curve into nine...
portions, each being one-half standard deviation wide. Stanines and the percentage of cases within each stanine are indicated below:

<table>
<thead>
<tr>
<th>Stanine</th>
<th>Percentage of cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4% (lowest)</td>
</tr>
<tr>
<td>2</td>
<td>7%</td>
</tr>
<tr>
<td>3</td>
<td>12%</td>
</tr>
<tr>
<td>4</td>
<td>17%</td>
</tr>
<tr>
<td>5</td>
<td>20%</td>
</tr>
<tr>
<td>6</td>
<td>17%</td>
</tr>
<tr>
<td>7</td>
<td>12%</td>
</tr>
<tr>
<td>8</td>
<td>7%</td>
</tr>
<tr>
<td>9</td>
<td>4% (highest)</td>
</tr>
</tbody>
</table>

Stanines have a mean equal to 5 and a standard deviation equal to 2. Each stanine is one-half standard deviation wide. Interpreting stanines is straightforward in that a student is simply described as being “in” the second stanine, ninth stanine, etc. A major advantage of stanines is that, since they are intervals or bands, they tend to minimize overinterpretation of data. Also, since they only require a single digit number, they are useful for situations where recording space is limited.

Standard scores represent the ultimate in standardized test score interpretation. However, there is one factor that limits their widespread adoption – most educators, parents, and students do not understand how to use standard scores. As a result, few schools or districts request standard scores from test publishers. However, standard scores save time and effort in determining aptitude-achievement discrepancies. They also allow for easy comparison of scores both within and across students either over time or across subjects.

You may keep in mind that such scores are not understood by most parents and students and as a result may not be a convincing way to use standardized test results in reporting them. What should you use, then? In our opinion, grade and age equivalents lend themselves too easily to misinterpretation and have too many limitations. As mentioned, standard scores would be our choice but may be too complicated for use by the general public. We, therefore, recommend that you use percentile ranks when reporting and interpreting standardized test results to parents. Be sure, however, to consider the limitations we mentioned regarding percentile ranks in making such interpretations.

**Check Your Progress 2**

**Notes:**

a) Write your answer in the space given below.

b) Compare your answer with that given at the end of the unit.

1. Which of the norms are recommended for interpreting standardized test results to public?

---------------------------------------------------------------------------

---------------------------------------------------------------------------
The first standardized test came into existence around the turn of the 20th century. These tests were tests of a single achievement area such as spelling. Single subject achievement tests are still used although they are largely confined to the secondary grades.

A variation of the single subject achievement test is the diagnostic achievement test. However, use of the diagnostic test is normally limited to those elementary and secondary school students who experience academic difficulty. These tests are administered to "diagnose" or indicate the specific cause or causes of a problem (for example, faulty letter identification) in some general academic area (for example, reading recognition). Seldom are such tests administered to an entire class or grade. Typically, students are selected for diagnostic testing after a single subject test or an achievement battery has indicated a problem in some general academic area.

The most frequently used type of achievement test is the achievement test battery, or survey battery. Such batteries are widely used, often beginning in the first grade and administered each year thereafter. There are several reasons why survey batteries are more popular than single subject achievement tests. The major advantages of survey batteries over single subject achievement tests are as follows:

- Each subtest is coordinated with every other subtest, resulting in common administration and scoring procedures, common format, and minimal redundancy.
- Batteries are less expensive and less time consuming to administer than several single subject tests.
- Each subtest is normed on the same sample, making comparisons across subtests, both within and between individuals, easier and more valid.

This last point is probably the major reason as to why batteries have come into such widespread use. We often use standardized tests to compare students, classes or schools. It takes less time to make these comparisons when a single norm group is involved than when several are involved. Furthermore, the likelihood of clerical errors is minimized when single, comprehensive score reports from a battery are used to make comparisons, as opposed to several single subject score reports.

Following are the disadvantages of batteries:

- The correspondence (content validity) of various subtests in the battery may not be uniformly high.
- The battery, emphasizing breadth of coverage, may not sample achievement areas in as much depth as a single subject achievement test.

13.5.1 Achievement Test Batteries or Survey Batteries

Various achievement test batteries are available to teachers these days. Some of the important batteries are discussed in this section:
California Achievement Test (CAT): This battery is published by CTB / McGraw Hill. It has five levels appropriate for various grades, and two alternate forms of the test are available. Scores are provided for reading (vocabulary and comprehension) language (mechanics, usage, structure and spelling) and Mathematics (computation, concepts and problems). The CAT has been standardized simultaneously with the Short Form Test of Academic Aptitude, facilitating identification of aptitude achievement discrepancies.

Comprehensive Tests of Basic Skills (CTBS): Like the CAT, the CTBS is published by CTB/McGraw-Hill. However, it is appropriate for students in grades K-12. Seven levels of the test are available for students in the various grades. An alternate form can be obtained. Level A is considered a pre-instructional or readiness test and provides scores for letter forms, letter names, and Mathematics. Level B provides scores for reading, language, Mathematics and Total Battery. Level B is designed to be administered to students who have completed their first year of instruction. The remaining levels, C, 1, 2, 3, and 4, yield scores in reading, language, Mathematics, reference skills (except for Level C), Science, and Social Studies. A total battery score is also provided, composed of reading, language, and Mathematics scores. Like the CAT, the CTBS has been standardized simultaneously with the Short Form Test of Academic Aptitude.

Iowa Tests of Basic Skills (ITBS): This battery is published by the Riverside Publishing Company. It is appropriate for students in grades K-8. The ITBS was normed on the same sample as the Cognitive Abilities Test (CogAT), an academic aptitude test. Thus, determination of aptitude-achievement discrepancies is facilitated when these two tests are used. Scores are provided for listening, word analysis, vocabulary, reading, comprehension, language (spelling, capitalization, punctuation, and usage), visual and reference materials, Mathematics (concepts, problem solving, computation), Social Studies, Science, writing and listening supplements, and basic and total battery.

Metropolitan Achievement Tests (MAT): Harcourt Brace Jovanovich publishes this battery which is appropriate for students in grades K-9. Six levels span the various grades, and two alternate forms of the Primer Level and three alternate forms of the other five levels are available. The Primer level includes scores for listening for sounds, reading, and numbers. The next level, Primary I, includes scores for word knowledge, word analysis, reading, Mathematics computation, and Mathematics concepts. Primary II includes these plus spelling and Mathematics problem solving. The remaining levels all provide scores for word knowledge, reading, language, spelling, Mathematics computation, Mathematics concepts and Mathematics problem solving. In addition, Science and Social Studies scores are available for the two higher levels.

Sequential Tests of Educational Progress (STEP): This battery is published by the Educational Testing Service. It is appropriate for grades 4 onwards and consists of four levels and two alternate forms. Scores are provided at the lowest three levels for English expression, reading, mechanics of writing, Mathematics computation, Mathematics basic concepts, Science and Social Studies. The highest level of the STEP does not include the mechanics of writing and Mathematics basic concepts subtests.
**SRA Achievement Series (SRA):** Published by Science Research Associates, the battery is appropriate for students in grades 1-9. Five levels cover the grade range, and no alternate forms are available. The two lowest levels include subtests for reading (word picture association, sentence-picture association, comprehension, vocabulary), Mathematics (concepts and computation) and language arts (alphabetization, capitalization, punctuation, spelling, and usage). The three highest levels include subtests for reading (comprehension, vocabulary total), language arts (usage, spelling, total), Mathematics (concepts, computation, total), Social Studies, Science and Uses of Sources.

**Standard Achievement Test Series:** Like the MAT, this battery is published by Harcourt Brace Jovanovich. Six levels are provided for the various grades and two alternate forms are available. Subtests for reading, Mathematics, and language arts are available at all levels. Except for the lowest level, scores are also provided for Science, Social Studies, and except at the highest level, listening comprehension. A unique feature of this Test is that it is available as either a basic battery, including only the reading, Mathematics, and language art subtests, or as a complete battery, including all the subtests. Practice tests are also available for all but the highest level.

**Tests of Achievement and Proficiency (TAP):** Published by Riverside Publishing Company, the TAP is appropriate for students in grades 9-12. The TAP is designed to allow for continuity with the ITBS, and also has been normed concurrently with the Cognitive Abilities Test (CogAT), an academic aptitude test. Thus, identification of aptitude-achievement discrepancies is facilitated when these two tests are used. Scores are provided for reading comprehension, Mathematics, written expression, using sources of information, Social Studies and Science; listening and writing supplements are also available.

### 13.5.2 Indian Achievement Tests

Some of the Indian standardized achievement tests are:

- C.I.E. Hindi Yogyata Pariksha for Class VII.
- Mother tongue test (Hindi) – B.K. Srivastava.
- Coimbatore Achievement Test (Social Science) – R.K. Mission.
- School Progress Record – L.N. Dubey.
Check Your Progress 3

Notes:  a) Write your answer in the space given below.

b) Compare your answer with that given at the end of the unit.

1. What are the advantages in using a test battery instead of tests selected from a number of sources?

...........................................................................
...........................................................................
...........................................................................
...........................................................................

13.6 ACADEMIC APTITUDE TESTS vs. ACHIEVEMENT TESTS

So far in this unit we have discussed tests that are used to measure past achievement. The intent of these tests is to identify what students have learned. At times, however, we are also interested in measuring an individual’s potential for learning or an individual’s academic aptitude. Such information is useful in making selection and placement decisions and to determine whether students are achieving up to their potential, that is, to indicate aptitude-achievement discrepancies. In short, aptitude tests are used to indicate aptitude-achievement discrepancies. Aptitude tests are used to predict future learning. While achievement tests are used to measure past learning.

Aptitude: Aptitude and potential can be considered synonymous for our purposes. They refer to the maximum we can expect from a student, as indicated by a student’s score on a test of academic aptitude or potential. Such tests are often referred to as IQ tests or intelligence tests. For now, all we need to know is that such tests provide us with a benchmark or standard against which to compare achievement test scores.

The academic aptitude test provides us with an estimated ceiling for a student’s academic performance. The academic achievement test, on the other hand, measures actual academic performance. Traditionally, students have been labeled overachievers or underachievers based on the relationship between their academic aptitude and academic achievement. Figure 13.2 illustrates an underachiever, an overachiever, and a student achieving at expectancy.

Student A in Figure 13.2 is a student with considerable potential who is not achieving up to his or her potential. Student B is a student with moderate potential who is achieving above his or her potential. More accurately, this “overachiever” is a student whose obtained aptitude score (not necessarily true score) is lower than his or her obtained achievement score. Student C represents a student achieving at the level we would expect, given his or her aptitude score. The obtained aptitude score is equivalent to the obtained achievement score.
Obviously, it is necessary to have an aptitude score to enable you to determine whether a student is achieving at expectancy (the level you would expect given the student’s aptitude). However, school district policies vary in requiring the administration of aptitude tests. Depending on your district, you may have aptitude tests administered to students every few years, or only in the fifth, ninth and eleventh grades, or not at all.

![Figure 13.2: Relative Levels of Aptitude and Achievement for an Underachiever, an Overachiever, and a Student Achieving at Expectancy.](image)

If you find aptitude test scores in your students’ folders, you can use them to enhance your achievement test interpretation. However, be careful not to simply label your students underachievers or overachievers.

Most aptitude tests yield more than one overall IQ score. Many yield a verbal and nonverbal score, or a language and non language score, or a verbal and a quantitative score. Quantitative scores represent general mathematical or number ability. When the aptitude or IQ test yields a verbal score and a nonverbal score or a quantitative score, more relevant comparisons are possible than when only one overall score is reported. Consider the following example:

Mona, a new seventh-grader, obtained the following scores on the cognitive abilities test (an aptitude test) at the beginning of sixth grade. (Note: $X = 100$, $SD = 15$).

Verbal = 100  
Quantitative = 130

Mona’s scores on the California Achievement Test (CAT) given at the end of sixth grade are as follows:

<table>
<thead>
<tr>
<th>Percentile rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Vocabulary</td>
</tr>
<tr>
<td>Reading Comprehension</td>
</tr>
<tr>
<td>Reading Total</td>
</tr>
<tr>
<td>Mathematical Concepts</td>
</tr>
<tr>
<td>Mathematical Computation</td>
</tr>
<tr>
<td>Mathematical Total</td>
</tr>
</tbody>
</table>
Mona’s parents have requested a meeting with you. They want you to push her harder in reading until her reading scores match her mathematics scores which have been superior.

What would you do? How would you interpret Mona’s scores? Would you push her in reading? Before you answer these questions, let’s make Mona’s data interpretable. We can do so by using bar graph comparisons to illustrate the concepts of underachievement and overachievement. This time we will add measurement scales to each histogram.

We know that an obtained IQ score of 100 is at the fiftieth percentile on an IQ test with a mean of 100 and a standard deviation of 15. We also know that on the same test an IQ score of 130 is at the ninety eighth percentile, two standard deviations above the mean. In the graphs in Figure 13.3 our percentile scales do not correspond directly to each other, but this is of little consequence since we are not interested in comparing reading with Mathematics. Rather, we are interested in comparing verbal aptitude with reading total, both of which are on the same scale, and quantitative aptitude with mathematics total which are also on a common scale.

From the graphs, we would conclude that Mona’s obtained Mathematics achievement score actually exceeds the obtained Mathematics aptitude score. According to our popular but somewhat misleading terminology, she is overachieving in Mathematics. Unless she is paying a high price socially or emotionally for working so hard in Mathematics, we see no problem here. Hence, the qualifier “over” in the word overachiever should not imply a negative valuation of what has been accomplished by this student. But Mona’s parents are not concerned with her Mathematics achievement.

<table>
<thead>
<tr>
<th>Verbal Aptitude</th>
<th>Reading Achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>99</td>
</tr>
<tr>
<td>50</td>
<td>98</td>
</tr>
<tr>
<td>40</td>
<td>97</td>
</tr>
<tr>
<td>30</td>
<td>96</td>
</tr>
<tr>
<td>20</td>
<td>95</td>
</tr>
</tbody>
</table>

Fig.13.3: A Comparison of Mona’s Aptitude and Achievement Scores Percentiles

They are concerned with her reading achievement. They want her pushed which suggests that they feel she can do better than she has achieved in the past. That is, Mona’s parents feel she is underachieving in reading. Is she? On the basis of a comparison of her obtained verbal aptitude score and her obtained reading achievement score, our conclusion would have to be “no”. In fact, Mona is “overachieving” in reading, too. That is, her obtained reading achievement score exceeds her obtained verbal aptitude score; she is actually
performing above expectancy. Would you agree that she needs to be pushed? By now, we should hope "not". In fact, you might suggest to Mona’s parents that they ease up, using your skills in the interpretation of standardized test results to substantiate your suggestion.

13.6.1 Aptitude-Achievement Discrepancies

What we have been doing is making a general or global decision about whether or not Mona is achieving at her expected level, as indicated by her aptitude score. In other words, are there any differences between her aptitude and her achievement? When these differences are large enough to indicate substantial variation in the traits being measured, we call them aptitude-achievement "discrepancies". But when is a difference a discrepancy? How large must the gap be before we call a difference a discrepancy? Does this begin to sound familiar? We hope so, but if it does not, the next question should help. How large a difference do we need between an aptitude score and an achievement score before we can conclude that the difference is due to a “real” discrepancy, rather than a “chance” difference? We learned how to use the standard error of measurement ($s_m$) and band interpretation to discriminate real from chance differences among subtests in an achievement test battery. The same principle can be applied to discriminate real discrepancies from chance differences when dealing with aptitude and achievement test scores.

**Check Your Progress 4**

**Notes:** a) Write your answer in the space given below.

b) Compare your answer with that given at the end of the unit.

1. What is the use of academic aptitude test?

..................................................................................................................................................

..................................................................................................................................................

**13.7 LET US SUM UP**

Standardized tests are carefully constructed by curriculum experts, teachers and school administrators which carry standardized administration and scoring procedures. These tests are frequently used to make comparisons over time or across schools. Although the standardized tests are not as useful as the teacher-made tests for classroom teacher but accountability requirement have made it necessary for teachers to administer and interpret them. A comparison of standardized achievement test and teacher-made test is made on same points as purpose, quality of test items, reliability and administration and scoring of tests. When administering standardized tests all test administrators should follow instructions in order to minimize error in test administration.

Grade and age norms/equivalents are much less commonly used and suffer from limitations. Percentile ranks are superior to these two norms and are also suitable for interpreting test results to parents. Standard scores also compare a student’s performance with that of his or her peers. Standard scores are
superior to percentile ranks for test interpretations but they tend to be not well understood by many educators and general public.

Standardized achievement test batteries are popular for school use. In these the advantage of unity in plan and standardization be weighed against a single achievement test. Some of the test batteries used are briefly described and enlisted in the unit. Standardized achievement tests tend to be carefully constructed and measure outcomes similar to those measured by academic aptitude tests. When an academic aptitude-achievement test discrepancy is found, the teacher's task is determine why the discrepancy exists, and then take appropriate steps to remedy it.

### 13.8 UNIT-END EXERCISES

1. How useful do you find standardized achievement tests for students?

2. Use a standardized achievement test and comment upon its utility for your students.

### 13.9 ANSWERS TO CHECK YOUR PROGRESS

#### Answers to Check Your Progress 1

1. Standardized achievement tests are based on common objectives and general content common to large number of schools over the country as compared to teacher-made test which is specific to his/her class. Standardized test is developed by specialists and is highly reliable as compared to teacher-made which is less reliable and developed by teacher himself or herself.

2. Standardized tests are used to:
   
   i) evaluate student’s progress during the academic year or over a period of years,
   
   ii) group students for instructional purposes,
   
   iii) evaluate general educational development of students in the basic skills.

#### Answer to Check Your Progress 2

1. Percentile ranks are recommended for interpreting standardized test results to public.

#### Answer to Check Your Progress 3

1. Achievement test battery (1) to evaluate a school's educational program and its component, (2) to help the teacher plan the work of his/her class and grouping of students within it, (3) the likelihood of clerical errors is minimized when single, comprehensive score reports from a battery are used to make comparisons as opposed to several single subject score reports.
1. Tests of academic aptitude predict academic achievement. But how well they do predict depends on what we use as our outcome measure. Carefully constructed standardized achievements tests measure outcomes similar to those measured by academic aptitude tests.

13.9 SUGGESTED READINGS


