

Assessment Brief



Public Schools of North Carolina

State Board of Education • Phillip J. Kirk, Jr., Chairman • North Carolina Department of Public Instruction • Michael E. Ward, Superintendent

Fall 1999
Vol. 6, No. 5

Understanding End-of-Course Testing: Scores and Standards

This publication is an update to the Fall 1995 Assessment Brief "Understanding End-of-Course Testing: Scores and Standards" (Vol. 2, No. 1). This update is also available at www.dpi.state.nc.us/accountability/testing.

Background and Introduction

The ABCs of Public Education is a comprehensive plan to reorganize public schools in North Carolina. The ABCs focus on strong accountability; a strong emphasis on the basics and on high educational standards; and maximum local control. The North Carolina End-of-Course (EOC) Tests were initiated in response to legislation passed by the North Carolina General Assembly--the North Carolina Elementary and Secondary Reform Act of 1984. This act called for the implementation of the Basic Education Program through the establishment of a core curriculum for all students for each content area and the development of tests to assess the implementation of each curriculum across the state. Based on this legislation, the North Carolina EOC Tests were developed for two purposes:

- To provide accurate measurement of individual student skills and knowledge specified in the North Carolina *Standard Course of Study*, and

- To provide accurate measurement of the knowledge and skills attained by groups of students for school, school system, and state accountability.

The scores for the multiple-choice North Carolina End-of-Course Tests are reported in several different formats:

- scale scores which measure subject-specific achievement and are standardized across tests (revised tests);
- percentiles which allow comparisons of achievement relative to the performance of others in the state; and
- achievement levels which allow for the comparison of student and group performance to predetermined standards based on what is expected in each course.

Table 1 below contains the descriptive statistics for the most recent norming year for each of the ten multiple-choice EOC tests. The most recent year that norms were established for each test is located in parentheses in the first column.

Scale Scores

Often it is desired to compare the results of two tests. For example, Jacob received a score of 65 on his Algebra I test and a score of 72 on his English I test. On which test did he do better when compared to other students taking

Table 1. Descriptive Statistics for the Current Norming Year for End-of-Course Tests.

EOC Test (Current Norming Year)	Mean	Standard Deviation	10th Percentile	50th Percentile	90th Percentile
Algebra I (1994)	55.1	9.1	42.8	55.3	66.7
Algebra II (1997)	58.8	10.5	44.8	59.0	72.8
Biology (1995)	55.5	8.7	43.8	55.9	66.7
Chemistry (1997)	56.8	8.5	45.6	56.7	68.0
ELP (1997)	53.8	9.3	40.7	54.6	65.3
English I (1995)	53.1	8.9	41.3	53.2	64.5
Geometry (1997)	57.0	10.4	42.6	57.2	70.7
Physical Science (1997)	53.7	9.4	40.6	54.1	65.8
Physics (1997)	55.9	9.0	44.0	56.2	67.5
US History (1995)	56.2	8.3	44.9	56.4	66.9

the tests? One might say that he did better on the English test if these scores are the percent of items he answered correctly on the tests. But if these are the raw scores (i.e., the number of questions answered correctly), then this conclusion may be true or it may not be true. His score of 72 on the English I test may mean that he answered 72 out of 100 questions correctly, or 72 out of 72 correctly or 72 out of 216 correctly. With only raw scores, it cannot be determined whether these scores are low, high, or intermediate. More information is needed.

To compare Jacob's two test scores, his scores need to be compared in relation to the distribution of all scores on each of the tests. Jacob's test scores must first be converted to a common scale using the means and standard deviations of each of the test score distributions. This common scale, z-score, has a mean of 0 and a standard deviation of 1. All other standard scores are derived from this common scale. For example, raw scores on the SAT are converted to the common scale and then converted to standard scale scores that have a mean of 500 and a standard deviation of 100.

Beginning in July 1996, the ten multiple-choice EOC tests are reported in terms of scale scores. For easier interpretation, the scale scores have a mean of 50 and a standard deviation of 10. The range of the scale scores is generally from 20 to 90.

Once Jacob's Algebra I and English I test scores have been converted to scale scores (Algebra I scale score would be 69 and English I scale score would be 83), it can be concluded that Jacob did better on the English I test than on the Algebra I test compared to the other students in North Carolina taking the tests.

There are many reasons to report the results from standardized tests as scale scores. Scale scores overcome the disadvantage of many other types of scores, such as percentiles, in that equal differences between score points represent equal differences in achievement. For example:

- *Scale scores can be used to compare test results when there have been changes in the curriculum and/or changes in the method of testing.* If scale scores had been used with earlier versions of the Algebra I tests (based on the 1985 *Standard Course of Study*), then scores on the 1993 forms could have been compared to scores on the 1994 revised forms (based on the 1992 *Standard Course of Study*). Information concerning trends in achievement could have been continued as indicators of long-term student achievement.

- *Scale scores on pretests or released test forms (created by the NCDPI Division of Accountability Services/Testing Section) can be related to secure test forms administered at the end of the course.* In addition, individual released items can be placed on the scale to help aid the interpretation of secure tests.
- *Scale scores can be used to compare the results of tests that measure the same content area but are composed of different formats of items.* If only raw scores (number correct) were reported, there would be no way to compare a student's score of 55 out of 81 on the multiple-choice part of a test to a score of 10 out of 20 on four open-ended questions. Transforming the scores to scale scores takes into account the difficulty of each test and the range of possible scores.
- *Scale scores can be used to minimize differences between various forms of a test.* In 1994 three forms of the Algebra I test were administered in each classroom and in 1995 three additional forms were administered. Even after equating there are slight differences between the raw scores of individuals with equal ability who took different forms of the tests (less the standard error of measurement). Even these slight differences can be minimized by converting the raw scores to scale scores.

Achievement Levels

Achievement levels allow comparisons of student and group performance to predetermined standards based on the expected level of performance in each course. Achievement levels were determined by relating the judgments of North Carolina teachers about the performance of each of their students to each of their student's performance on the EOC tests. Four levels are reported for each course. The contrasting groups method of performance standard setting was used to establish the achievement levels for EOC tests.

Table 2 (on the following page) reflects the range of scores at each of the four achievement levels for the ten multiple-choice end-of-course tests. A description of each of the four levels is also located on the following page.

Across all multiple-choice EOC tests, about 52% of the students score at Levels III and IV.

Percentiles

A percentile shows the relative position of an individual's test score within the test scores of the standardization group. Percentiles indicate the proportion of the population tested that received a lower score for the norming (baseline) year of administration. For example, in 1995 a student performing at the 62nd percentile in Algebra I scored better than 62% of the students taking the Algebra I test in 1994 (the first year the revised test was administered and norms established).

Table 2. Range of Scores Associated with Each Achievement Level for EOC Tests.

Subject Area	Level I	Level II	Level III	Level IV
Algebra I	23-44	45-54	55-65	66-87
Algebra II	23-45	46-57	58-68	69-88
Biology	26-46	47-54	55-64	65-87
Chemistry	23-47	48-55	56-64	65-88
ELP*	21-42	43-51	52-60	61-87
ELP**	0-29	30-42	43-53	54-67
English I	22-42	43-51	52-60	61-85
Geometry	23-45	46-56	57-66	67-87
Physical Science	23-43	44-53	54-63	64-87
Physics	23-42	43-51	52-62	63-87
US History	27-47	48-56	57-64	65-88

* Achievement level ranges for ELPS for the 1996-97 school year and beyond are based on scale scores.

** Achievement level ranges for ELPS for 1995-96 and earlier are based on raw scores.

Percent Proficient

Table 3 contains the original norming year for each of the ten multiple-choice EOC tests.

EOC Achievement Levels

Level I	Students performing at this level do not have sufficient mastery of knowledge and skills of the course to be successful at a more advanced level in the content area.
Level II	Students performing at this level demonstrate inconsistent mastery of knowledge and skills of the course and are minimally prepared to be successful at a more advanced level in the content area.
Level III	Students performing at this level consistently demonstrate mastery of the course subject matter and skills and are well prepared for a more advanced level in the content area.
Level IV	Students performing at this level consistently perform in a superior manner clearly beyond that required to be proficient in the course subject matter and skills and are very well prepared for a more advanced level in the content area.

Table 3. Original Norming Year of EOC Tests.

EOC Test	Original Norming Year
Algebra I	1987
Algebra II	1988
Biology	1987
Chemistry	1989
ELP	1991
English I	1990
Geometry	1989
Physical Science	1991
Physics	1990
US History	1988

Table 4 (on the following page) contains the trend data for students performing at Level III or above on the multiple-choice EOC tests. Only data collected and reported by the Division of Accountability Services/Testing Section are included. Complete statewide trend data are not available for the years when certain EOC tests were designated local option tests. These are EOC tests for Algebra I, Chemistry, Geometry, Physical Science, and Physics. Beginning in 1998-99, all ten multiple-choice EOC tests were mandated to be administered to all students enrolled in the course. The 1998-99 statewide preliminary data is included in Table 4. (Final 1998-99 results will be available in the Fall of 1999.)

The data reported in Table 4 show that across the years, more and more North Carolina students are performing at Level III and above in each of the ten EOC subjects.

Table 5 (on the following page) shows the increase in the number of students performing at Level III or above when the 1993-94 data are compared to the preliminary 1998-99 data.

Taking into account the increase in the total number of students enrolled in these courses, a total of 152,253 more students scored at Level III or above in 1998-99. When interpreting the data by course, one needs to consider the number of students enrolled in each course.

In conclusion, the data for the ten multiple-choice EOC tests reflect the progress made in key high school subjects as measured by the ten EOC tests.

EOC Test of English II

Information regarding the EOC Test of English II is located in a separate *Assessment Brief* because this test is an on-demand writing assessment, i.e., not a multiple-choice test.

Table 4. Percent Proficient Trend Data for EOC Tests Based on Achievement Level III and Above.

EOC Test	Proficiency Level Score	Percent 1993-94	Percent 1994-95	Percent 1995-96	Percent 1996-97	Percent 1997-98	Percent 1998-99*
Algebra I	55	45.3	54.0	53.1	55.5	61.6	65.4
Algebra II	58	43.3	**	**	**	**	59.0
Biology	55	37.1	56.1	55.8	57.0	59.0	57.7
Chemistry	56	36.9	**	**	**	**	60.4
ELP	52	40.1	61.7	57.0	62.6	66.9	67.4
English I	52	40.7	58.2	58.1	58.5	60.7	64.6
Geometry	57	41.2	**	**	**	**	58.3
Physical Science	54	30.4	**	**	**	**	55.6
Physics	52	57.8	**	**	**	**	72.1
US History	57	35.6	49.7	49.4	49.5	49.6	51.0

*The results for 1998-99 are based on preliminary data reported in the publication *The 1998-99 North Carolina Preliminary State Testing Results* (August 6, 1999). Final results will be available Fall 1999.

**These years are those in which a specific multiple-choice EOC test was not mandated for statewide test administration. At that time, these tests were designated as local option tests. Beginning in 1998-99, the ten multiple-choice EOC tests must be administered to students enrolled in the designated courses.

Table 5. Increase in Number of Students at Achievement Level III and Above. From 1993-94 to 1998-99

EOC Test	1993-94 Number Tested	1993-94 Number at Level III or Above	1993-94 Percent at Level III or Above	1998-99 Number Tested*	1998-99 Number at Level III or Above*	1993-94 Percent at Level III or Above	Increase in Number at Achievement Level III or Above
Algebra I	69,162	31,354	45.3	87,449	57,158	65.4	25,804
Algebra II	42,497	18,408	43.3	48,957	28,885	59.0	10,477
Biology	74,840	27,742	37.1	76,950	44,398	57.7	16,656
Chemistry	38,462	14,178	36.9	41,262	24,922	60.4	10,744
ELP	81,290	32,616	40.1	77,740	52,378	67.4	19,762
English I	81,685	33,237	40.7	89,775	57,954	64.6	24,717
Geometry	53,932	22,245	41.2	60,413	35,221	58.3	12,976
Physical Science	65,777	19,990	30.4	66,838	37,162	55.6	17,172
Physics	10,803	6,241	57.8	11,223	8,092	72.1	1,851
US History	65,872	23,462	35.6	69,701	35,556	51.0	12,094
Total	584,320	229,473		630,308	381,726		152,253

*The results for 1998-99 are based on preliminary data reported in the publication *The 1998-99 North Carolina Preliminary State Testing Results* (August 6, 1999). Final results will be available Fall 1999.