DOCUMENT RESUME

ED 304 474 TM 012 872

TITLE	Technical Characteristics of the North Carolina Test of Algebra II, Forms B-E.
INSTITUTION	North Carolina State Dept. of Public Instruction, Raleigh. Div. of Research.
PUB DATE	88
NOTE	61p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Achievement Tests; *Algebra; Curriculum Evaluation; Mathematics Achievement; *Mathematics Tests; Secondary Education; Secondary School Mathematics; Secondary School Students; *State Programs; Test Format; Testing Programs; Test Norms; *Test Validity
IDENTIFIERS	*North Carolina Test of Algebra II

ABSTRACT

The North Carolina Test of Algebra II was developed for use as an achievement test following the completion of the Algebra II course of study. Its design serves two purposes: (1) a normative measure of student achievement; and (2) an objective-based measurement of curriculum coverage. The test's curricular validity, content validity, instructional validity, item pool content validity, standardization sample, and concurrent validity are discussed. The measurement of student achievement is attained by administering a set of items based on a core of objectives and a set of variable objectives (one item per objective tested). This test design was necessitated by a large number of objectives to be covered in a limited administration time. The measurement of curriculum is met by the same set of items that cover the entire range of objectives taught in the Algebra II course of instruction. Four unique forms of the test (Forms B through E) have been developed to afford a broad curriculum coverage. Normative student scores are based on the 56-item total score. Curriculum assessment is achieved by combining the results from all forms of the test administered at one time--a total of 224 items. This booklet is designed to facilitate proper use of test scores by describing the curricular and psychometric characteristics of the test. Methods for deriving scores, reliability, test norms, test content, and curricular assessment are discussed. Twenty-six graphs and four tables are presented. Outlines of goals and objectives of item content and those rejected for use are appended. (TJH)

Forms B-E

Technical Characteristics of the

North Carolina Test of ALGEBRA II

NCTests

North Carolina Department of Public Instruction
Division of Research
Raleigh, NC 27603-1332

Published 1988

FOREWORD

NCDPI Division of Research, in cooperation with NCDPI Instructional Services Area, has developed diagnostic achievement tests of basic skills for public school students in Grades 3, 6, and 8; survey achievement tests of Science and Social Studies for students in Grades 3, 6, and 8; and high school course tests for students taking Algebra I, Algebra II, Biology, Chemistry, Geometry, and U.S. History. Physics and English I achievement tests will be added in 1990, and other tests are being planned.[a]

To facilitate the proper technical use of the test scores obtained from the administrations of the tests, the curricular and psychometric characteristics of the tests will be described in a series of technical manuals. This manual contains a description of the characteristics of the North Carolina Test of Algebra II.

[a]Readers who have an interest in the origins of the test development program are referred to the North Carolina Elementary and Secondary School Reform Act of 1984, the North Carolina Basic Education Program, the North Carolina Standard Course of Study, and the Teacher Handbook.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (cont.)

DESCRIPTION

The North Carolina Test of Algebra II was developed for use as an achievement test following the completion of the Algebra II course of study. Its design serves a dual purpose: that of a normative measurement of student achievement, and of an objective-based measurement of curriculum coverage.

The measurement of student achievement is attained by administering a set of items based on a basic core of objectives and a set of variable objectives (one item per objective tested). This test design was necessitated because of a large number of objectives to be covered in a limited administration time. The measurement of curriculum is met by the same set of items that cover the entire range of objectives taught in the Algebra II course of instruction.

The first test administration design for the North Carolina Test of Algebra II consisted of one form of the test to be administered each year. During the second year of administration of the test an alternative strategy of test administration was utilized. In this strategy, four unique forms of the test are administered in each classroom, one form per student, such that 25% of the students in the classroom will take Form 1, 25% Form 2, and so on (see Table 1). This alternative administration strategy afforded a broader curriculum coverage than the first strategy of administering only one form of the test at a time.

The normative student scores are based on the 56 item total score. Curriculum assessment is achieved by combining the results from all forms of the test administered at one time—a total of 224 items.

Table 1

Organization of the North Carolina Test of Algebra II

50 Core Objectives			
6 Variable Objectives	6 Variable Objectives	6 Variable Objectives	6 Variable Objectives
56 Items Form B	56 Items Form C	56 Items Form D	56 Items Form E

VALIDITY

The purpose in developing an Algebra II achievement test is to obtain test scores from which inferences may be drawn concerning the degree of success a particular student, classroom, school, or school district has had in mastering the Algebra II curriculum. Another purpose is to discern the degree to which the curriculum has been mastered by the students in the aggregate. To the extent this can be done meaningfully, test scores may be said to be valid. Thus one inference drawn from a test score may be valid, while another inference may not be valid.

Theoreticians insist correctly that only inferences concerning test scores can be said to have validity. Generally, readers understand this, and here the convenient shorthand will be employed of speaking about "test validity" rather than "inferences about achievement drawn from scores obtained from tests."

Test validity is a predominant theme in test development, from the time the idea for a test is conceived until the final test scores have been analyzed and interpreted. For convenience, the various components of test validity will be described as if they were unique, independent components rather than interrelated parts. The first component of test validity to be described will be curricular validity.

Curricular validity. If a test is to be used to measure the degree to which a course of study has been mastered, the first step is to define the curriculum. In the case of Algebra II, curriculum definition was done through a cooperative effort, led by the NCDPI Instructional Services Area, involving curriculum specialists, teachers, administrators, university professors, and others. The result was a list of 16 goals encompassing 106 objectives. Supported by expert opinion and a statewide consensus, these goals and objectives were approved by the State Board of Education as the basis for instruction in Algebra II. Curricular validity, the first step in establishing test validity, was established by this process.

Instructional validity. A basic course of study may not include all of the objectives taught under various circumstances in Algebra II. For example, some advanced classes may cover some material that would be beyond the reach of 95% of all Algebra II students. For this reason it is important to know just what is being taught in the majority of Algebra II courses in the state. To determine this, all of the Algebra II teachers in the state (approximately 750 in 1985) were surveyed concerning the topics they taught every year in their Algebra II classes. The analysis of the Algebra II curriculum was based on 644 responses, or 86% of all of the possible responses.

Each Algebra II teacher examined half of the 106 objectives and noted whether they taught the objective every year and, if so, was it basic to Algebra II instruction. The two answers turned out to be equivalent: If the objective was taught every year, it was considered to be basic to Algebra II instruction. Fourteen of the goals comprising 80 of the objectives were considered to be basic to the Algebra II curriculum (evaluated as basic by at least 65.2% of the responding teachers or considered an essential aspect of the Algebra II curriculum). The objectives used in test development are listed in Appendix A and the objectives that were rejected are listed in Appendix B, together with the proportion of teachers that judged each objective as basic.

Instructional validity, the second step in defining test validity, was established by this procedure. It limits inferences drawn from the Algebra II test scores to the basic instructional program comprising the 80 objectives.

In summary, it was concluded that curricular and instructional validity depended jointly on the 80 objectives and 14 goals under which they were collected, and that the Algebra II test should be built on that foundation.

Content validity of the item pool. Content validity—the degree to which test items reflect the basic instructional program—was defined through a number of operations:

First, the item pool for the Algebra II test was created. It consisted of 1,231 items and contained 12 to 24 items per objective. The items were developed by six North Carolina Algebra II teachers from across the state trained in the technical aspects of item-writing. The use of classroom teachers helped to insure that instructional validity was maintained, since the items would be drawn from their classroom experiences.

Second, the item pool was edited for grammar, syntax, psychometric form, and linguistic bias.

Third, the item pool was analyzed by curriculum specialists and classroom teachers to assure that the items were valid representations of the objectives for which they were written. Each item was reviewed by at least four classroom teachers from different educational regions across the state. The criteria for evaluating each item included the following: curriculum match (objective, difficulty level, thinking skills, and vocabulary), format (familiarity, print size/style, and mechanics), art (clarity, accuracy, and labeling), item bias (gender, ethnic, SES/geographic, or other), stem (accuracy, ambiguity, single problem, wordiness, and complete statement), and answer choices (one best, homogeneity, logical order, clues/cues, and answer key).

Fourth, the items were collected into 14 test forms for field testing. Although the forms were not the final forms of the North Carolina Test of Algebra II, they were organized in such a way that the objectives were represented equitably across all forms. Each form contained 97 or 98 items, 10 of which were common across all forms for the purpose of ability equating should that become necessary.

Fifth, test administration instructions were written, distribution procedures were organized, and administrators were trained to conduct the test administration. The experienced test administration organization used to administer statewide tests in North Carolina was employed to accomplish the field testing. The administration of the test forms followed the routine eventually expected to be used when the test of record was given.

Sixth, a sample of 8,296 students was selected to take the 14 field test forms containing a total of 1,231 items. To insure broad representation, schools were selected from the eight North Carolina educational regions. All Algebra II students in a school took one form of the field test and every test form was administered in all educational regions. Consequently, each item was answered by approximately 593 students (the number of students per field test form ranged from 554 to 638). Since each form contained 97 or 98 items—too many to be answered during the prescribed administration time of two hours—each student was instructed to answer the 10 common items and the even or odd variable items for a total of 54 items. The resulting sample size varied from 276 to 303 per item.

Seventh, the field test data were analyzed using both the classical psychometric model and the one-parameter Rasch model (results were generated from the BICAL computer program). Sixteen statistics were assembled for each item. The item psychometric information was then placed on the item record, which became the basic document to which all other records were referenced. The item record contains the item number, goal, objective, historical information, a copy of the item itself, and the item statistics. Each item has a separate item record.

Eighth, the item statistics were submitted to a computer analysis using a program designed to scan a range of statistics and print out an appropriate notation based on the criteria that had been built into the program specifically for Algebra II. An item was classified as "too easy" if the p-value was greater than .93 or it was classified as "too hard" if the p-value was less than .26. An item was said to have "weak prediction" if the point-biserial correlation was less than .16. An item was said to have "poor discrimination" if the discrimination index was less than .45. An item was said to have an "entrapment choice", a "marginal top group", or an "inverted ICC" if the Item Characteristic Curve groupings displayed certain irregularities. The item psychometric notations were then placed on the item record.

Algebra II cannot be represented by a single factor. Therefore, the maximization of item-total correlations was not a primary goal of item development. Once an item was shown to have at least a modest correlation with a corrected total score (point-biserial greater than .15) and was judged to measure an objective, it was included in the item pool (unless it exhibited other psychometric inadequacies). While this may have reduced the potential internal reliability as measured by coefficient alpha, it increased the validity of the test by allowing for an objective factor structure that was not expected to be unitary.

Gender and race bias analyses were conducted on each item in the Algebra II item pool. The following statistics were computed for each item: the partial correlation between success on the item score and group membership (with the total score partialed); the chi square statistic comprising observed chi squares for contingency tables (group x item-success) for three cuts of the total-score distribution; and the delta distance statistic, which is the distance (of the point representing the item) from the main axis of the ellipse in the plot of normalized item difficulties (group 1 versus group 2). No conclusive results were obtained due to the small sample sizes (N's ranged from 276 to 303 students per item) and the large proportion of field-tested items that exhibited poor statistical properties. Therefore, the extent of bias in each item due to gender or race was judged on the qualitative assessments made by teachers during the item review process (step 3 described previously).

These statistical notations were reviewed and decisions were made about the psychometric adequacy of the items. The decisions were then conveyed to curriculum specialists, who also reviewed the items and reached a decision about the curricular adequacy of the items. The psychometric and curricular decisions concerning the item's suitability for use in a test were then placed on the item record.

Of the 1,231 field tested items, 306 (24%) were deleted due to statistical flaws, i.e., too easy or too difficult, weak prediction, ICC irregularities, or lack of validity as evidenced by the item's relationship with the total score; and 20 (2%) were deleted due to curricular deficiencies (objectives 6.8 and 12.3). This left 905 items (74%) in the Algebra II item pool for future test development.

Content validity of the test. After a consideration of the logistics involved, it was decided to prepare one complete 56-item test for administration in May 1987, and to field test four additional core tests of 56 items each for use in succeeding years. The core tests were based on the same core objectives as the complete test and variable objectives were chosen at random. Each test was composed of one item per objective.

This method of item selection is a modified domain sampling model, with the various cores randomly equivalent. The domain sampling model in its pure form is highly inefficient because it allows the entry of items that are grossly inappropriate for normative measurement—items that no one can answer or that everyone can answer, or items that have psychometric deficiencies of a more complex form. In the modification used here, the domain of items was limited to those items that had satisfactory psychometric characteristics and were relevant to the Algebra II curriculum. This was determined by the analysis of the item field test data, which was used to verify the psychometric adequacy of the item pool and to direct where item revisions should be made.

After the test was assembled into test forms, they were reviewed by one curriculum supervisor and two teachers in each of the eight educational regions. The criteria for evaluating each form of the test included the following: (1) the content of the test reflects the goals and objectives taught; (2) the items appear to reflect several levels of mastery; (3) the test layout is pleasing, the format is consistent from item to item, and the symbols and figures are precise; and (4) each item has one and only one answer that is correct. The ratings for the 1987 North Carolina Test of Algebra II were generally average to superior on all of the criteria.

Although the initial equating of the core tests depended upon random selection of the items from the item pool, the final equating was based on the statistics obtained at the time the first test of record was administered. This second psychometric analysis, described next, was used to eliminate random differences among the cores and thus facilitate the precision of measurement from one year to the next.

Standardization sample. The first North Carolina Test of Algebra II consisted of one form (Form A) containing 56 items measuring the 50 core objectives and 6 randomly-selected variable objectives. This test was administered to 36,633 North Carolina Algebra II students in May 1987.

At the same time, four additional core tests (Forms B-E) of 56 items (one item measuring each of the 50 core objectives and six randomly-selected variable objectives) were administered as separate forms to 2,663 students (662 to 672 students per form). These students attended schools randomly-selected to be representative of the state on the basis of criteria that were judged to be related at least partially to Algebra II ability levels—school performance on the North Carolina Competency Test, for example. (Prior to the first Algebra II test, no comparative information existed on Algebra II achievement.)

The four future core tests (Forms B-E) were interleaved in all student samples. This produced an even spread of ability across all four future core tests. The agreement of the mean test scores on all four future core tests (see Table 2) supported the view that the samples were representative of all North Carolina Algebra II students. The disagreement of the mean test scores of the four future core tests with the mean test score from the first test (see Table 2) indicate that some other variable was operating to produce the discrepancy.

At this point a decision was made to adopt a different test administration strategy in order to generate more data for curriculum analysis. The first test of record, administered May 1988, consisted of four cores (Forms B-E) administered simultaneously. These forms were administered to 35,146 North Carolina Algebra II students. The state norm population comprises these 35,146 students.

As in the previous year, four additional core tests (Forms 1-4) of 58 items (one item measuring each of the 50 core objectives and eight randomly-selected variable objectives) were administered as separate forms to a sample of 2,624 students (621 to 686 students per form). These students attended schools selected to be representative of the state on the basis of criteria that were judged to be related to Algebra II ability levels—school performance on the 1987 North Carolina Test of Algebra II, for example.

The four future core tests (Forms 1-4) were interleaved in all student samples. This produced an even spread of ability across all four tests. The agreement of mean test scores on the four future core tests and the agreement of these mean test scores with the state norm means for the first test of record (see Table 2) support the view that the samples were representative of all North Carolina Algebra II students.

Concurrent validity of the test. When the 1988 test of record was administered, Algebra II teachers were asked to indicate the expected final letter grade for each student in their classes. Figures 1-5 display a comparison of letter grades and the mean Algebra II test score corresponding to each letter grade for the overall student population and for each of the four core tests administered. All of the figures are consistent and add to the equivalence of the four core tests. The figures exhibit a positive relationship between performance on the North Carolina Test of Algebra II and the subsequent grade in an Algebra II course.

The concurrent validity, as determined by the relationship of Algebra II test scores to Algebra II course grades, conforms closely to expectations and contributes to the validity of inferences concerning student achievement as measured by the North Carolina Test of Algebra II.

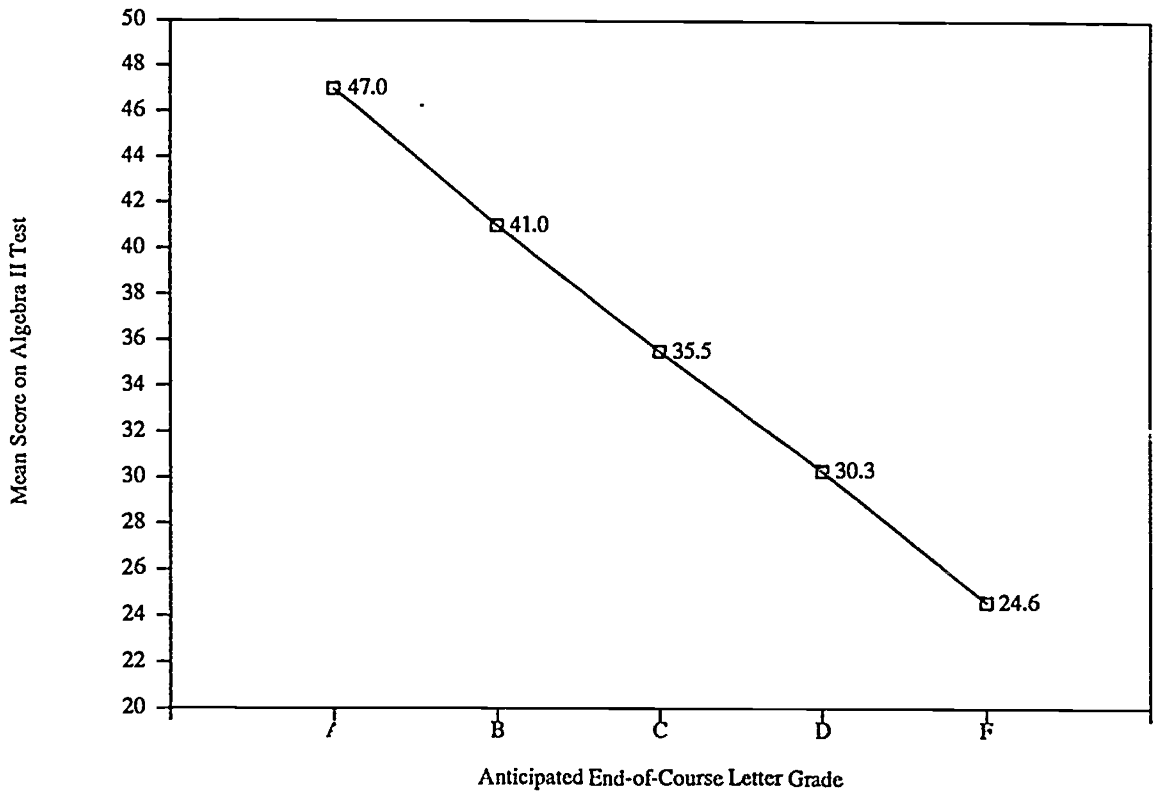

Figure 1. Comparison of letter grades teachers expected students to receive and scores subsequently earned on the 56-item 1988 North Carolina Test of Algebra II (N = 35,146 students).

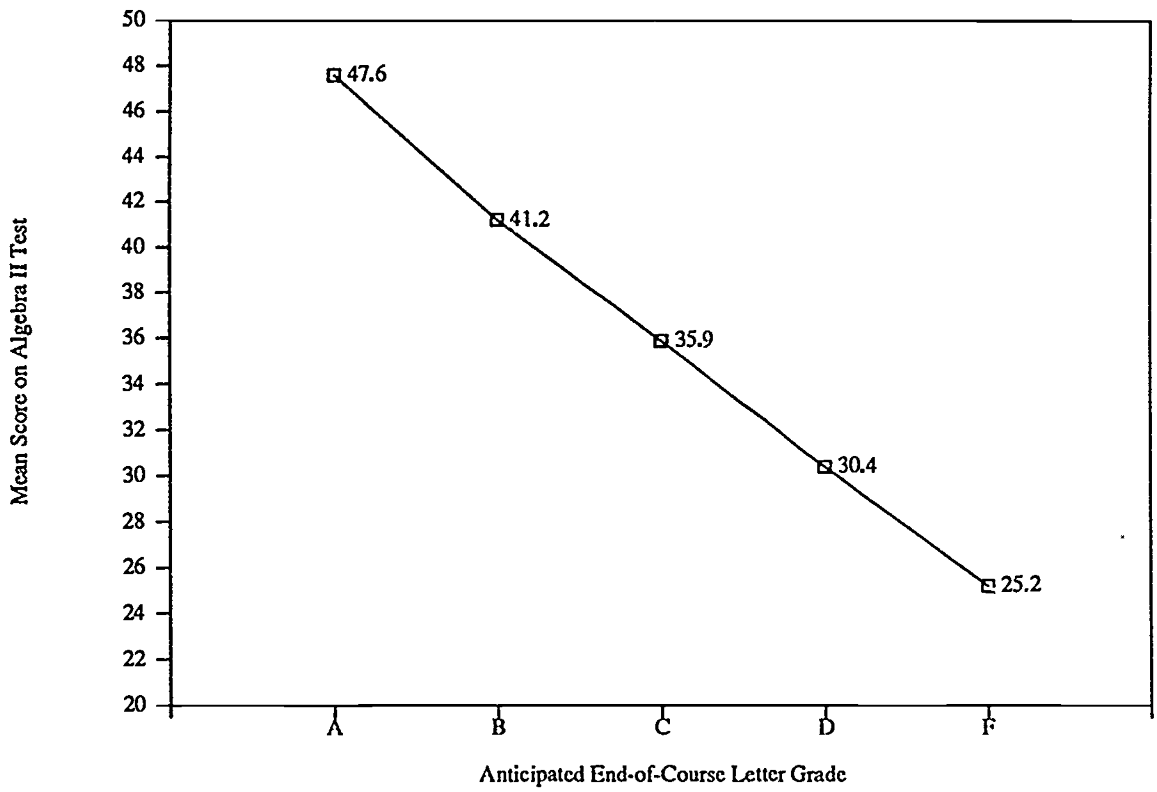

Figure 2. Comparison of letter grades teachers expected students to receive and scores subsequently earned on the 56-item 1988 North Carolina Test of Algebra II Form B (N = 8,917 students).

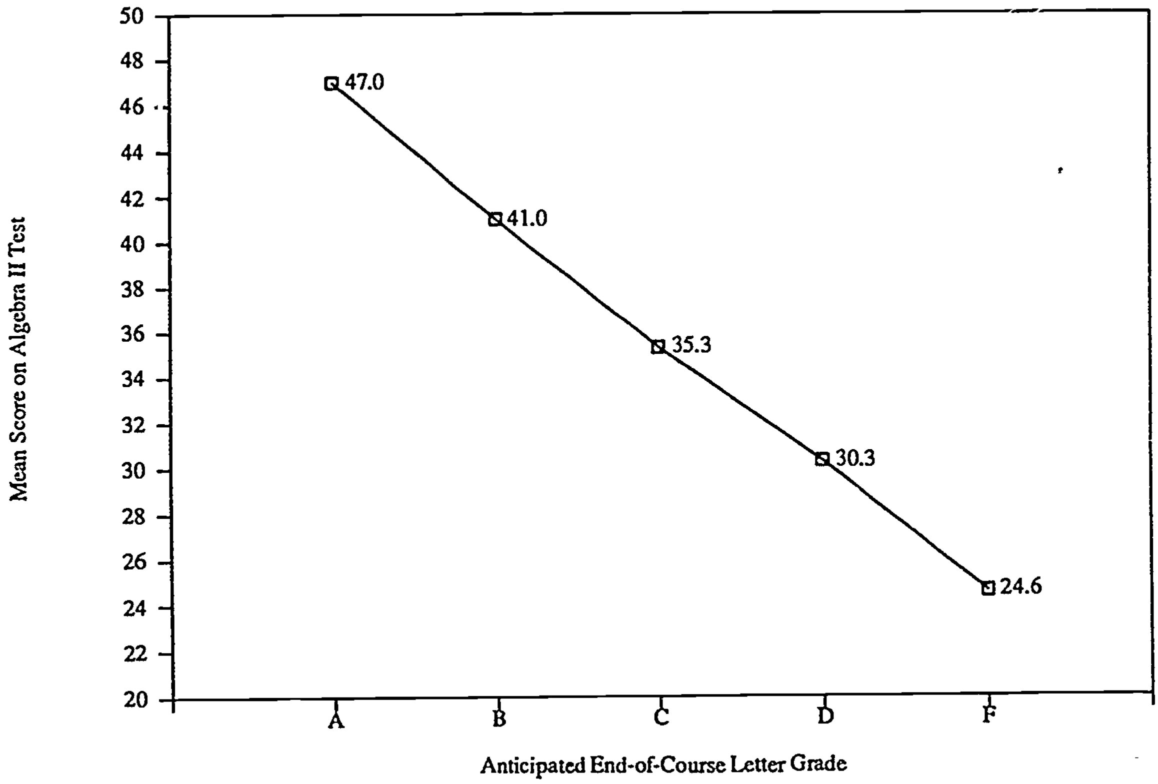

Figure 3. Comparison of letter grades teachers expected students to receive and scores subsequently earned on the 56-item 1988 North Carolina Test of Algebra II Form C (N = 8,829 students).

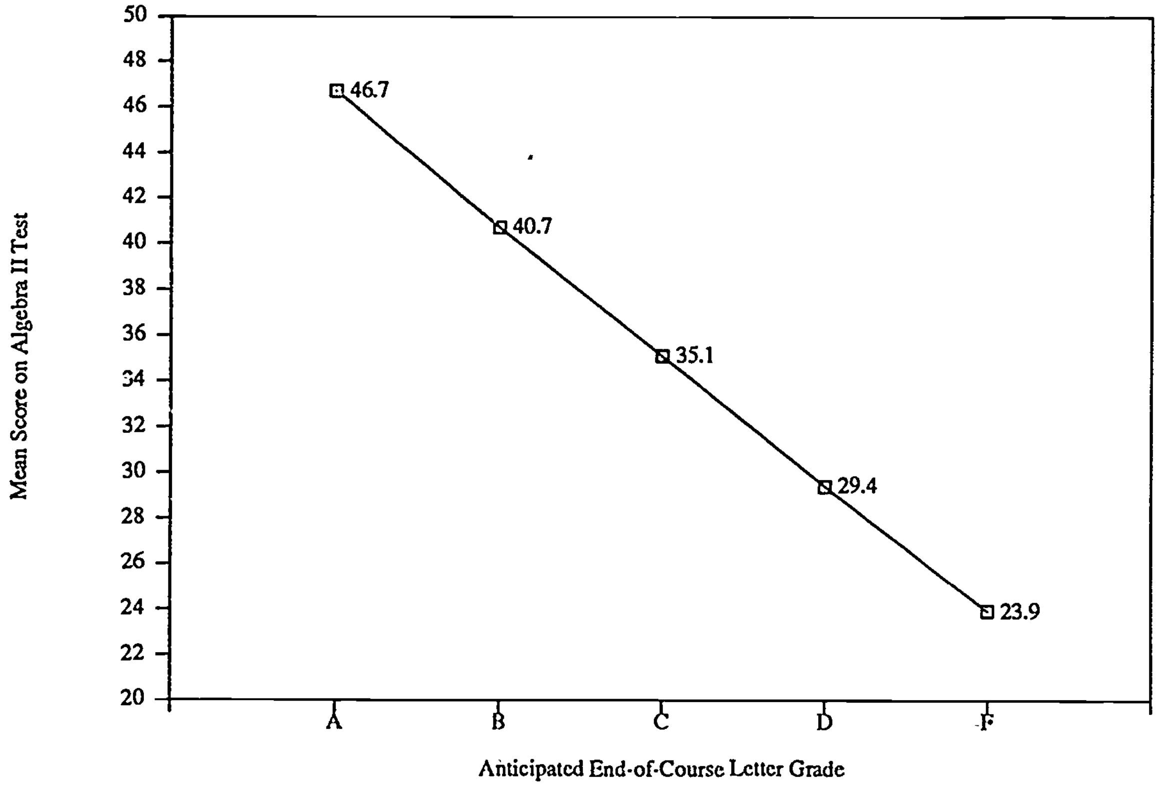

Figure 4. Comparison of letter grades teachers expected students to receive and scores subsequently earned on the 56-item 1988 North Carolina Test of Algebra II Form D (N = 8,750 students).

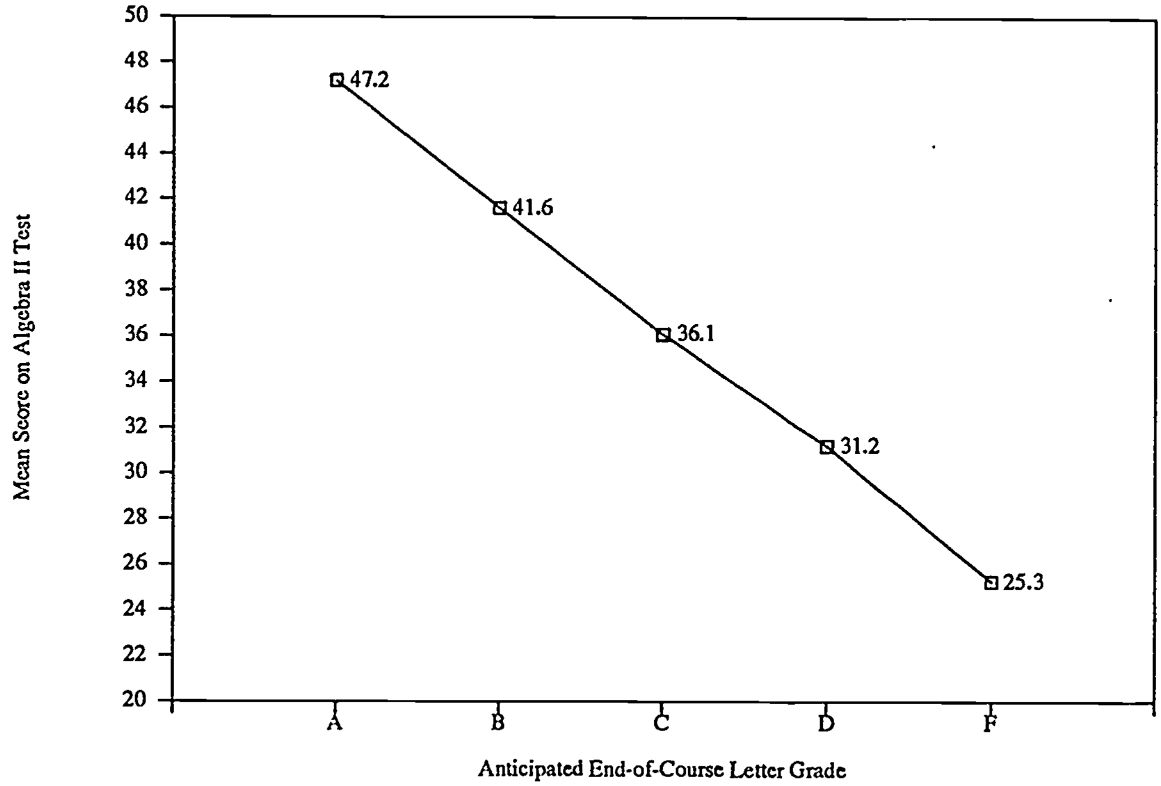

Figure 5. Comparison of letter grades teachers expected students to receive and scores subsequently earned on the 56-item 1988 North Carolina Test of Algebra II Form E (N = 8,650 students).

METHOD FOR DERIVING TEST SCORES

Item information was available to support the classical scoring model and the Rasch scoring model. The classical model gives a unitary weight to each item; a correct choice adds 1 to the total score, an incorrect choice adds 0. The one-parameter Rasch model also uses unitary weighting. (The two- and three-parameter item response models given more credit for answering some items correctly, and less credit for answering other items correctly. These models assume that each item has a fundamental, unchanging difficulty level.)

The classical scoring model was utilized to score the North Carolina Test of Algebra II because it is fundamentally sound, simple to use, and easy to interpret. A total score is computed consisting of the sum of the 56 items on each core test.

RELIABILITY AND OTHER STATISTICS

The descriptive statistics, the standard error or measurement, the alpha reliability coefficients, and the alternate form reliability estimates (correlations between the first statewide core test and the four future core tests) for the first statewide administration of the North Carolina Test of Algebra II in May 1987 (Form A), for the second statewide administration in May 1988 (Forms B-E) and for the four future core tests administered in May 1988 (Forms 1-4) are given in Table 2.

TABLE 2

Descriptive Statistics for the North Carolina Test of Algebra II

Form	Mean	s	Median	se_{meas}	Reliability Coefficient Alpha	Alternate Form
A(ST)	37.63	10.29	38	4.12	0.91	0.84,0.83,0.85,0.83
B-E(ST)	36.07	10.69	36			
B(FT)	34.54	10.05	35	4.06	0.90	0.84
B(Eqt)	34.44	10.10				
B(ST)	36.24	10.69	37			
C(FT)	34.08	10.26	34	4.24	0.91	0.83
C(Eqt)	34.42	10.20				
C(ST)	35.93	10.63	36			
D(FT)	34.50	10.46	34	4.02	0.91	0.85
D(Eqt)	34.46	10.30				
D(ST)	35.50	10.91	36			
E(FT)	34.63	9.74	35	3.97	0.90	0.83
E(Eqt)	34.44	10.30				
E(ST)	36.60	10.50	37			
1(FT)	34.41	8.99	34	4.38	0.87	
1(Eqt)	34.52	10.19				
2(FT)	35.96	9.80	36	4.37	0.90	
2(Eqt)	34.63	10.79				
3(FT)	34.66	9.72	35	4.30	0.90	
3(Eqt)	34.67	10.62				
4(FT)	34.33	9.73	34	4.34	0.90	
4(Eqt)	34.63	10.25				

FT: Field Test administration.
Eqt: Equated to the mean of Forms B-E.
ST: Statewide Test administration.

The total core scores are symmetrical about a mean score of 36, or 64 % correct (see Figures 6 through 10). The alpha reliability estimates have a mean value of .91; and the alternate form reliability estimates have a mean value of .84.

For practical purposes, the proper measure of reliability is the alternate form reliability. The calculation of this statistic requires that two or more equivalent test forms be developed. The older procedure required the development of one form, which was then "cloned" to obtain a second, alternate form of the test. A judicious selection of alternate items was recommended to prevent direct memory transfer from an administration of one test to its alternate form. But the possibility remained that errors of selection in the first form would be duplicated in the second form. A newer procedure requires that the tests be truly equivalent—that is, that the two or more tests be developed in exactly the same way, but independently of one another. This permits the reliability coefficient to reflect any random errors of selection made in the development of either of the test forms.

The alternate forms developed for the North Carolina Test of Algebra II reflect this newer procedure. That is, each test form is developed from the domain of items in exactly the same manner. The alternate form reliability of these tests reflects:

- trait instability not following from maturation or instruction
- instrument instability resulting from fallible test development procedures
- administrative instability reflecting different testing occasions

The square of the alternate-form reliability coefficient accurately reflects the maximum proportion of variance one can legitimately expect to predict from the administration of the North Carolina Test of Algebra II ($r^2 = .84^2 = .71$) when test scores are compared across time or with other measures of student abilities or personality traits that have similar reliabilities. In brief, the alternate form reliability correlation coefficient is the statistic to use when correcting for attenuation.

With the test design employed in the North Carolina Test of Algebra II, a major concern was the equivalence of the four core tests administered each year. From an examination of the descriptive statistics in Table 2, it can be concluded that the means of the four core tests were not statistically different when they were first field tested in 1987 (for Forms C and E, $t = 1.003$, $n_c = n_e = 666$, $t(.05) = 1.96$).

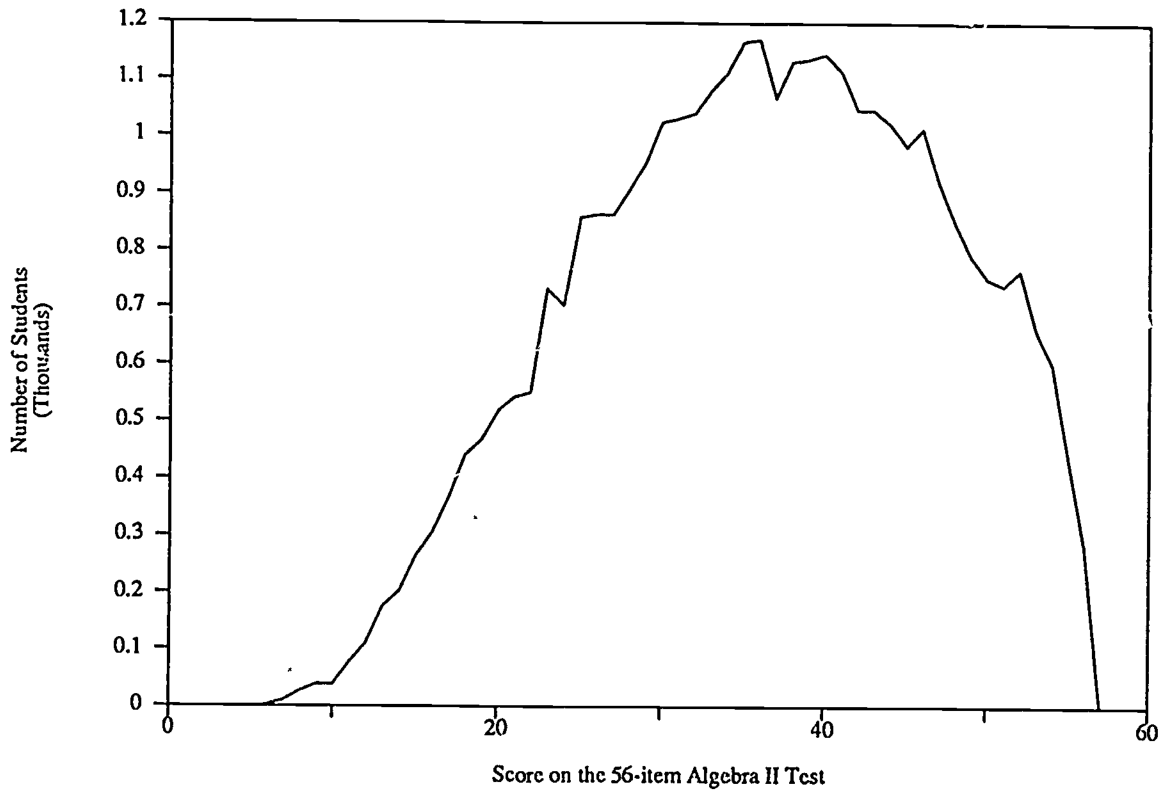

Figure 6. Frequency distribution of scores on all forms of the 1988 North Carolina Test of Algebra II 56-item core test (N = 35,146).

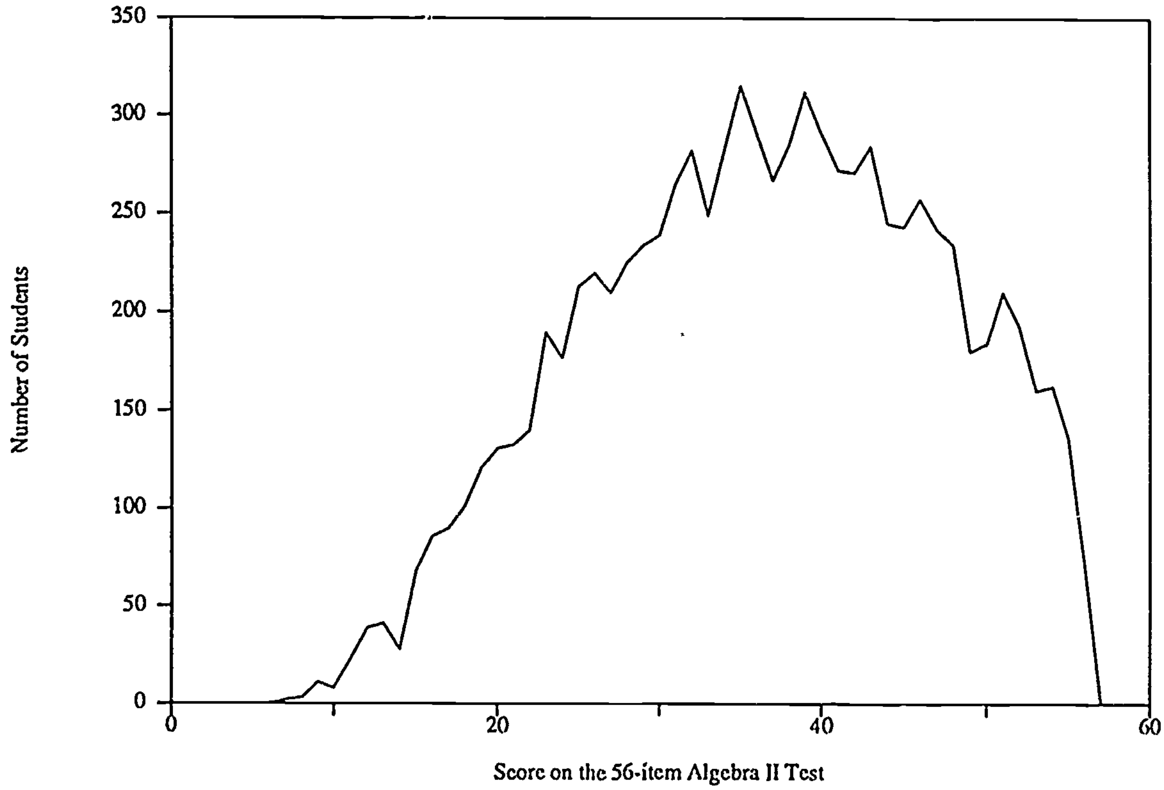

Figure 7. Frequency distribution of scores on Form B of the 1988 North Carolina Test of Algebra II 56-item core test (N = 8,917).

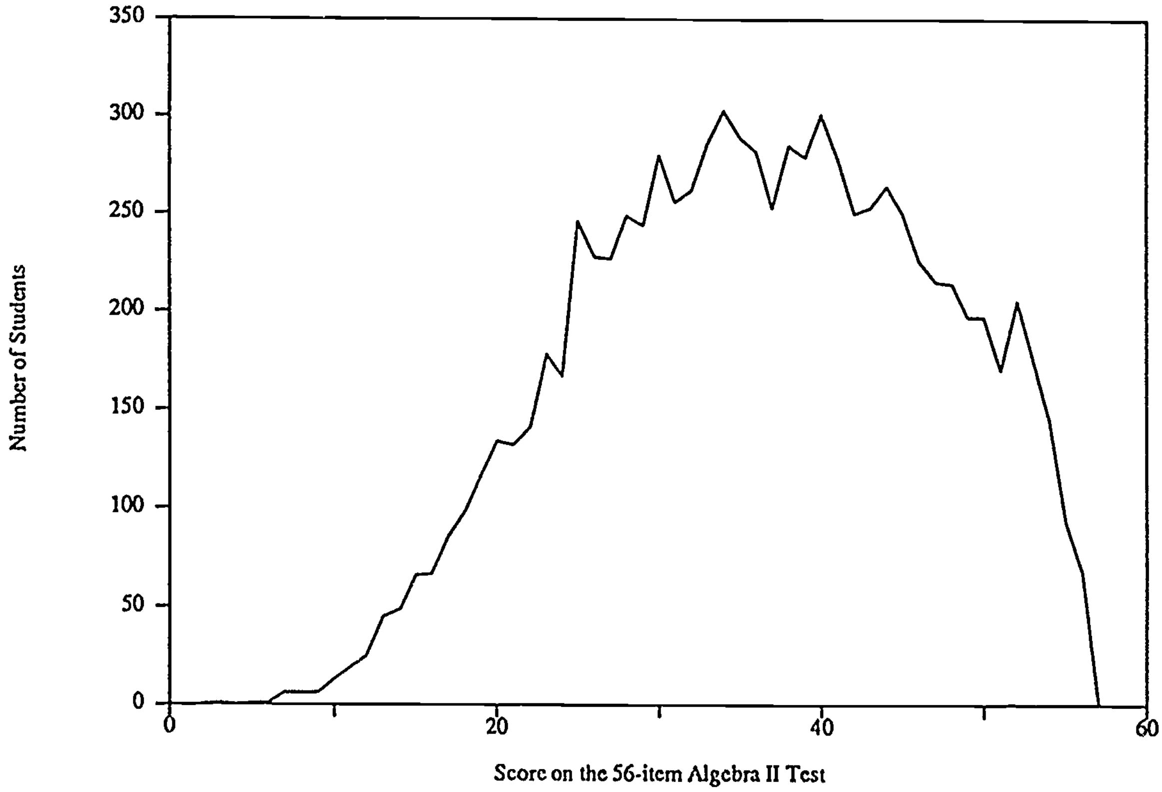

Figure 8. Frequency distribution of scores on Form C of the 1988 North Carolina Test of Algebra II 56-item core test (N = 8,829).

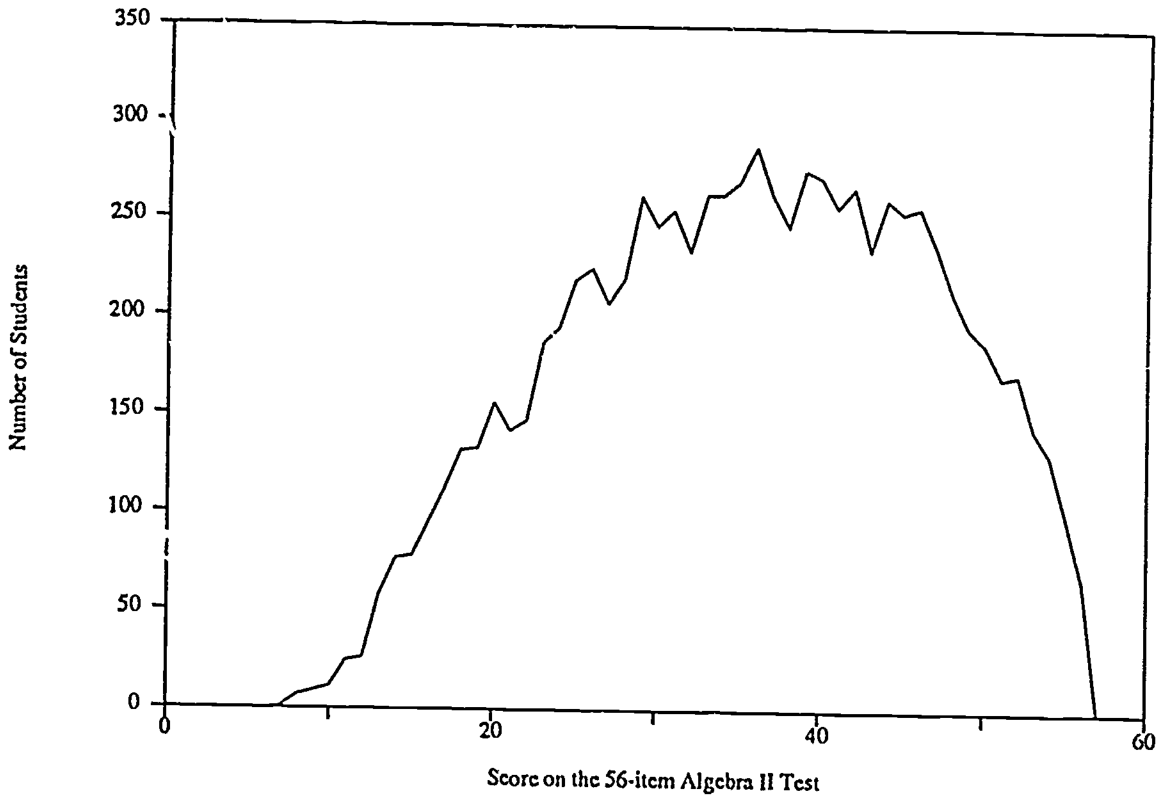

Figure 9: Frequency distribution of scores on Form D of the 1988 North Carolina Test of Algebra-II 56-item core test (N = 8,750).

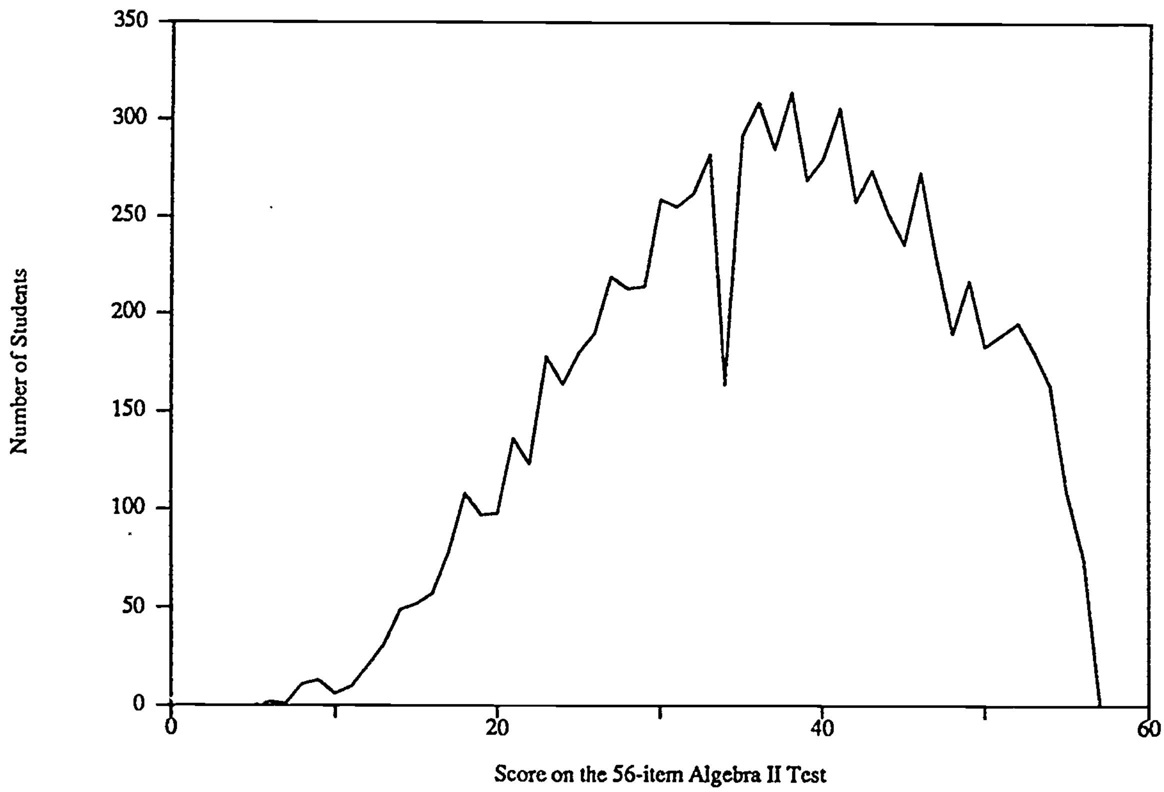

Figure 10. Frequency distribution of scores on Form E of the 1988 North Carolina Test of Algebra II 56-item core test (N = 8,650).

To examine further the equivalence of the four core tests, an equipercentile comparison of the four core tests and perfect agreement (a slope of 1.00 determined from the mean of the four core tests) was conducted. To make this comparison, the mean of a block of scores within successive five percentile points was taken to compare with the mean of the block of scores within the same five percentile points on the second test. This yielded twenty reasonably reliable points of comparison. The results are displayed in Figures 11 through 14.

The differences of the data points of the four core tests from perfect agreement are small (see Figures 11 through 14). These differences could be adjusted statistically by providing a separate set of norms for each form. A simple and efficient alternative is to redevelop the core tests slightly so that even small differences disappear. With this technique, a single norms table can be used for all four core tests and all future core tests. To accomplish this transformation, the test developer had available the 224 items on the four core test forms for which comparable psychometric data were available.

The results of the adjustments for core tests B through E, employed in 1988, are given in Table 2 and Figures 15 through 18. The required changes (the substitution of four items) were minimal.

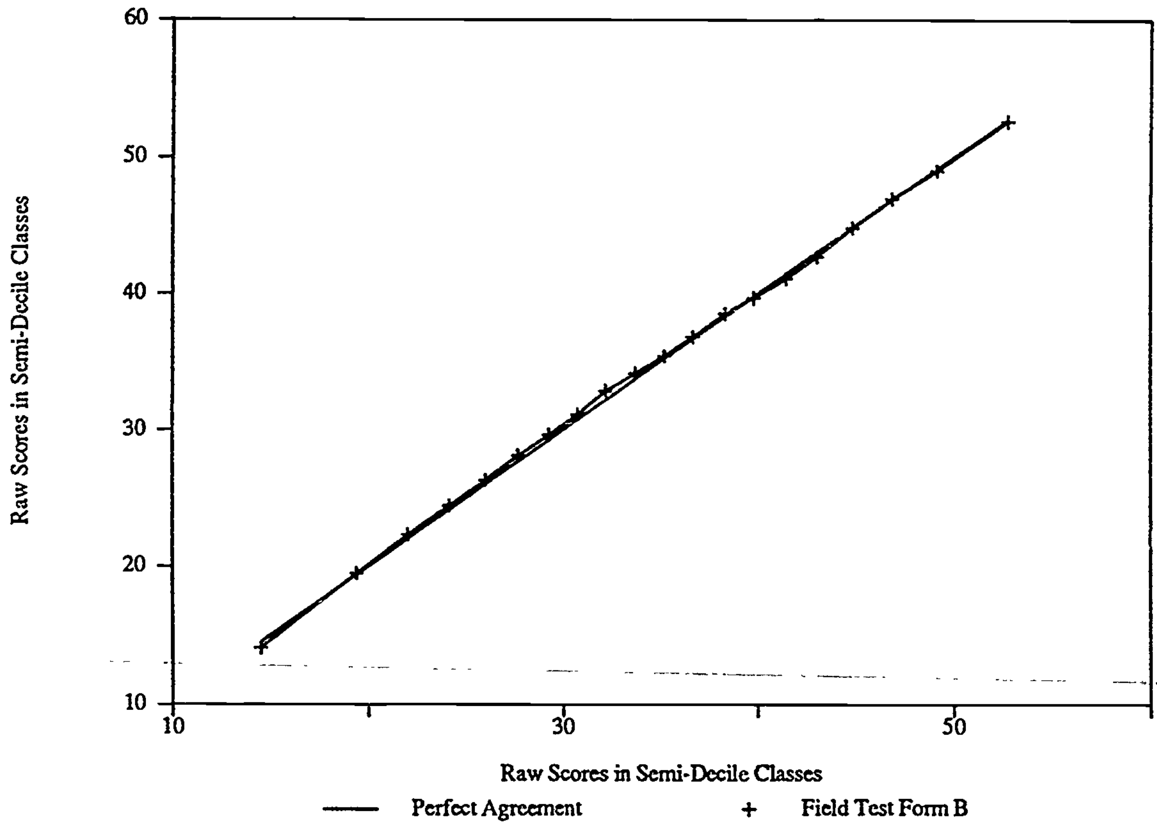

Figure 11. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core B (unadjusted).

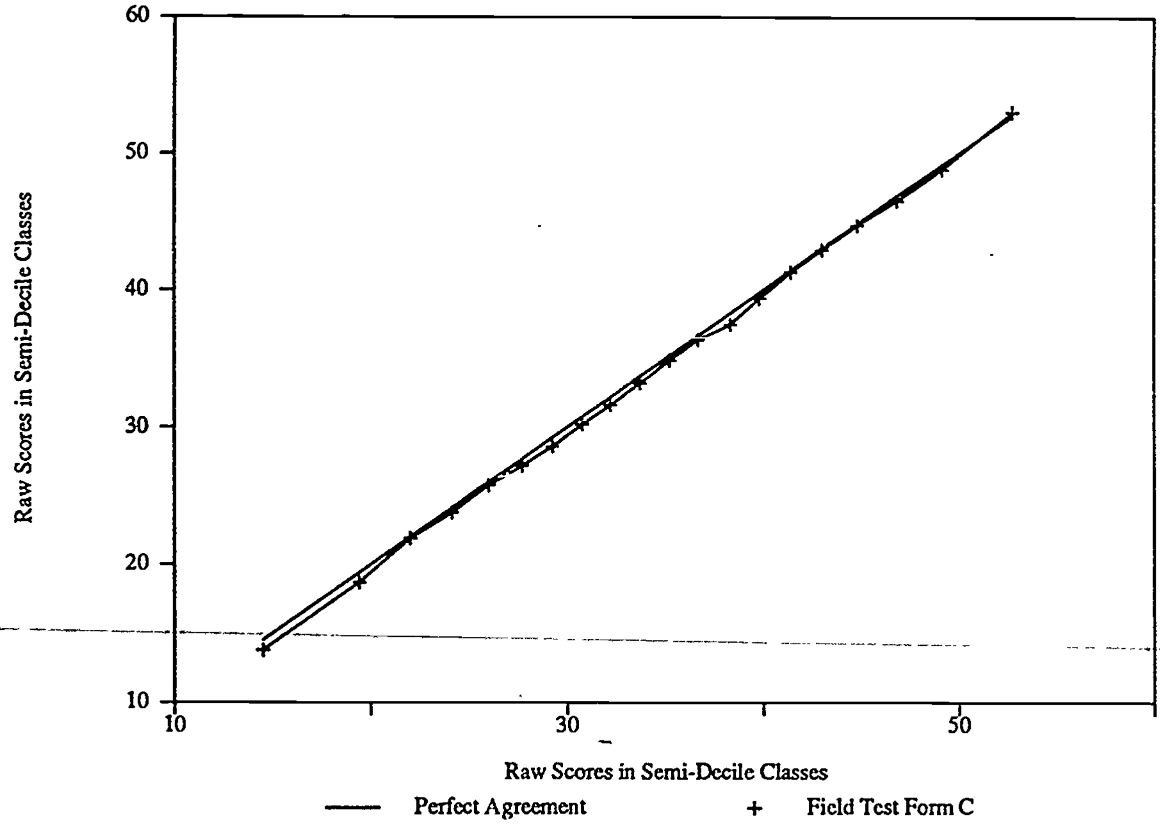

Figure 12. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core C (unadjusted).

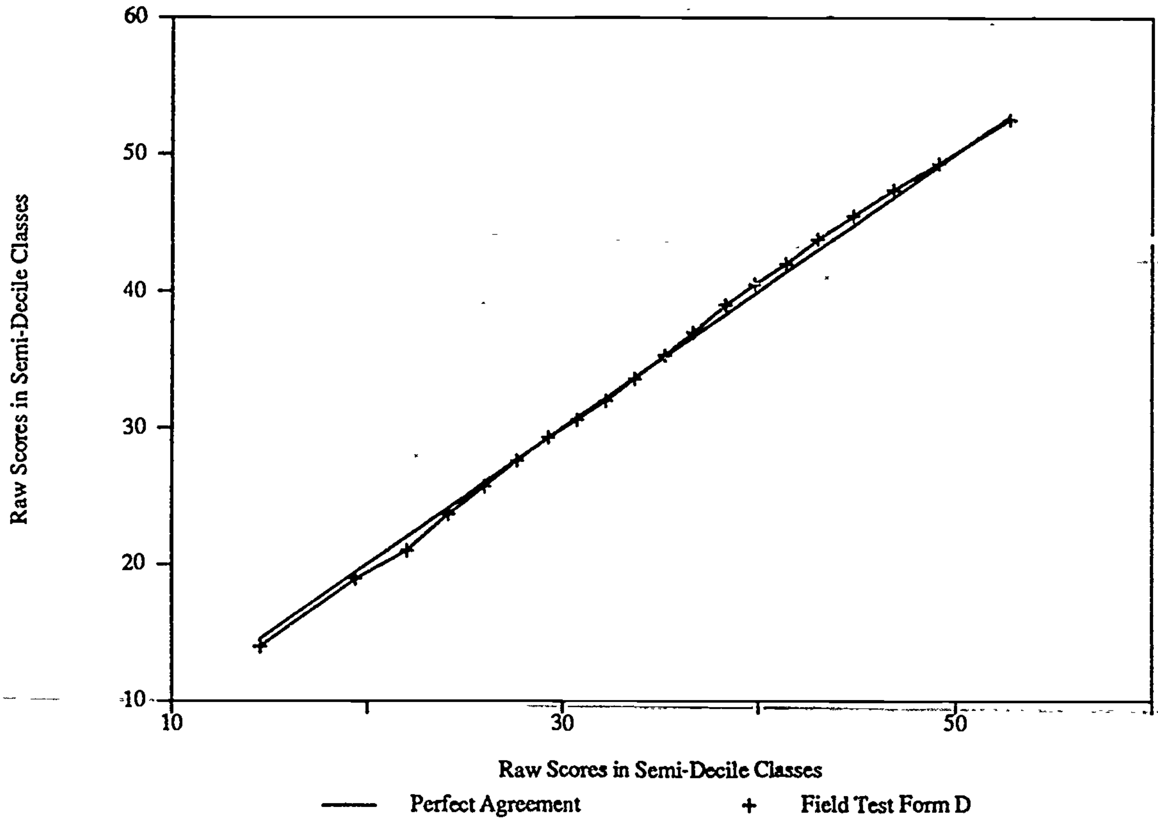

Figure 13. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core D (unadjusted).

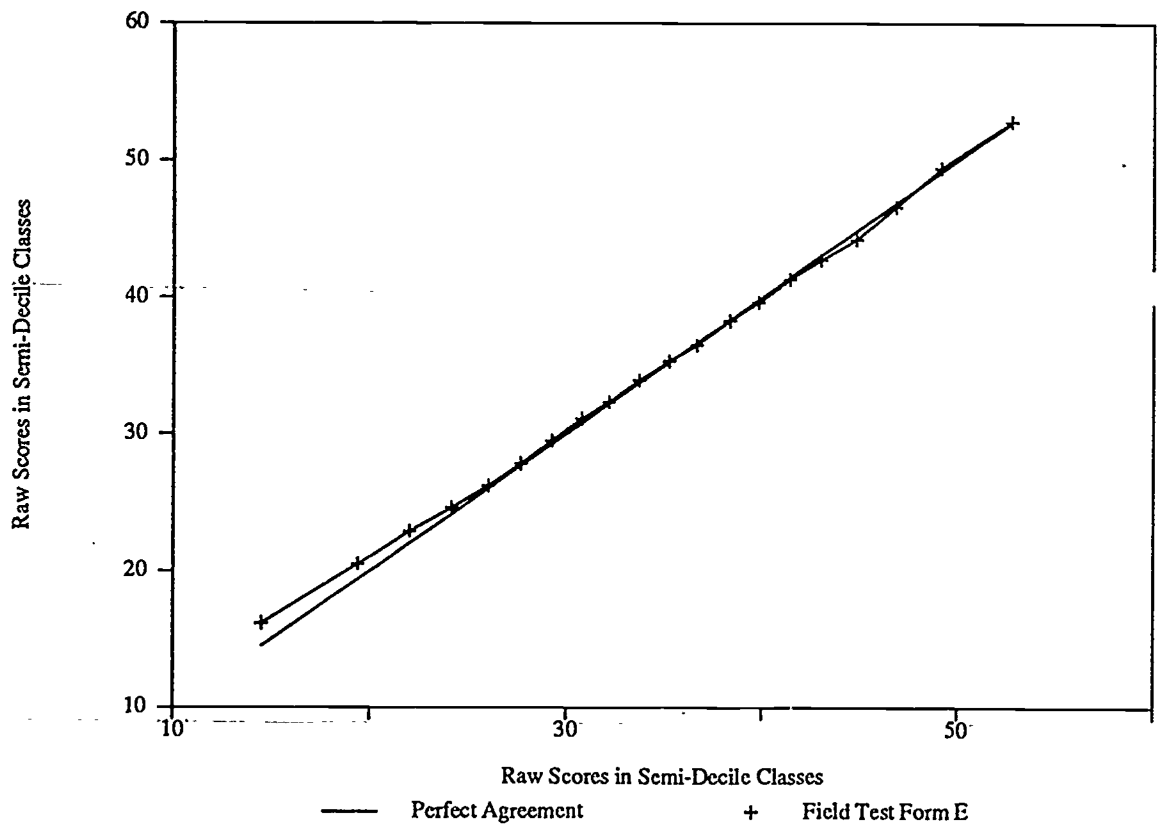

Figure 14. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core E (unadjusted).

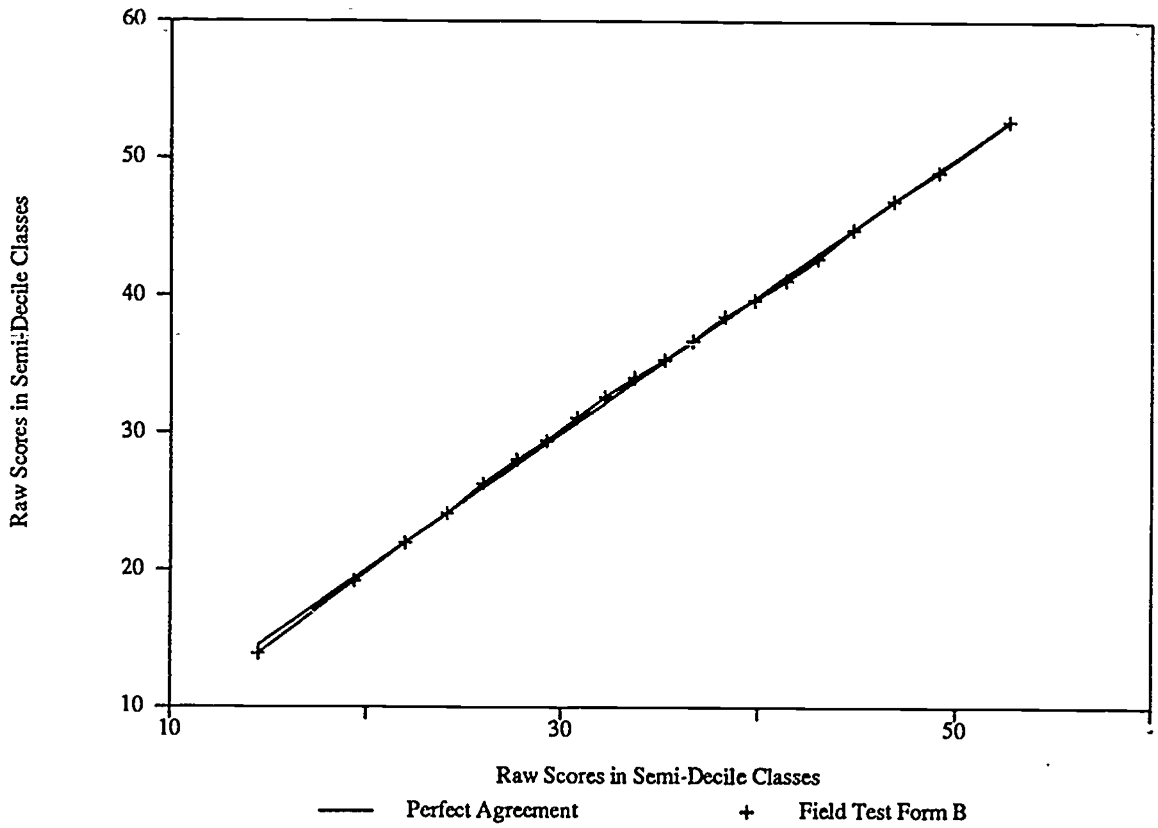

Figure 15. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core B (adjusted).

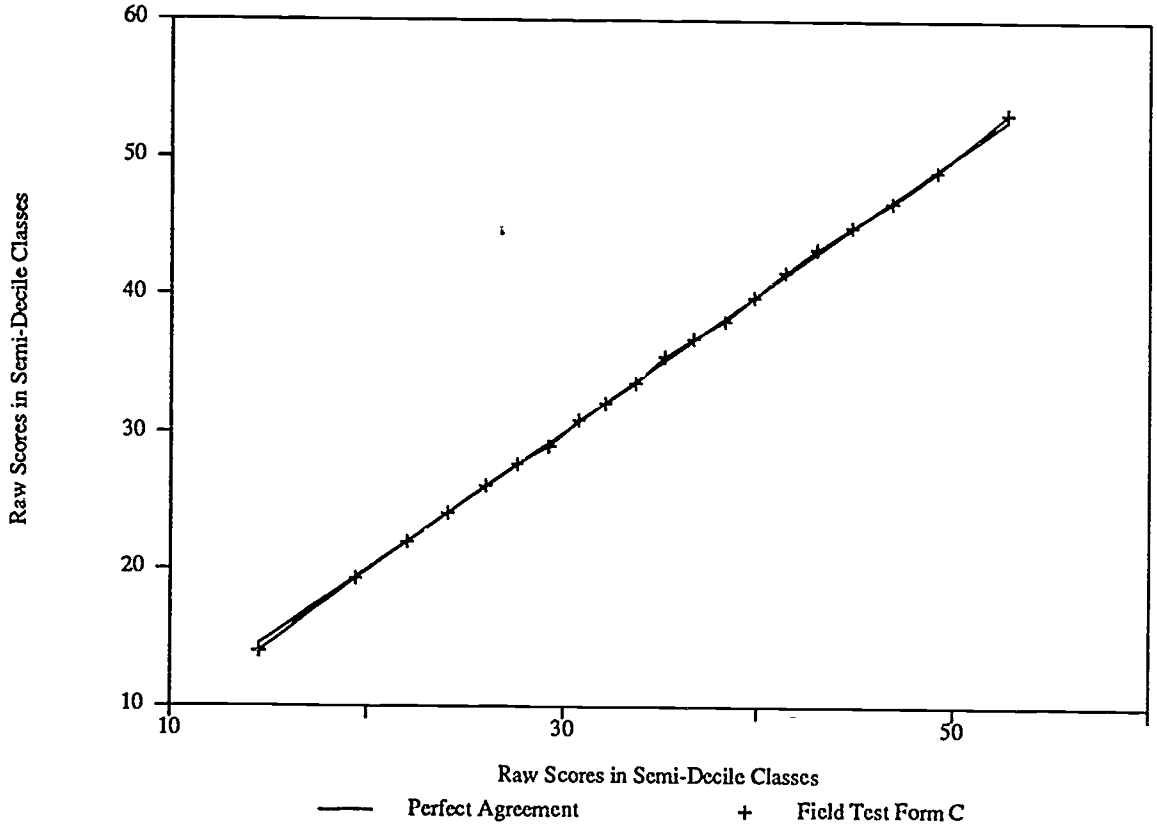

Figure 16. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core C (adjusted).

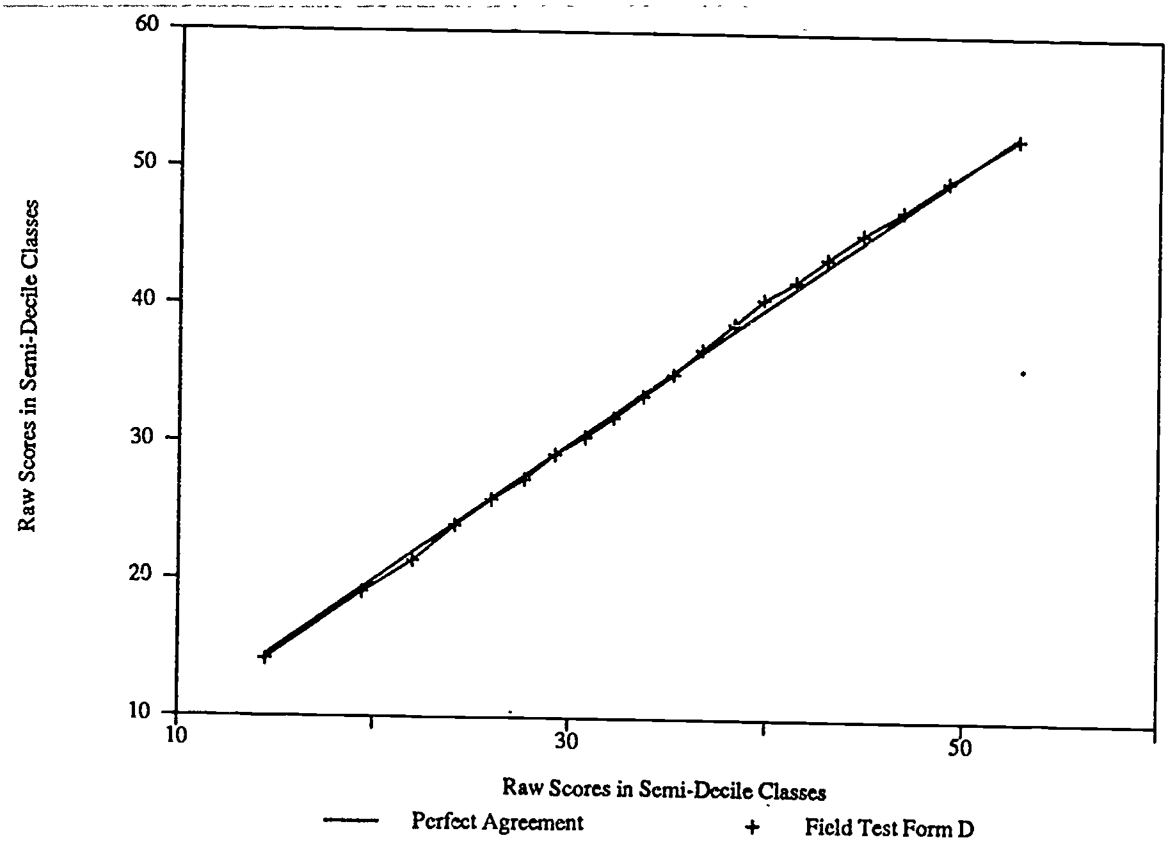

Figure 17. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core D (adjusted).

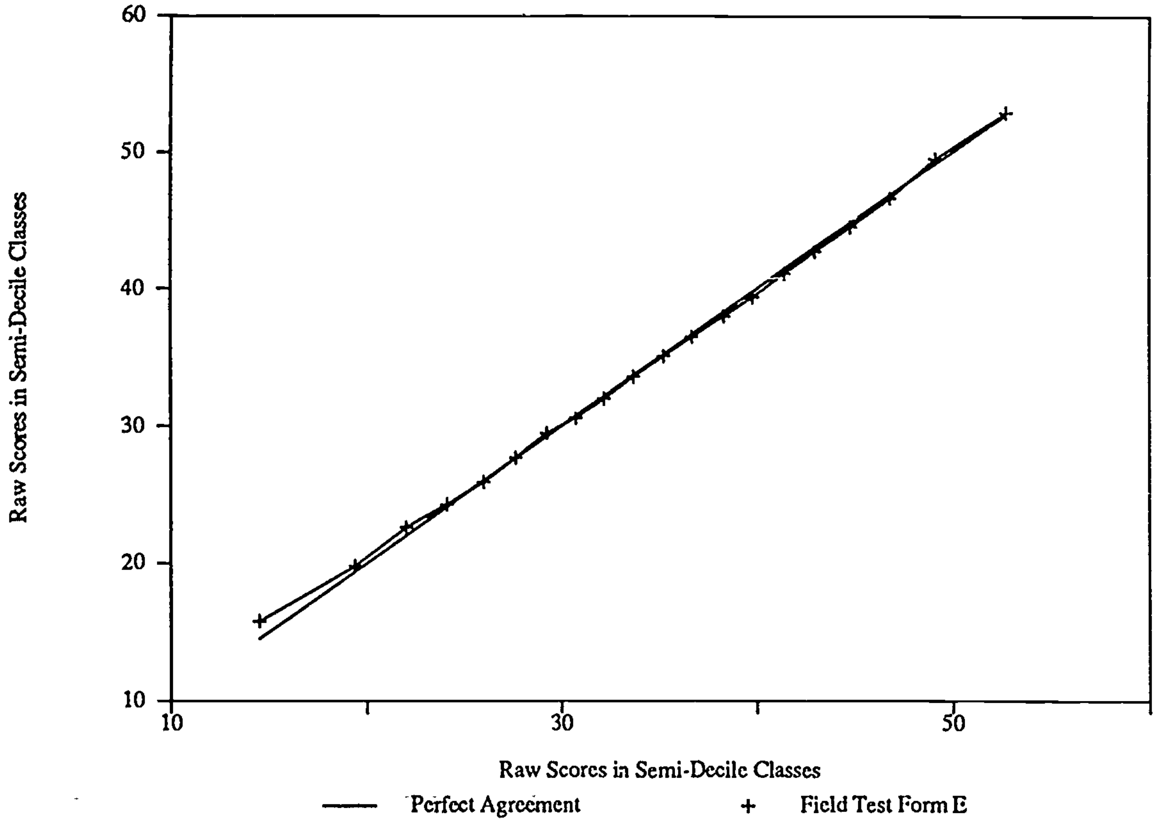

Figure 18. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core E (adjusted).

Of special significance to comparability of student scores across the years is the equivalence of the four future cores to the mean of the first test of record. An equipercentile analysis was made of the relationship of the four future core tests to the mean of the 1988 core tests of record. In Figures 19 though 22, the differences of the data points from perfect agreement are small. The results of the adjustments for each of the four future core tests are given in Table 2 and Figures 23 through 26. The required changes (the substitution of approximately eight items per form) were minimal. The length of the test was also changed from 58 to 56 items to be comparable with Forms B-E.

The adjustments to the core test assure the continuity of the norms table for future years while providing new test items each year. The new test items will prevent the loss of confidentiality, and therefore validity, that occurs with the continued use of the same items. Student scores will have a common reference point from 1988 onward, barring changes in the definition of the basic instructional program.

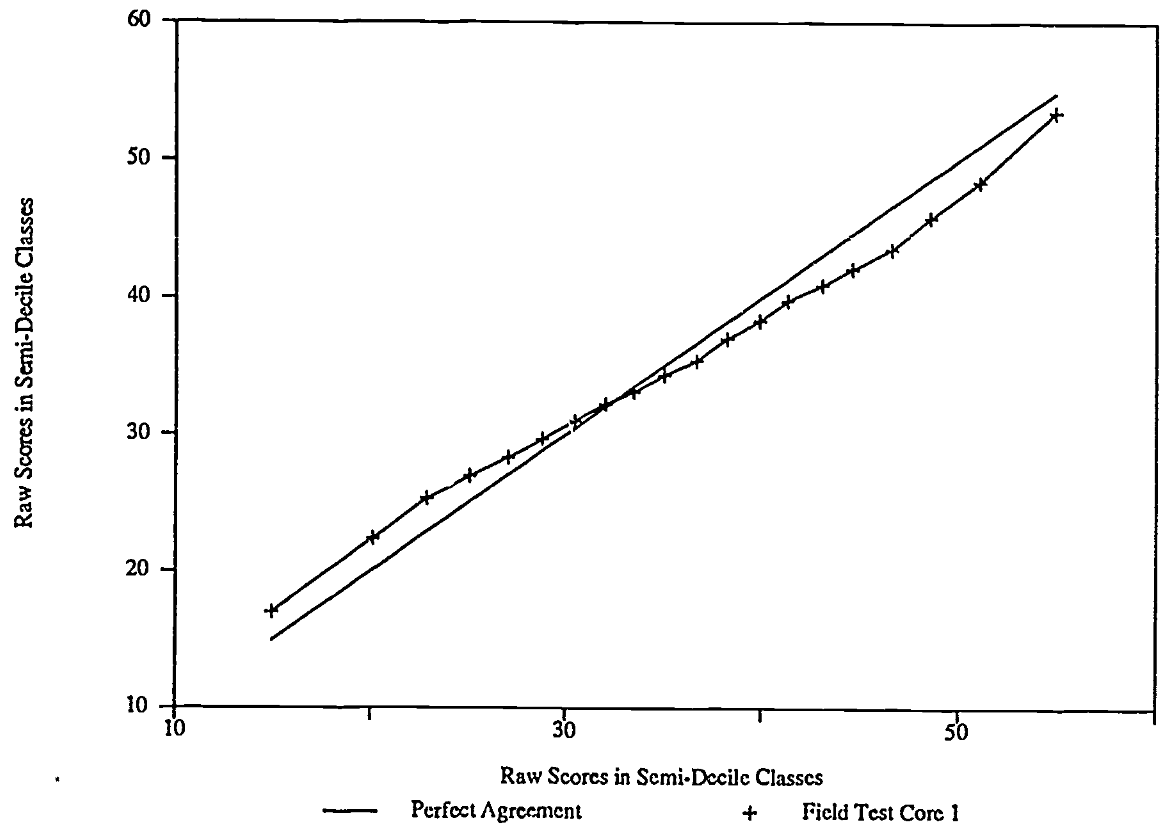

Figure 19. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core 1 (unadjusted).

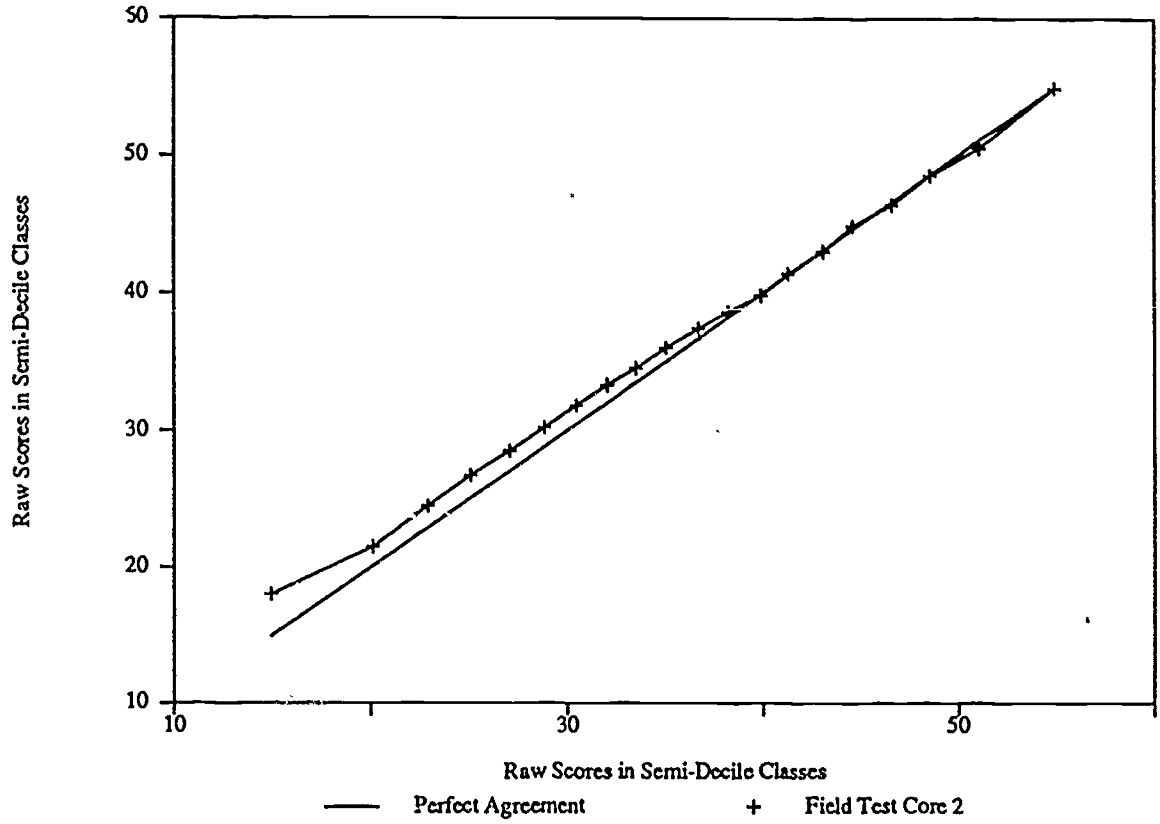

Figure 20. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core 2 (unadjusted).

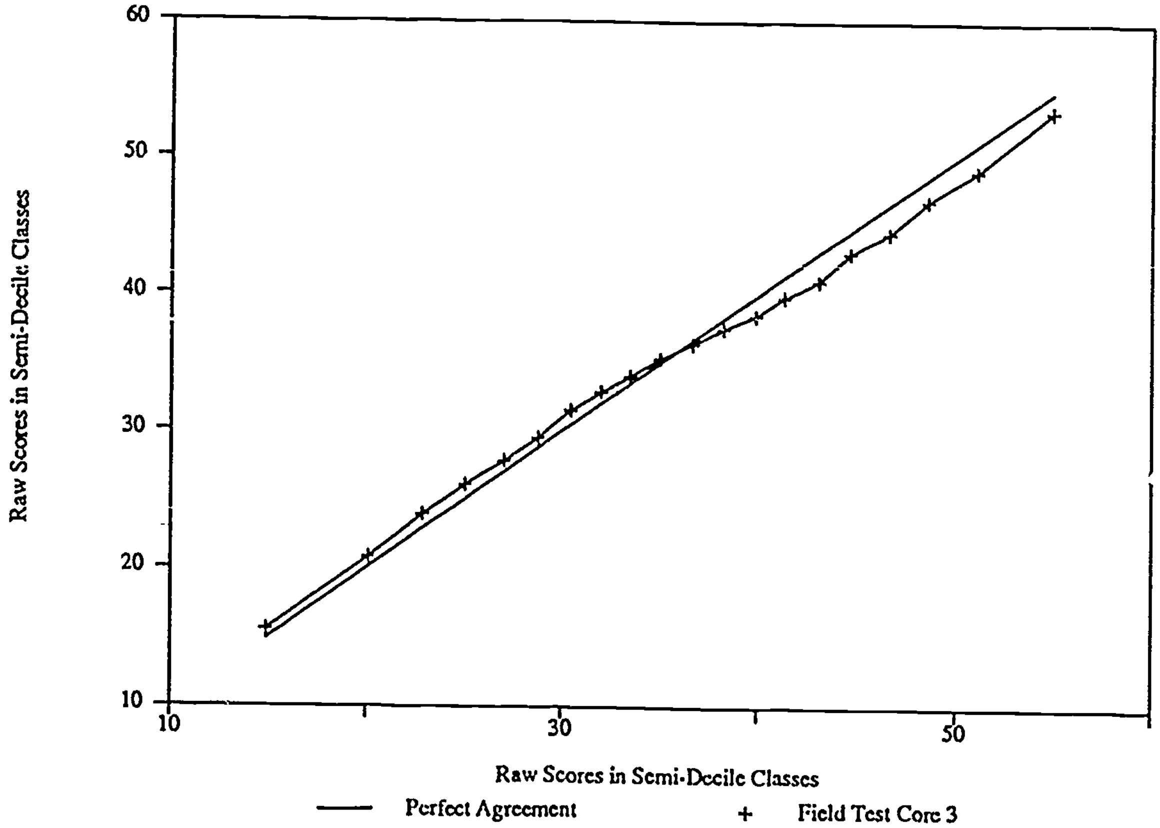

Figure 21. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core 3 (unadjusted).

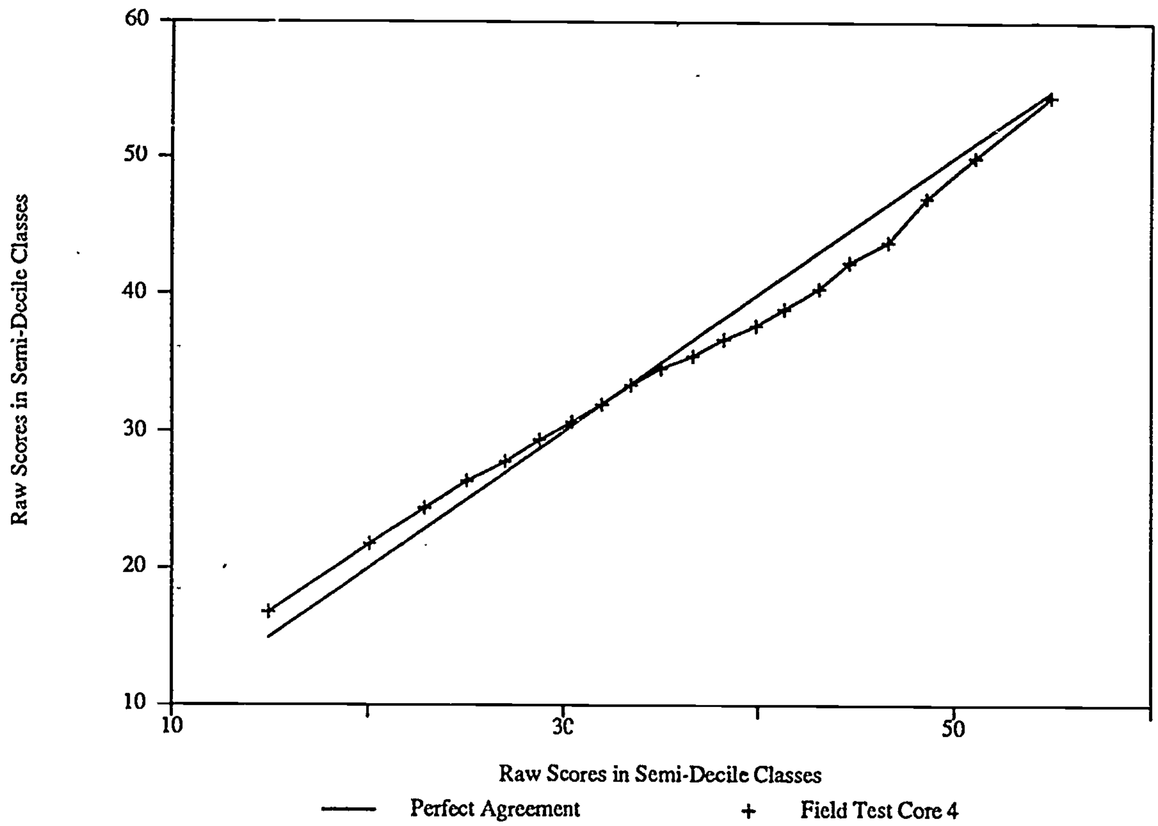

Figure 22. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core 4 (unadjusted).

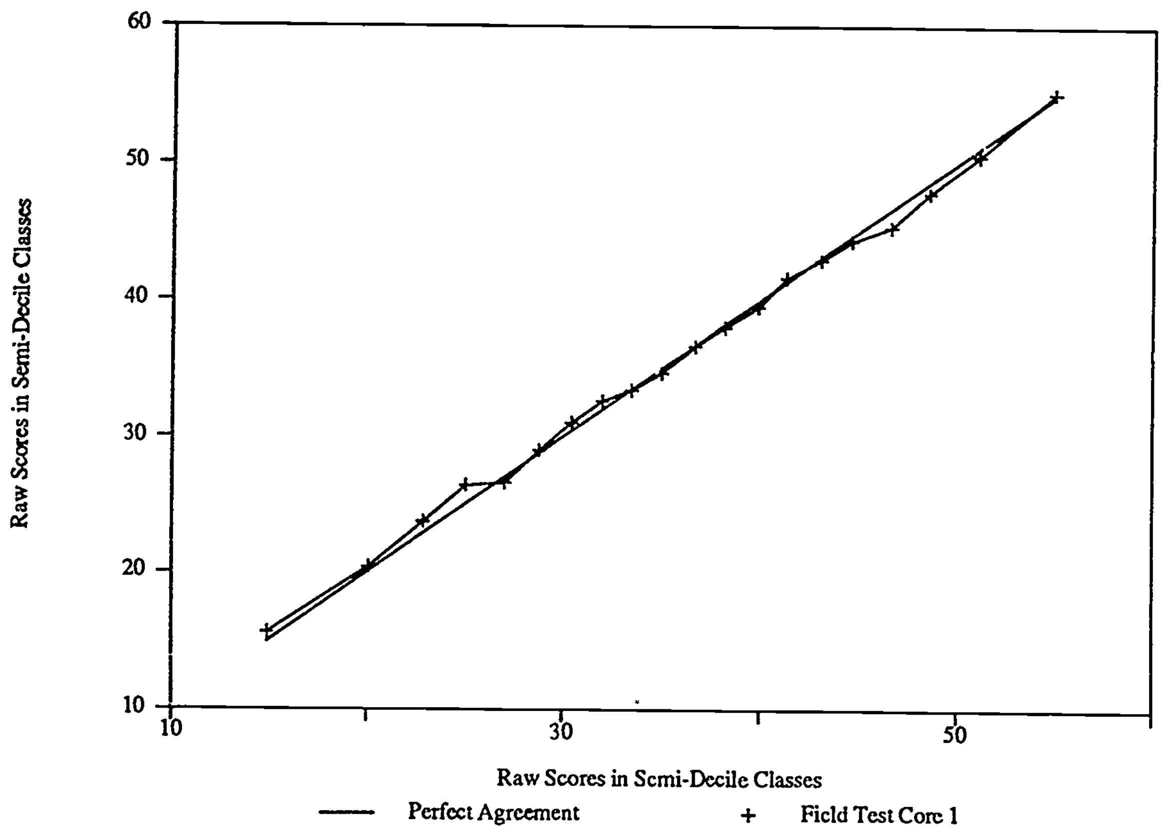

Figure 23. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core 1 (adjusted).

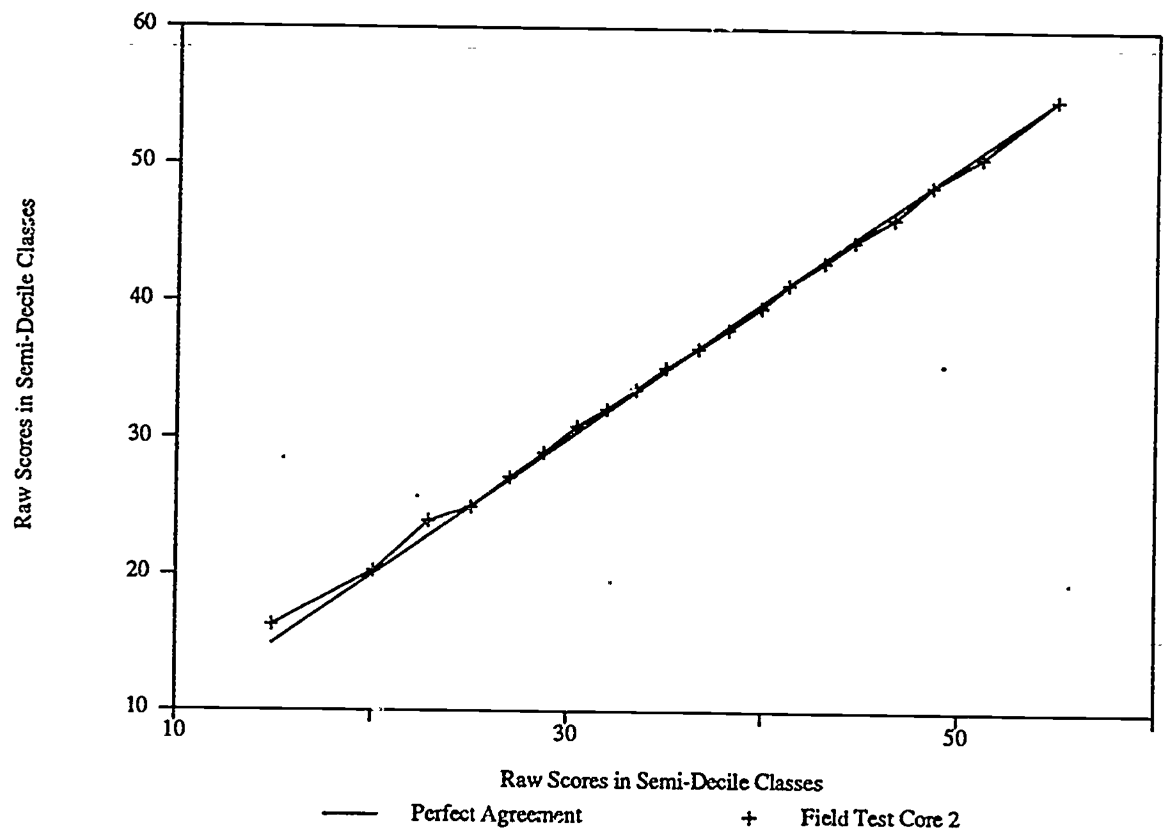

Figure 24. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core 2 (adjusted).

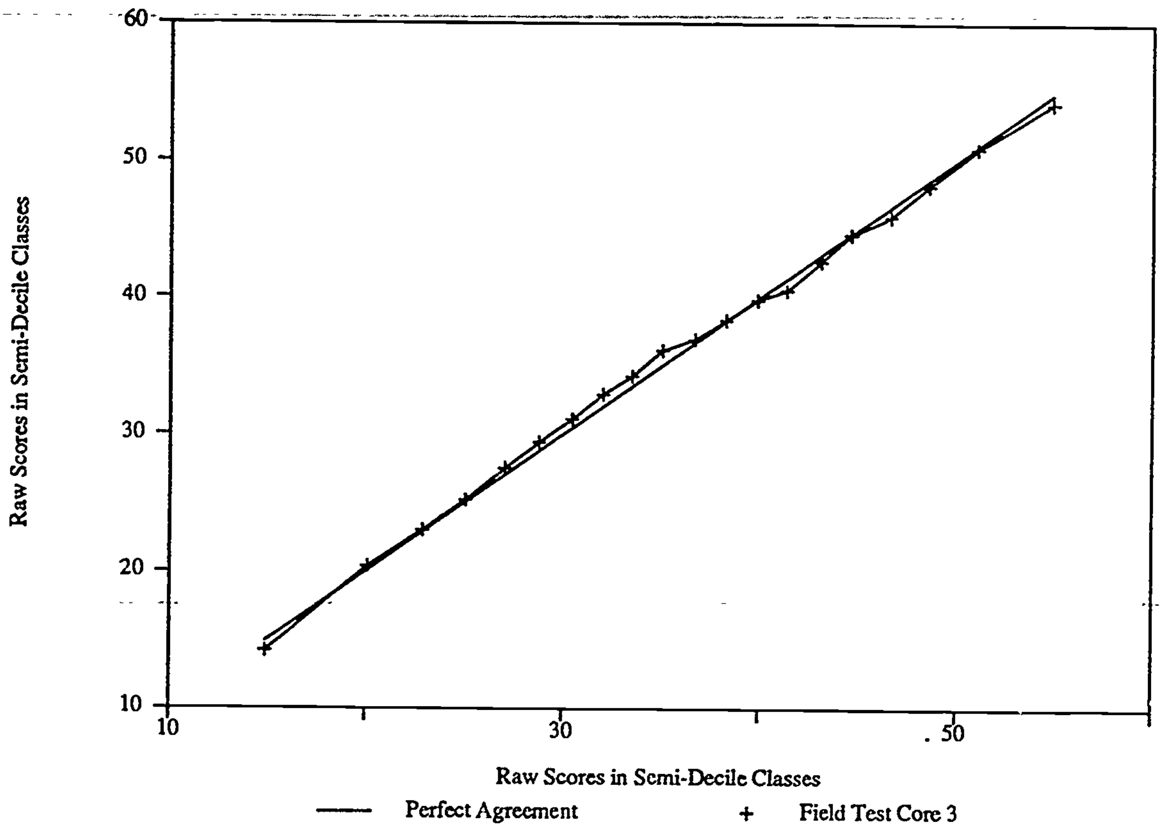

Figure 25. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core 3 (adjusted).

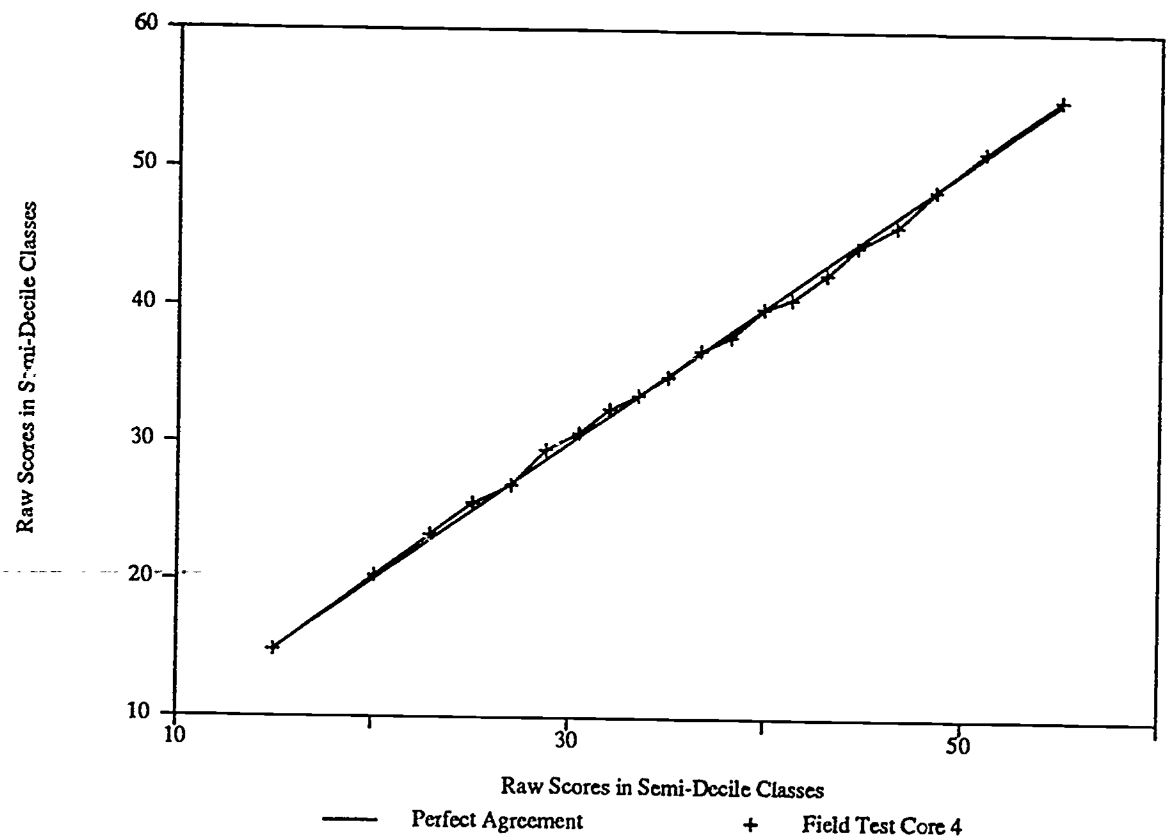

Figure 26. Equipercentile comparison of the 1987 North Carolina Test of Algebra II (all field-tested core tests) with field test core 4 (adjusted).

TEST NORMS

Students who answer all 56 of the Algebra II items correctly could be assumed to be excellent Algebra II students. If everyone answered all of the items correctly, however, a different interpretation would have to be placed on the scores. At some point, scores must have a reference grounded in the experience of all students. In some respects, at least, everything is good or bad by comparison. Norms tables provide that reference. Given a norms table, a student's score can be compared with other students' scores.

Norms tables commonly have two points of reference: a scale of percentiles and a scale of standard scores. The former permits the location of a score within percentile ranks; thus a student is said to have exceeded the performance of 80% of the students in the norm group (in this case, Algebra II students taking the North Carolina Test of Algebra II in May 1988). The latter, standard scores, permits the location of a score within normally-distributed standard scores. This reference is appropriate if the student abilities are believed to be normally distributed. In a normal distribution, raw scores are given greater and greater weight as they diverge from the mean in either direction.

The choice of a metric for the standard scores is arbitrary. To avoid inappropriate and confusing comparisons with some of the more common metrics, such as those employed in IQ scores or NCE scores, a metric having a mean of 50 and a standard deviation of 10 was chosen. Most curriculum research studies involving the summation of scores will find the standard score to be the statistic of choice.

The norms table for student scores on the North Carolina Test of Algebra II is given in Table 3. These scores set a baseline of comparison for present and future achievement in Algebra II. Thus a student score in 1988, 1989, and future years can be referenced to the scores of all 1988 Algebra II students in North Carolina.

Table 3

Norms for Student Scores on the North Carolina Test of Algebra II

Raw Score	Percentile	Standard Score[a]
56	99	68.5
55	99	67.6
54	97	66.6
53	95	65.7
52	93	64.8
51	91	63.8
50	89	62.9
49	87	62.0
48	84	61.0
47	82	60.1
46	79	59.2
45	76	58.2
44	73	57.3
43	70	56.4
42	67	55.4
41	64	53.6
40	61	52.6
39	58	51.7
38	55	50.8
37	52	49.8
36	48	48.9
35	45	47.9
34	42	47.0
33	39	46.1
32	36	45.1
31	33	44.2
30	30	43.3
29	27	42.3
28	24	41.4
27	22	40.5
26	19	39.5
25	17	38.6
24	15	37.7
23	13	36.7
22	11	35.8
21	9	34.9
20	8	33.9
19	6	33.0
18	5	32.1
17	4	31.1
16	3	30.2
15	2	29.3
Less than 15	2	

[a]Adjusted to a mean of 50 and a standard deviation of 10.0.

CONTENT OF THE TEST

The North Carolina Test of Algebra II is objective-referenced; that is, its reference is to a domain of objectives. This domain is mapped over a domain of items, where the items reflect the objectives, equal in kind and number except for random fluctuations. The Algebra II tests were designed to achieve an even assessment across all objectives; in short each objective was to be represented by the same number of items. This design is consistent with the concept of a domain of objectives mapped over by a domain of items. Although the objectives have unit weighting, the goals are weighted by the number of objectives assigned to them. From empirical analyses, this is a traditional aspect of curriculum development: the more important a goal is considered to be, the greater the number of objectives that will be developed for it. Thus, an underlying system of weights exists for curricular goals.

Appendix A lists each goal and objective and the numerical item representation for each objective as it appears on the 1988 North Carolina Test of Algebra II (forms B-E). In addition, the proportion of teachers rating each objective as basic to the Algebra II curriculum is listed.

Table 4 lists the difficulty level for all items on the North Carolina Test of Algebra II tested in 1988 in terms of p-values (proportion of all students answering the item correctly).

Table 4

Item Difficulty by Item Number for the North Carolina Test of Algebra II
Form B

Item	P-value	Item	P-value	Item	P-value	Item	P-value
B1	0.75	B15	0.86	B29	0.35	B43	0.79
B2	0.82	B16	0.50	B30	0.68	B44	0.37
B3	0.80	B17	0.79	B31	0.60	B45	0.50
B4	0.43	B18	0.62	B32	0.65	B46	0.83
B5	0.90	B19	0.63	B33	0.50	B47	0.58
B6	0.82	B20	0.80	B34	0.74	B48	0.43
B7	0.79	B21	0.40	B35	0.75	B49	0.60
B8	0.54	B22	0.77	B36	0.85	B50	0.75
B9	0.88	B23	0.47	B37	0.80	B51	0.61
B10	0.91	B24	0.54	B38	0.81	B52	0.70
B11	0.56	B25	0.38	B39	0.73	B53	0.41
B12	0.70	B26	0.49	B40	0.65	B54	0.62
B13	0.76	B27	0.35	B41	0.50	B55	0.33
B14	0.80	B28	0.64	B42	0.85	B56	0.68

Item Difficulty by Item Number for the North Carolina Test of Algebra II
Form C

Item	P-value	Item	P-value	Item	P-value	Item	P-value
C1	0.77	C15	0.73	C29	0.66	C43	0.66
C2	0.92	C16	0.75	C30	0.50	C44	0.50
C3	0.69	C17	0.56	C31	0.73	C45	0.93
C4	0.87	C18	0.68	C32	0.54	C46	0.59
C5	0.91	C19	0.66	C33	0.47	C47	0.48
C6	0.67	C20	0.84	C34	0.86	C48	0.50
C7	0.56	C21	0.57	C35	0.35	C49	0.57
C8	0.76	C22	0.72	C36	0.89	C50	0.69
C9	0.47	C23	0.74	C37	0.88	C51	0.49
C10	0.71	C24	0.41	C38	0.83	C52	0.74
C11	0.53	C25	0.55	C39	0.94	C53	0.37
C12	0.50	C26	0.54	C40	0.43	C54	0.69
C13	0.62	C27	0.45	C41	0.60	C55	0.39
C14	0.94	C28	0.39	C42	0.79	C56	0.45

Table 4 (cont.)

Item Difficulty by Item Number for the North Carolina Test of Algebra II
Form D

Item	P-value	Item	P-value	Item	P-value	Item	P-value
D1	0.81	D15	0.29	D29	0.56	D43	0.59
D2	0.69	D16	0.70	D30	0.68	D44	0.46
D3	0.45	D17	0.77	D31	0.49	D45	0.70
D4	0.79	D18	0.46	D32	0.55	D46	0.76
D5	0.87	D19	0.71	D33	0.70	D47	0.44
D6	0.73	D20	0.84	D34	0.68	D48	0.51
D7	0.71	D21	0.52	D35	0.94	D49	0.42
D8	0.45	D22	0.70	D36	0.80	D50	0.71
D9	0.90	D23	0.49	D37	0.71	D51	0.53
D10	0.60	D24	0.51	D38	0.86	D52	0.60
D11	0.84	D25	0.49	D39	0.56	D53	0.50
D12	0.58	D26	0.53	D40	0.68	D54	0.82
D13	0.51	D27	0.56	D41	0.83	D55	0.40
D14	0.77	D28	0.69	D42	0.67	D56	0.46

Item Difficulty by Item Number for the North Carolina Test of Algebra II
Form E

Item	P-value	Item	P-value	Item	P-value	Item	P-value
E1	0.89	E15	0.80	E29	0.50	E43	0.68
E2	0.45	E16	0.75	E30	0.52	E44	0.51
E3	0.90	E17	0.50	E31	0.76	E45	0.59
E4	0.78	E18	0.72	E32	0.43	E46	0.65
E5	0.97	E19	0.73	E33	0.45	E47	0.55
E6	0.91	E20	0.81	E34	0.88	E48	0.48
E7	0.64	E21	0.51	E35	0.79	E49	0.38
E8	0.50	E22	0.59	E36	0.91	E50	0.73
E9	0.93	E23	0.50	E37	0.78	E51	0.60
E10	0.81	E24	0.49	E38	0.86	E52	0.47
E11	0.78	E25	0.42	E39	0.48	E53	0.61
E12	0.55	E26	0.55	E40	0.89	E54	0.74
E13	0.96	E27	0.50	E41	0.69	E55	0.51
E14	0.44	E28	0.47	E42	0.85	E56	0.57

On the 1988 North Carolina Test of Algebra II, the 50 core objectives were accompanied by 6 variable objectives that varied across the four forms (as discussed earlier—see Table 1). These items contributed to the individual student scores and to curricular assessment. Each variable objective was answered by one-fourth of the students.

At the classroom level, 224 items were answered during each test administration by an average of four students. This procedure provided a database of approximately four items per core objective and one item per variable objective across four students. From this database of information, estimates of how various portions of the curriculum are being mastered in the classroom may be drawn. At the school, school district, and state level, the 224 items were answered by a much larger number of students: up to 9,000 students per item. This assures a more stable measurement, but does not include a larger number of objectives or items. The accumulation of item and objective information depends upon measurement across successive years.

The measurement afforded by the 24 variable objectives is critical to assessing curriculum mastery at the classroom, school, school district, and state levels. Each year of test administration adds to the database and gives a more detailed and comprehensive picture of curriculum success.

In summary, the utility of a test is its initial norms table, its statistical equivalence of core tests during each administration and from year to year, and its broad sampling of the curriculum across time.

APPENDICES

Appendix A

Test Content - Item Representation by Goal and Objective

Goal/Obj	Description	No. Items 1988	% Teachers Rating as Basic[a]
Goal 1	The learner will use the language of algebra.		
1.1	Use the order of operations and evaluate algebraic expressions.	4	99.2
1.2	Translate English words and phrases into mathematical language.	4	97.1
1.3	Use the properties of addition to simplify arithmetic and algebraic expressions (Additive Identity, Commutative, Associative, Additive Inverse).	4	98.4
1.4	Use the properties of multiplication to simplify arithmetic and algebraic expressions (Multiplicative Identity, Commutative, Associative, Multiplicative Inverse, Multiplication Property of Zero).	4	98.4
1.5	Use the Distributive Property of Multiplication over Addition to simplify arithmetic and algebraic expressions.	4	98.8
Goal 2	The learner will locate numbers on the number line and the coordinate graph.		
2.1	Graph sets of real numbers on the number line.	4	99.6
2.2	Graph ordered pairs of numbers on the coordinate plane and find the coordinates of points on the plane.	4	99.6
2.3	Graph linear equations in two variables.	4	97.1
2.4	Graph a relation on the cordinate plane.	1	92.6
2.5	Graph the solution sets of systems of linear inequalities in two variables.	1	93.4
2.6	Graph a function on the coordinate plane.	1	92.6
2.7	Graph the equations of a parabola, circle, ellipse, and hyperbola.	1	72.1

Goal/Obj	Description	No. Items 1988	% Teachers Rating as Basic[a]
Goal 3	The learner will perform operations with real numbers.		
3.1	Add real numbers.	4	97.1
3.2	Subtract real numbers.	4	96.3
3.3	Multiply real numbers.	4	96.7
3.4	Divide real numbers.	4	94.2
3.5	Use < or > to compare two numbers.	4	93.8
3.6	Simplify expressions involving positive, negative, and zero exponents.	4	95.1
3.7	Multiply and divide numbers written in scientific notation.	1	72.6
3.8	Write a rational number as a terminating or repeating decimal.	1	86.0
Goal 4	The learner will solve linear equations and inequalities.		
4.1	Solve equations in one variable.	1	96.7
4.2	Solve equations involving absolute value.	4	94.3
4.3	Solve equations with rational coefficients.	4	96.7
4.4	Solve literal equations and formulas.	4	90.2
4.5	Solve inequalities in one variable.	4	95.9
4.6	Solve inequalities involving absolute value.	1	93.4
Goal 5	The learner will solve systems of linear equations.		
5.1	Find the solution sets of open sentences in two variables with given replacements for the variables.	4	97.9
5.2	Find the solution sets of systems of two linear equations in two variables.	4	95.5
5.3	Use systems of two linear equations in two variables to solve problems.	4	89.0
5.4	Find the solution sets of systems of three linear equations in three variables.	4	69.8
5.5	Solve systems of linear equations by using Cramer's Rule.	4	36.0

Goal/Obj	Description	No. Items 1988	% Teachers Rating as Basic[a]
Goal 6	The learner will perform operations with polynomials.		
6.1	Add polynomials.	4	95.7
6.2	Subtract polynomials.	4	95.2
6.3	Multiply a polynomial by a monomial.	4	94.7
6.4	Multiply two binomials by using special product formulas (square of a binomial, product of the sum and difference of two binomials).	4	93.7
6.5	Multiply a binomial and a polynomial.	4	98.5
6.6	Find the quotient of two monomials.	4	99.0
6.7	Divide one polynomial by another of lower degree.	4	97.4
6.9	Factor monomials and find the greatest common factor (GCF) and least common multiple (LCM) of two or more monomials.	0	99.5
6.10	Factor special polynomials (perfect square trinomials, difference of two squares, sum or difference of two cubes).	4	97.9
6.11	Factor quadratic polynomials.	4	98.4
6.12	Use factoring to solve an equation.	4	100.0
6.13	Use polynomial equations to solve problems.	1	89.5
6.14	Use factoring to solve inequalities.	0	88.4
6.15	Factor polynomials completely.	4	95.8
Goal 7	The learner will perform operations with algebraic fractions.		
7.1	Write algebraic fractions in lowest terms.	4	99.0
7.2	Simplify products and quotients or rational algebraic expressions.	4	99.5
7.3	Simplify sums and differences of rational algebraic expressions.	4	99.5
7.4	Simplify complex fractions.	4	93.7
7 5	Solve fractional equations.	4	97.4

Goal/Obj	Description	No. Items 1988	% Teachers Rating as Basic[a]
Goal 8	The learner will solve problems involving radical expressions.		
8.1	Simplify roots of real numbers.	4	98.4
8.2	Simplify expressions involving fractional exponents.	4	92.2
8.4	Simplify expressions involving sums and differences of radicals.	4	96.3
8.5	Simplify expressions involving products and quotients of radicals.	4	96.9
8.6	Indicate the square root of a negative number as a complex number.	4	91.6
8.7	Solve equations which contain radical expressions.	1	88.5
Goal 9	The learner will solve quadratic equations.		
9.1	Complete the square to solve quadratic equations.	1	89.1
9.2	Use the Quadratic Formula to solve quadratic equations.	4	97.4
9.3	Use the discriminant of a quadratic equation to determine the nature of the roots.	1	77.6
9.4	Write a quadratic equation given its solution set.	1	75.0
9.6	Solve a system of two equations in which one or both of the equations are quadratic.	0	57.1
Goal 10	The learner will solve problems involving complex numbers.		
10.1	Add and subtract complex numbers.	1	88.5
10.2	Simplify expressions involving products and quotients of complex numbers.	0	88.0
10.3	Solve quadratic equations involving complex numbers.	1	80.1

Goal/Obj	Description	No. Items 1988	% Teachers Rating as Basic[a]
Goal 11	The learner will use analytic geometry to solve problems.		
11.1	Use the distance formula.	1	84.8
11.2	Determine the coordinates of the midpoint of a segment.	0	83.6
11.3	Find the slope of a line given two points on a line, an equation of the line, or the graph of a line.	4	98.4
11.4	Find an equation of a line given its slope and the coordinates of a point on the line, the coordinates of two points on the line, or its slope and y-intercept.	4	98.4
11.5	Determine if two lines are parallel or perpendicular by examining their slopes.	4	95.1
11.6	Use the Pythagorean Theorem and its converse to solve problems.	1	80.3
11.7	Write the equation of a circle from its geometric properties.	1	71.0
11.8	Identify parabolas, circles, elipses, and hyperbolas from their equations.	1	67.9
Goal 12	The learner will solve problems involving variation.		
12.1	Use direct variation to solve problems.	1	75.3
12.2	Use inverse variation to solve problems.	1	65.6
Goal 14	The learner will solve problems involving logarithmic and exponential functions.		
14.1	Write an exponential function as a logarithmic function and write a logarithmic function as an exponential function.	1	69.9
14.5	Solve problems using laws of exponents.	4	57.0

Goal/Obj	Description	No. Items 1988	% Teachers Rating as Basic[a]
Goal 15	The learner will investigate some techniques for problem solving.		
15.1	Solve "word problems" (number, age, coin, perimeter, digit, work, uniform motion).	4	96.2
15.2	Use inequalities as well as equations to solve "word problems".	0	82.2
15.3	Solve "word problems" involving fractional equations.	1	90.9
15.4	Use quadratic equations to solve verbal problems.	1	88.0

[a]Percentage of Algebra II teachers rating the objective as basic to instruction in Algebra II.

Appendix B

Goals and Objectives Rejected for Use

Goal/Obj	Description	% Teachers Rating as Basic[a]
Goal 5	The learner will solve systems of linear equations.	
5.5	Solve systems of linear equations by reducing the matrix of coefficients to triangular form.	23.6
Goal 6	The learner will perform operations with polynomials.	
6.8	Use synthetic division to divide a polynomial by a linear binomial.	55.2
6.16	Expand the powers of binomials using Pascal's triangle or the binomial theorem.	31.1
6.17	Use the binomial theorem to find a specified term of an expansion.	23.8
Goal 8	The learner will solve problems involving radical expressions.	
8.3	Estimate the value of radicals representing irrational numbers.	56.5
Goal 9	The learner will solve quadratic equations.	
9.5	Find the sum and product of the solutions of a quadratic equation.	65.1
Goal 12	The learner will solve problems involving variation.	
12.3	Use joint variation to solve problems.	52.7
Goal 13	The learner will solve problems involving arithmetic and geometric sequences and series.	
13.1	Complete arithmetic sequences (progressions) and find arithmetic means.	32.8
13.2	Find a given term in an arithmetic sequence (progression).	33.3

Goal/Obj	Description	% Teachers Rating as Basic[a]
13.3	Find the sums of arithmetic series and use summation notation.	29.6
13.4	Complete geometric sequences (progressions) and find geometric means.	30.1
13.5	Find a given term in a geometric sequence (progression).	29.6
13.6	Find the sum of a finite geometric series.	29.0
13.7	Find the sum of a geometric series having infinitely many terms.	23.7
Goal 14	The learner will solve problems involving logarithmic and exponential functions.	
14.2	Use a scientific calculator to evaluate products, quotients, powers, and roots.	45.2
14.3	Find the logarithms of numbers by using a base other than 10.	50.0
14.4	Solve problems involving exponential equations.	54.3
Goal 15	The learner will investigate some techniques for problem solving.	
15.5	Use the Fundamental Counting Principle to solve problems.	24.7
15.6	Solve problems involving permutations.	8.7
15.7	Solve problems involving combinations.	9.7
15.8	Solve simple probability problems.	11.4
Goal 16	The learner will use trigonometry to solve problems.	
16.1	Find the sine and cosine of an angle in standard position, given a point, other than the origin of the terminal side of the angle.	33.0
16.2	Find the values of the six trigonometric functions of an angle, given a point other than the origin on the terminal side of the angle, or given the value of one of its functions and quadrant in which its terminal side lies.	31.9
16.3	Find the trigonometric function values of acute angles, in particular those of 30°, 45°, and 60° angles.	35.1

Goal/Obj	Description	% Teachers Rating as Basic[a]
16.4	Use a scientific calculator or tables to find approximations of the values of trigonometric functions for an angle.	32.4
16.5	Solve a right triangle, given either the measures of a side and an acute angle or the measures of two sides and learn to solve problems involving right triangles.	36.8

[a]Percentage of Algebra II teachers rating the objective as basic to instruction in Algebra II.