DOCUMENT RESUME

ED 324 212 SE 051 591

TITLE	The North Carolina Test of Geometry. Technical Characteristics. Forms 1-8.
INSTITUTION	North Carolina State Dept. of Public Instruction, Raleigh. Div. of Accountability Services/Research.
PUB DATE	90
NOTE	42p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Curriculum Evaluation; *Geometry; Mathematics Education; Mathematics Skills; Secondary Education; *Secondary School Mathematics; *Test Construction; Test Content; *Testing Programs; Test Norms; Test Reliability; *Test Theory; Test Validity
IDENTIFIERS	*North Carolina Test of Geometry

ABSTRACT

To facilitate the proper technical use of the test scores obtained from the administration of the tests, the curricular and psychometric characteristics of the tests are described in a series of technical manuals. This manual, the eighth in the series, contains a description of the characteristics of the North Carolina Test of Geometry. This paper describes the performance of test items individually and the test generally. The test was developed for use as an achievement test following the completion of the Geometry course of study. Its design was to serve as a normative measurement of student achievement and as an objective-based measurement of curriculum coverage. The curricular, instructional, content, and concurrent validity of the test are discussed. Included are: (1) the method for deriving test scores; (2) reliability and other statistics; (3) a curriculum assessment; (4) content of the test; and (5) test norms. The utility of the test has been determined by its statistical equivalence of core tests from year to year, its broad sampling of the curriculum across time, and its initial norms table. The goals and objectives used to assemble the test and the number of questions associated with each is appended. The percent of teachers reporting each goals/objective as basic to instruction in Geometry as well as a list of goals/objectives that were rejected for use are also appended. (KR)

Forms 1 – 8

Technical Characteristics of the North Carolina Test of Geometry

NCTests

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

Published 1990

Technical Characteristics of the

North Carolina Test of

Geometry

Prepared by
Eleanor E. Sanford

NCTests

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

FOREWORD

The NCDPI Division of Accountability Services/Research, in cooperation with the NCDPI Division of Curriculum and Instruction Services, 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 achievement tests for students taking Algebra I, Algebra II, Biology, Chemistry, English I, Geometry, Physics, and U.S. History. Physical Science and Economic, Legal, and Political Systems will be added in 1991, and other tests are being planned.[a]

To facilitate the proper technical use of the test scores obtained from the administration of the tests, the curricular and psychometric characteristics of the tests are described in a series of technical manuals. This manual, the eighth in the series, contains a description of the characteristics of the North Carolina Test of Geometry.

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

Table of Contents

Foreword ii

List of Tables iv

List of Figures v

Description 1

Validity 2

- Curricular Validity 2
- Instructional Validity 2
- Content Validity 3
- Concurrent Validity 8

Method for Deriving Score 9

Reliability and Other Statistics 10

Curricular Assessment 15

Content of the Test 16

Norms 25

Appendices 27

- A: Test Content—Item Representation by Goal and Objective 28
- B: Goals and Objectives Rejected for Use 33

List of Tables

1 Organization of the North Carolina Test of Geometry 1

2 Core Development of the North Carolina Test of Geometry 7

3 Descriptive Statistics of the 1989 North Carolina Test of Geometry 11

4 Item Difficulty by Item Number of the NCT–Geometry—Core 1 17

5 Item Difficulty by Item Number of the NCT–Geometry—Core 2 18

6 Item Difficulty by Item Number of the NCT–Geometry—Core 3 19

7 Item Difficulty by Item Number of the NCT–Geometry—Core 4 20

8 Item Difficulty by Item Number of the NCT–Geometry—Core 5 21

9 Item Difficulty by Item Number of the NCT–Geometry—Core 6 22

10 Item Difficulty by Item Number of the NCT–Geometry—Core 7 23

11 Item Difficulty by Item Number of the NCT–Geometry—Core 8 24

12 Norms for Student Scores on the North Carolina Test of Geometry 26

List of Figures

1 Comparison of letter grades teachers expected students to receive and scores subsequently earned on the 60-item 1989 North Carolina Test of Geometry—Cores 1–8 (N = 43,067 students) 8

2 Frequency distribution of scores on the 60-item 1989 North Carolina Test of Geometry—Cores 1–8 (N = 43,325 students) 12

Description

The North Carolina Test of Geometry (NCT–Geometry) was developed for use as an achievement test following the completion of the Geometry course of study. Its design serves a dual purpose: that of a normative measurement of student achievement and that 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 55 objectives and a set of five objectives selected randomly from 10 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 basic set of objectives taught in the Geometry course of instruction. Four forms of the test are administered in each classroom, one form per student. Under this system, a fourth of the students in the classroom will take Form 1 of the test, a fourth will take Form 2, and so on (see Table 1).

Table 1
Organization of the North Carolina Test of Geometry

55 Basic Objectives			
5 Variable Objectives	5 Variable Objectives	5 Variable Objectives	5 Variable Objectives
60 Items Form 1	60 Items Form 2	60 Items Form 3	60 Items Form 4

The normative student scores are based on the 60-item total score. Curriculum assessment is achieved by combining the results from all four forms of the test administered at one time for a total of 240 items.

Validity

The development of a Geometry achievement test has two purposes. The first is to obtain 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 Geometry curriculum. The second is to assess the degree to which the curriculum has been mastered by 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 may not.

Theoreticians state that only inferences concerning test scores may be said to have validity. Generally, readers understand this, and this manual will employ the convenient shorthand 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. The validity of a test concerns what the test measures and how well it does so. 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 Geometry, this was done through a cooperative effort, led by NCDPI Program Services/Division of Curriculum and Instruction Services, involving curriculum specialists, teachers, administrators, university professors, and others. The result was a list of 14 goals encompassing 97 objectives (see Appendices A and B). Supported by expert opinion and a statewide consensus, these goals and objectives were approved by the State Board of Education in 1985 as the basis for instruction in Geometry. Curricular validity, the first step in establishing test validity, was established by this method.

Instructional validity

A basic course of study may not include all of the objectives taught under various circumstances in Geometry. For example, some advanced classes may cover material that would be beyond the reach of 95% of all Geometry students. For this reason, it becomes important to know just what is being taught in the majority of Geometry classes in the state. To determine this, all Geometry teachers in North Carolina were surveyed in May 1986 (N = 915). The analysis of results was based on 749 responses, or 82% of all possible responses.

The Geometry teachers examined the 97 objectives and noted whether they taught each objective every year and whether each objective was basic to instruction. Of the objectives, 86 were rated as basic to instruction by at least 50% of the responding teachers. After deliberation by curriculum specialists, 22 of the Geometry objectives were removed from further assessment for various reasons (for example, they involved geometric constructions or they concerned advanced topics in coordinate geometry). During the item writing process it was determined that one goal encompassing seven objectives (geometric constructions) and three other objectives (two-column format proofs, indirect proofs, and trigonometric relationships) should not be assessed by a multiple-choice achievement test. It was determined that the remaining 13 goals and 65 objectives formed the basic curriculum for Geometry.

The objectives used in test development are listed in Appendix A and the objectives that were rejected for use are listed in Appendix B, together with the proportion of teachers judging each one as basic.

Instructional validity, the second step in defining test validity, was established by these procedures. It defines the inferences that can be drawn from the Geometry scores.

In summary, it was concluded that curricular and instructional validity depended jointly on the 13 goals and 65 objectives under which they were collected and that the Geometry test should be built on that foundation.

Content validity

Content validity—the degree to which test items reflect the basic instructional program—is a quality commonly referenced in evaluating achievement tests. Content validity is built into a test from the beginning, and the procedures relating to the content validation of the North Carolina Test of Geometry are described below.

Content validity of the item pool. The content validity of the item pool was defined through a number of operations:

First, the item pool for the Geometry test was created in 1987. It was specified that the pool would have 750 items, with 10 items per objective. For four objectives, the items to be written were specified to be in proof format (3.5 and 3.6—10 out of 10 items; 8.2—4 out of 10 items; and 9.6—5 out of 10 items). The items were developed by ten North Carolina Geometry teachers trained in the technical aspects of item-writing. The use of classroom teachers from across the state helped to insure that instructional validity was maintained, since the items would be drawn from their classroom experiences. All total, 721 multiple choice items and 29 proof items were written for the Geometry item pool.

Second, the item pool was edited for grammar, syntax, psychometric form, linguistic bias, and subject area content. Four multiple choice and nine proof items were deleted from the Geometry item pool at this time.

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 eight classroom teachers—one from each of the educational regions across the state. The criteria for evaluating each item included the following:

- conceptual—objective match, fair representation, lack of cultural bias, clear statement, single problem, one best answer, common context in foils, each foil credible
- language—appropriate for age; correct punctuation, spelling, and grammar; lack of excess words; no stem/foil clues; no negatives in foils
- format—logical order of foils; familiar presentation style, print size, and type; correct mechanics and appearance; equal foil lengths
- diagram—necessary, clean, relevant, unbiased

Four multiple choice items were deleted from the Geometry item pool at this time. The 80 items written for objectives 10.4 and 13.8-13.14 dealing with trigonometry and geometric constructions were held for development at a later time.

Fourth, the items were collected into ten test forms for field testing. Although the forms were not the final forms of the North Carolina Test of Geometry, they were organized in such a way that the objectives were represented equitably across all forms. Each form contained 72 or 73 multiple choice items and two proof items. Ten of the multiple choice items were common across all forms of the test for the purpose of ability equating should that become necessary. Due to the procedure and time involved in scoring the two proof items, the decision was made that proofs would be assessed by a separate five-item test administered in the early spring. These items will not be discussed further in this manual.

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

Sixth, a sample of 7,973 students was selected to take the ten field test forms containing a total of 623 items. To insure broad representation, schools were selected from each of the eight North Carolina educational regions and were representative of the state based on criteria that were judged to be at least partially related to Geometry ability levels—school performance on the 1987 NCT–Algebra I and NCT–Algebra II. The ten field tests were interleaved in all student samples, producing an even spread of ability across all of the tests. Consequently, each item was answered by approximately 707 students (the number of students per field test form ranged from 655 to 744).

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). Eighteen statistics were assembled for each item, i.e., p-value, Rasch difficulty index, adjusted Rasch difficulty index, standard error of the mean, fit mean-square, item validity (point-biserial correlation), and the item characteristic curve groupings. Item bias due to gender or ethnicity was examined by computing the partial correlation between the item score and gender/ethnicity while controlling for total score.

The item statistics were submitted to computer analysis using a program designed to scan a range of statistics for an item and print out an appropriate psychometric notation based on the criteria that had been built into the program specifically for Geometry. An item was classified as "too hard" if the p-value was less than .30 or as "too easy" if the p-value was greater than .93. An item was said to have "weak prediction" if the point-biserial correlation was less than .18. 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. An item was said to exhibit potential "gender" bias if the partial correlation with gender was more extreme than ± .133, and to exhibit potential "ethnic" bias if the partial correlation with ethnicity was more extreme than ± .1375.

The content of Geometry cannot be represented by a single factor. Therefore, maximization of item-total (point-biserial) correlations was not a goal of item development. Once an item was shown to have at least modest correlation with a corrected total score (.18 or greater) and was judged to measure an objective, it was included in the item pool. 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.

This information was placed on the item record, which became the basic document to which all other records are referred. The item record contains the goal, objective, historical information, a copy of the item itself, the item field-test statistics, and the psychometric notations. Each item has a separate item record.

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

Of the 623 items field-tested in 1988, 74 (12%) were deleted from the item pool. In content areas, these items are usually discarded. In Geometry, it was felt that an analysis of defective items was possible, and that the curriculum specialist could revise or rewrite the items to bring them within a usable range of values. Of the 74 deleted items, 68 were revised or rewritten for future use. The revised items were employed randomly throughout the selection of items for tests in order to assure that their effects would fall equally across all forms. The revised Geometry item pool contained 617 items for future test development.

Content validity of the test. After a consideration of the logistics involved, it was decided to prepare eight complete core tests (60 items) for administration in May 1989. Each core test was based on a random selection of items from the 55 Geometry core objectives, for a total of 55 items randomly chosen from the approved item pool (one item per objective). The five additional items were randomly chosen from the ten variable objectives (one item per objective). Thus, the content of the test cores directly reflected all of the decisions that had been made earlier.

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 and curricular characteristics. This was determined by the analyses of the item field-test data, which was used to verify the psychometric and curricular adequacy of the item pool and to direct where item revisions should be made.

After the 60-item core tests were assembled, they were reviewed by one curriculum supervisor and two teachers in each of the eight educational regions. The criteria for evaluating each core test included the following:

- that the content of the test should reflect the goals and objectives taught
- that the items should be clearly and concisely written, and the vocabulary appropriate to the target age level
- that the content should be balanced in relation to ethnicity, sex, socioeconomic status, and geographic district of the state
- that each item should have one and only one answer that is right; however, the distracters should appear plausible for someone who has not achieved mastery of the represented objective

The ratings of the 1989 North Carolina Test of Geometry (Cores 1–8) were high to superior on all of the criteria.

Although the initial equating of the core tests depended upon the random selection of items from the item pool, the final equating was based on the statistics obtained at the time the first test of record was administered (see Table 2). 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 another and from one form to another.

Standardization sample. The first North Carolina Test of Geometry consisted of eight forms (Cores 1–8) administered simultaneously, each form containing 60 unique items. This test was administered to 43,325 North Carolina Geometry students in May 1989. The state norm population comprises these 43,325 students.

The eight forms of the test (Cores 1–8) were interleaved in all classrooms, and this produced an even spread of ability across all of the forms. The agreement of the mean core tests supported the view that the initial equating process described above was successful (see Tables 2 and 3).

Table 2
Core Development of the North Carolina Test of Geometry

Core	Process	Mean of P-values for All Items	SD of P-values of All Items	Sum of P-values of All Items[a]
1	Design	.602	.199	32.52
	Field Test	.630	.155	37.77
	Statewide Test	.622	.198	37.29
	Equate	.627	.196	37.60
2	Design	.602	.173	32.50
	Statewide Test	.642	.172	38.49
	Equate	.627	.170	37.60
3	Design	.601	.175	33.03
	Statewide Test	.624	.171	37.46
	Equate	.627	.173	37.60
4	Design	.600	.185	32.97
	Statewide Test	.622	.191	37.33
	Equate	.627	.188	37.60
5	Design	.601	.179	33.63
	Statewide Test	.604	.177	36.24
	Equate	.627	.183	37.60
6	Design	.601	.152	33.64
	Statewide Test	.628	.152	37.69
	Equate	.627	.152	37.61
7	Design	.601	.168	33.04
	Statewide Test	.622	.169	37.34
	Equate	.627	.171	37.60
8	Design	.600	.187	32.98
	Statewide Test	.645	.184	38.70
	Equate	.627	.182	37.60

[a] Sum of p-values for Design Process based on all non-revised items. The number of revised or new items per form ranged from four to six.

Concurrent validity of the test

When the 1989 North Carolina Test of Geometry was administered, Geometry teachers were asked to indicate the expected final letter grade for each student in their class. Figure 1 displays a comparison of letter grades in Geometry and the mean NCT–Geometry core score corresponding to each letter grade for the overall student population. The figure corresponds closely to expectation and adds to the evidence concerning the validity of the test.

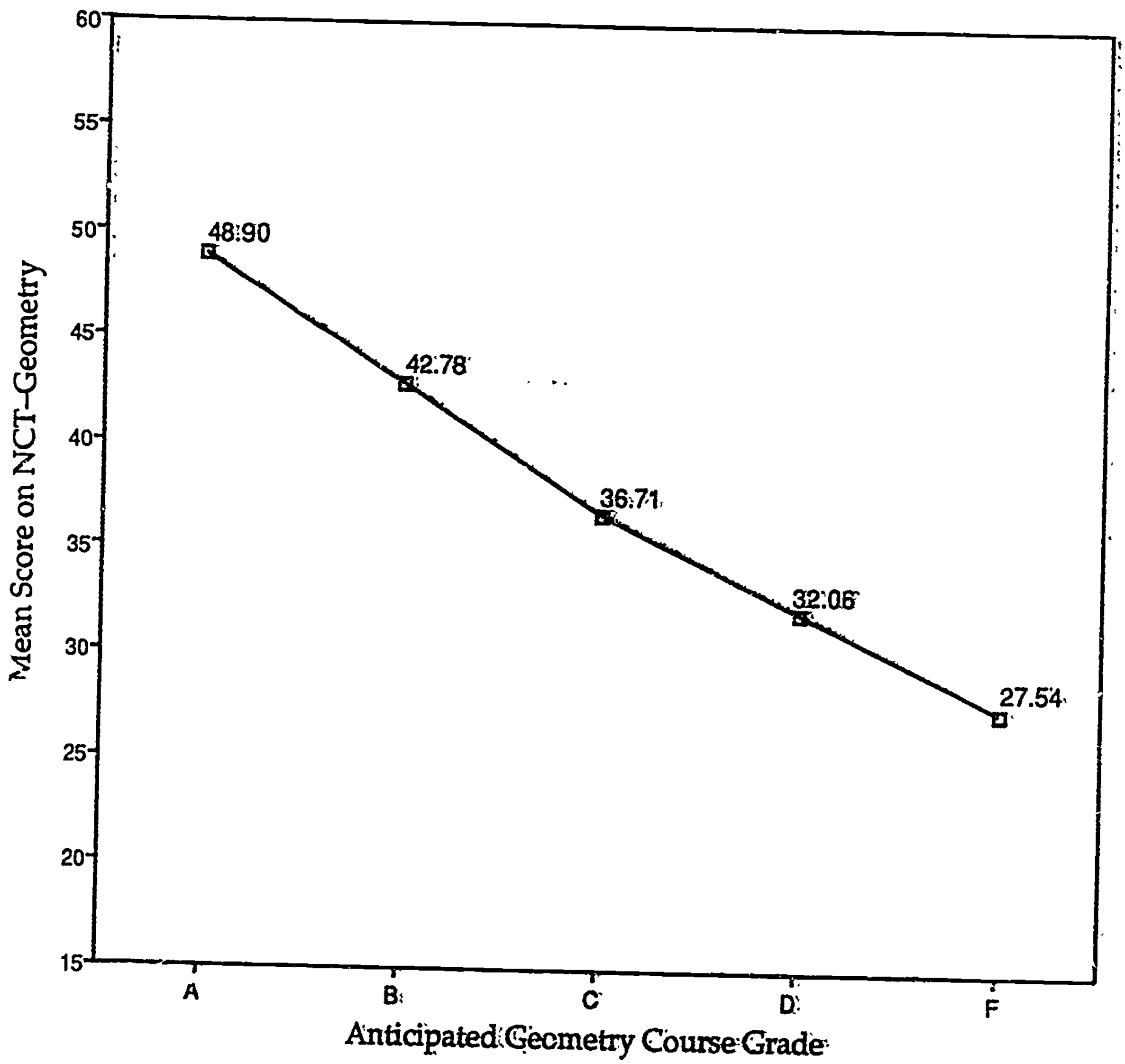

Figure 1. Comparison of letter grades teachers expected students to receive and scores subsequently earned on the 60-item 1989 North Carolina Test of Geometry—Cores 1–8 (N = 43,067).

Method for Deriving Test Scores

Item information was available to support the classical scoring model and the Rasch scoring model. The classical scoring model gives unitary weight to each item; a correct choice adds one to the total score; an incorrect choice adds zero. The one-parameter Rasch model also uses unitary weighting. (The two- and three-parameter item response models give 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 Geometry because it is fundamentally sound, simple to use, and easy to interpret. Each student's score consists of the sum of the right answers to the 60 items.

Reliability and Other Statistics

The descriptive statistics, the standard errors of measurement, the alternate form reliability estimates, and the alpha reliability coefficients for the first statewide administration of the North Carolina Test of Geometry in May 1989 (Cores 1–8) are given in Table 3.

To determine the alternate form reliability, a random sample of 827 students was selected to participate in the field testing. To insure broad representation, schools were selected from each of the eight North Carolina educational regions and were representative of the state based on criteria that was judged to be at least partially related to Geometry ability levels—school performance on the 1988 NCT–Algebra I and NCT–Algebra II. A total of 713 students in 13 schools actually participated in the core field testing and were administered Core 1 in May 1989. The agreement between the mean core field test (Core 1) of the sample group (mean = 37.77 and $s = 9.32$) and the state norm mean (aggregated across Cores 1-8—see Table 3) indicates that the field core sample was representative of the North Carolina Geometry student population. Sample students' scores on Core 1 and the eight cores administered statewide were matched and correlated to obtain estimates of the alternate form reliability of the North Carolina Test of Geometry.

The alternate form reliability estimates obtained from the small matched samples were averaged through z-score transformations to obtain a more stable estimate of the reliability. The average weighted alternate form reliability of the NCT–Geometry is .83. The alpha reliability estimates have a mean value of .90. The 1989 core scores are symmetrically distributed about a mean of 37.5, or 62.5% correct (see Figure 2).

Table 3
Descriptive Statistics of the 1989 North Carolina Test of Geometry

Core	N	Mean	SD	Median	se_{meas}[a]	Reliability Alternate Form[b]	Reliability Coefficient Alpha
89/1–8	43,325	37.5	10.3	37	4.25	.83	
1	5,630	37.4	9.85	37	3.27		.89
2	5,588	38.5	10.06	39	3.18		.90
3	5,552	37.5	10.36	37	3.28		.90
4	5,446	37.3	10.03	37	3.17		.90
5	5,403	36.2	11.08	36	3.32		.91
6	5,315	37.7	10.52	38	3.33		.90
7	5,243	37.3	10.17	37	3.22		.90
8	5,148	38.7	9.76	38	3.24		.89

[a] Based on the internal consistency estimate of reliability where available.
[b] Matched samples of students administered Field Core 1 and one of the eight cores statewide (n = 713). Form 1, n = 90; Form 2, n = 87; Form 3, n = 88; Form 4, n = 82; Form 5, n = 89; Form 6, n = 83; Form 7, n = 85; and Form 8, n = 77.

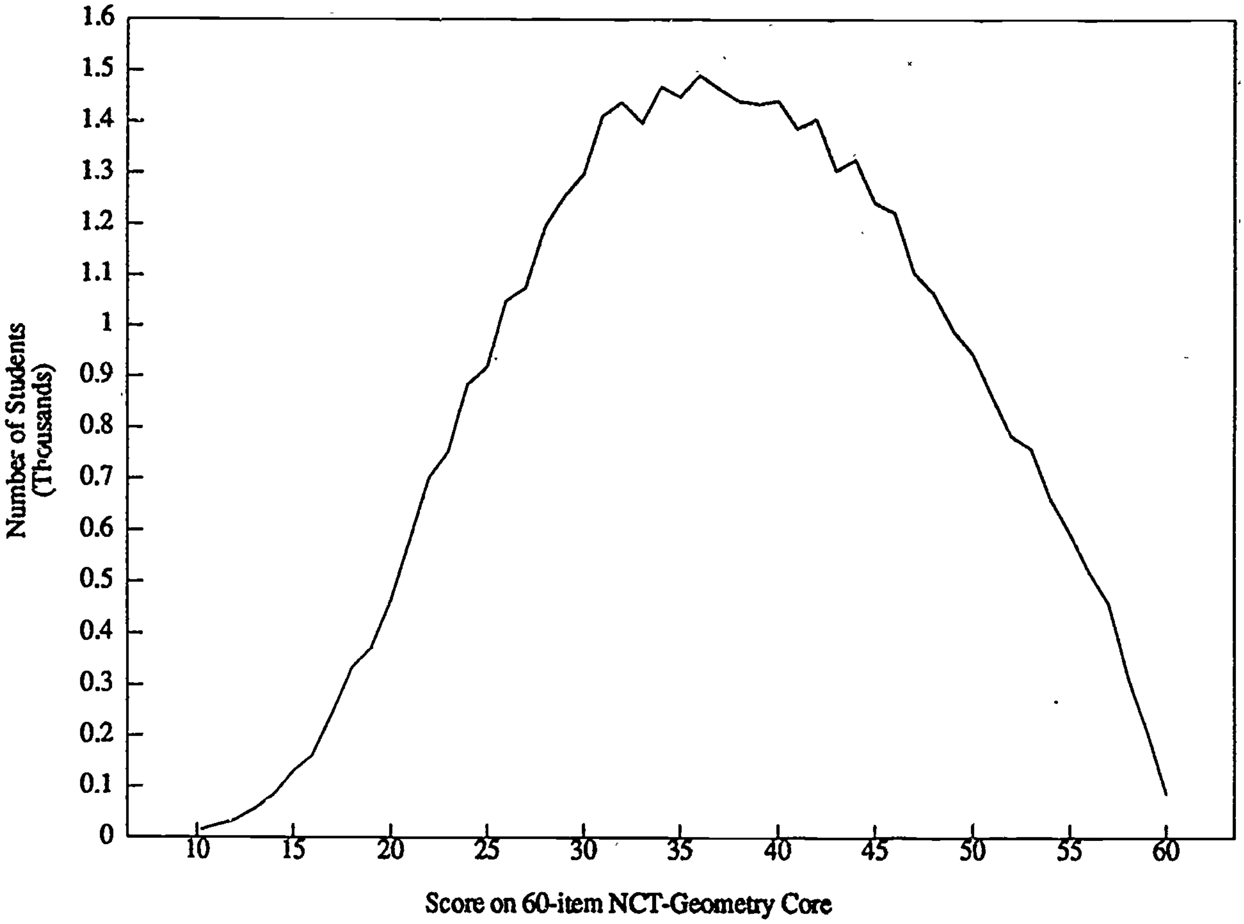

Figure 2. Frequency distribution of scores on the 1989 North Carolina Test of Geometry—Cores 1–8 (N = 43,325).

For practical purposes, the proper measure of reliability is the alternate form reliability. The calculation of this statistic requires that two or more equivalent forms of the test be developed. The older alternate form reliability 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 the 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 two or more tests be developed in exactly the same way, but independent of one another. This permits the reliability coefficient to reflect any random errors in selection made in the development of either of the test forms.

The alternate forms developed for the North Carolina Test of Geometry reflect this newer procedure. That is, each test form is developed from the domain of items in exactly the same manner. Any failure of the alternate form reliability to be 1.00 reflects the following:

- 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 Geometry ($r^2 = .83^2 = .69$) 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 coefficient is the statistic to use when correcting for attenuation.

Of special significance to the comparison of students scores across time is the equivalence of the core tests (Cores 1–8). While the initial equating process was successful, the eight cores were not exactly equivalent. Small random differences in item difficulty that generally occur between field testing and statewide testing were exhibited. These differences could be adjusted statistically by providing a separate set of norms for each test. A simple and efficient alternative is to redevelop the cores slightly so that even small differences are eliminated. With this technique, a single norms table can be used with all eight core tests.

The eight North Carolina Test of Geometry cores were administered in May 1989. The item difficulties for all 480 items based on large random samples (approximately 5,400 students responded to each item on the eight forms) were available to the test developer. To accomplish the final equating, items were exchanged between core tests to yield total scores that were equivalent for all eight tests. A total of 24 items were exchanged, while maintaining all of the decisions concerning the design of the test that had been made previously (see Table 2).

The adjustments to the core tests assure continuity of the norms table for future years while providing different test items each year. The typical administration procedure for the NCT–Geometry will be to administer four of the eight forms of the test per year (see Table 1). The different test items prevent the loss of confidentiality, and therefore validity, that occurs with the continued use of the same items. Students scores have a common reference point from 1989 onward, barring changes in the definition of the basic instructional program.

Curriculum Assessment

The North Carolina Test of Geometry is composed of items representing the 55 basic objectives and five items representing the ten variable objectives (as discussed earlier—see Table 1). All 60 items contribute to the individual student scores and to curricular assessment. Items representing a variable objective were answered by half of the students.

At the classroom level, 240 items are answered during each test administration by an average of six students. This procedure provides a database of approximately four items per objective across six students. From this database of information, evidence 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 240 items are answered by a much larger number of students: over 11,000 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 ten variable objectives (five items per form, for a total of 20 items) 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.

Content of The Test

The North Carolina Test of Geometry 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 Geometry tests were designed to achieve an even assessment across all objectives considered basic to instruction in Geometry; in short, each basic objective was 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 number of objectives that will be developed for it. Thus, an underlying system of weights exists for the curricular goals.

Appendix A lists each goal and objective and the numerical item representation for each objective as it appeared on the 1989 North Carolina Test of Geometry (Cores 1–8). In addition, the proportion of teachers rating each objective as basic to instruction in the Geometry curriculum is listed.

Tables 4 through 11 list the difficulty level for all items on the 1989 North Carolina Test of Geometry (Cores 1–8) in terms of p-values (proportion of all students answering the item correctly).

Table 4
Item Difficulty by Item Number for the 1989 NCT–Geometry—Core 1

Item #	P-value	Item #	P-value	Item #	P-value
1	.89	21	.59	41	.58
2	.89	22	.71	42	.42
3	.82	23	.55	43	.53
4	.82	24	.89	44	.35
5	.47	25	.57	45	.77
6	.62	26	.74	46	.36
7	.36	27	.67	47	.61
8	.78	28	.86	48	.79
9	.87	29	.42	49	.55
10	.60	30	.51	50	.36
11	.36	31	.81	51	.37
12	.89	32	.63	52	.66
13	.29	33	.82	53	.44
14	.95	34	.74	54	.42
15	.80	35	.56	55	.12
16	.87	36	.89	56	.73
17	.90	37	.91	57	.46
18	.54	38	.71	58	.36
19	.73	39	.45	59	.56
20	.52	40	.40	60	.45

Table 5
Item Difficulty by Item Number for the 1989 NCT-Geometry—Core 2

Item #	P-value	Item #	P-value	Item #	P-value
1	.59	21	.44	41	.55
2	.90	22	.75	42	.50
3	.27	23	.58	43	.73
4	.76	24	.69	44	.44
5	.65	25	.40	45	.55
6	.52	26	.53	46	.69
7	.44	27	.89	47	.80
8	.74	28	.86	48	.86
9	.76	29	.43	49	.52
10	.62	30	.42	50	.69
11	.52	31	.45	51	.76
12	.82	32	.95	52	.94
13	.94	33	.81	53	.50
14	.74	34	.61	54	.60
15	.72	35	.42	55	.78
16	.85	36	.47	56	.24
17	.71	37	.70	57	.62
18	.73	38	.75	58	.71
19	.93	39	.53	59	.47
20	.70	40	.53	60	.42

Table 6
Item Difficulty by Item Number for the 1989 NCT–Geometry—Core 3

Item #	P-value	Item #	P-value	Item #	P-value
1	.64	21	.75	41	.51
2	.85	22	.55	42	50
3	.77	23	.73	43	.66
4	.81	24	.85	44	.54
5	.45	25	.61	45	.50
6	.62	26	.80	46	.44
7	.79	27	.59	47	.72
8	.61	28	.59	48	.78
9	.80	29	.45	49	.26
10	.51	30	.65	50	.53
11	.36	31	.67	51	.49
12	.59	32	.97	52	.71
13	.86	33	.66	53	.40
14	.88	34	.57	54	.52
15	.89	35	.66	55	.55
16	.88	36	.38	56	.20
17	.97	37	.59	57	.71
18	.73	38	.55	58	.47
19	.48	39	.46	59	.86
20	.57	40	.60	60	.37

Table 7
Item Difficulty by Item Number for the 1989 NCT-Geometry—Core 4

Item #	P-value	Item #	P-value	Item #	P-value
1	.74	21	.73	41	.42
2	.36	22	.88	42	.53
3	.26	23	.54	43	[illegible]
4	.90	24	.91	44	.44
5	.84	25	.76	45	.77
6	.52	26	.85	46	.50
7	.70	27	.92	47	.83
8	.93	28	.65	48	.50
9	.58	29	.68	49	.84
10	.63	30	.43	50	.72
11	.32	31	.44	51	.28
12	.85	32	.81	52	.95
13	.77	33	.38	53	.47
14	.70	34	.47	54	.57
15	.32	35	.53	55	.35
16	.61	36	.33	56	.41
17	.78	37	.79	57	.55
18	.62	38	.86	58	.73
19	.56	39	.64	59	.52
20	.68	40	.78	60	.34

Table 8
Item Difficulty by Item Number for the 1989 NCT–Geometry—Core 5

Item #	P-value	Item #	P-value	Item #	P-value
1	.52	21	.46	41	.42
2	.69	22	.44	42	.58
3	.59	23	.49	43	.39
4	.85	24	.73	44	.52
5	.63	25	.73	45	.45
6	.92	26	.56	46	.65
7	.89	27	.68	47	.90
8	.78	28	.62	48	.42
9	.89	29	.61	49	.59
10	.53	30	.50	50	.81
11	.51	31	.42	51	.39
12	.85	32	.68	52	.74
13	.77	33	.82	53	.39
14	.75	34	.61	54	.51
15	.53	35	.89	55	.39
16	.75	36	.51	56	.32
17	.43	37	.78	57	.51
18	.75	38	.55	58	.38
19	.74	39	.51	59	.48
20	.91	40	.39	60	.14

Table 9
Item Difficulty by Item Number for the 1989 NCT–Geometry—Core 6

Item #	P-value	Item #	P-value	Item #	P-value
1	.51	21	.76	41	.51
2	.52	22	.60	42	.58
3	.48	23	.66	43	.53
4	.77	24	.64	44	.44
5	.34	25	.64	45	.63
6	.74	26	.62	46	.31
7	.78	27	.69	47	.61
8	.87	28	.59	48	.73
9	.89	29	.57	49	.81
10	.63	30	.65	50	.65
11	.78	31	.52	51	.50
12	.57	32	.89	52	.45
13	.92	33	.81	53	.68
14	.80	34	.70	54	.70
15	.50	35	.93	55	.67
16	.52	36	.52	56	.32
17	.53	37	.86	57	.52
18	.78	38	.41	58	.88
19	.54	39	.47	59	.50
20	.67	40	.54	60	.46

Table 10
Item Difficulty by Item Number for the 1989 NCT–Geometry—Core 7

Item #	P-value	Item #	P-value	Item #	P-value
1	.81	21	.76	41	.58
2	.81	22	.71	42	.83
3	.82	23	.37	43	.43
4	.53	24	.79	44	.46
5	.76	25	.79	45	.50
6	.91	26	.87	46	.66
7	.75	27	.57	47	.31
8	.34	28	.65	48	.55
9	.91	29	.43	49	.93
10	.51	30	.81	50	.75
11	.59	31	.46	51	.35
12	.52	32	.55	52	.41
13	.90	33	.65	53	.51
14	.77	34	.69	54	.53
15	.60	35	.39	55	.60
16	.70	36	.55	56	.50
17	.90	37	.39	57	.84
18	.71	38	.44	58	.63
19	.61	39	.46	59	.48
20	.73	40	.55	60	.43

Table 11
Item Difficulty by Item Number for the 1989 NCT–Geometry—Core 8

Item #	P-value	Item #	P-value	Item #	P-value
1	.54	21	.71	41	.55
2	.49	22	.81	42	.75
3	.88	23	.56	43	.51
4	.89	24	.84	44	.51
5	.82	25	.49	45	.37
6	.33	26	.74	46	.41
7	.40	27	.47	47	.69
8	.87	28	.67	48	.51
9	.91	29	.36	49	.53
10	.61	30	.88	50	.50
11	.48	31	.43	51	.55
12	.85	32	.94	52	.52
13	.90	33	.66	53	.83
14	.92	34	.85	54	.54
15	.90	35	.50	55	.53
16	.94	36	.45	56	.58
17	.89	37	.64	57	.85
18	.70	38	.53	58	.45
19	.83	39	.53	59	.63
20	.82	40	.58	60	.28

Test Norms

Students who answer all 60 of the Geometry items correctly could be assumed to be excellent Geometry students. If everyone answered all of the items correctly, however, a different interpretation would be placed on the scores. At some point, scores must have a reference group grounded in the experience of all students. In some respect, everything is good or bad by comparison. Norms tables provide this 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, Geometry students who took the North Carolina Test of Geometry in May 1989). 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 score 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 most useful statistic.

The norms table for student scores on the North Carolina Test of Geometry is given in Table 12. These scores set a baseline of comparison for present and future achievement in Geometry. Thus, a student score in 1990, 1991, and future years, can be referenced to the scores of all 1989 Geometry students in North Carolina.

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

Table 12
Noms for Student Scores on the North Carolina Test of Geometry

Raw Score	Percentile	Standard Score[a]
60	99	71.8
59	99	70.9
58	99	69.9
57	98	68.9
56	97	68.0
55	96	67.0
54	94	66.0
53	93	65.0
52	91	64.1
51	89	63.1
50	87	62.1
49	85	61.2
48	82	60.2
47	80	59.2
46	77	58.3
45	74	57.3
44	71	56.3
43	68	55.3
42	65	54.4
41	62	53.4
40	59	52.4
39	55	51.5
38	52	50.5
37	49	49.5
36	45	48.5
35	42	47.6
34	38	46.6
33	35	45.6
32	32	44.7
31	28	43.7
30	25	42.7
29	22	41.7
28	20	40.8
27	17	39.8
26	15	38.8
25	12	37.9
24	10	36.9
23	8	35.9
22	7	35.0
21	5	34.0
20	4	33.0
19	3	32.0
Less Than 19	2	

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

Appendices

Appendix A
Test Content—Item Representation by Goal and Objective

Goal/Obj	Description	No. Items 1989	% of Teachers Reporting as Basic[a]
Goal 1	The learner will state the characteristics of sets of points.		
1.3	Identify and name unions and intersections of sets of points.	8	91.08
1.5	Find the length of a segment.	8	99.39
1.7	Identify the midpoint of a given segment.	8	99.09
1.9	Determine when two angles are congruent.	8	99.70
1.10	Identify interiors and exteriors of geometric figures.	8	96.93
1.11	Identify the bisector of an angle.	8	99.39
Goal 2	The learner will use the structural properties of the real number.		
2.1	State and use the properties of equality.	8	96.04
2.2	State and use the properties of inequality.	8	86.28
Goal 3	The learner will develop geometric proofs.		
3.1	Translate a geometric statement into an "If-Then Statement."	8	94.21
3.2	State the converse of a conditional statement.	8	92.38
3.3	State the hypothesis and conclusion for a conditional statement.	8	94.80
3.4	Use the process of deductive reasoning in mathematical and non-mathematical situations.	8	84.31

[a]Percentage of North Carolina Geometry teachers rating the objective as basic to instruction in Geometry.

Goal/Obj	Description	No. Items 1989	% of Teachers Reporting as Basic[a]
Goal 4	The learner will use some of the properties of angles and lines to develop proofs and solve exercises.		
4.2	Classify an angle.	8	100.00
4.3	Identify adjacent and vertical angles.	8	99.69
4.4	Determine the complement and supplement of a given angle.	8	99.69
4.5	Apply the Angle Addition Postulate.	8	96.31
4.6	Apply the Segment Addition Postulate. (Definition of Betweenness)	8	92.28
Goal 5	The learner will recognize perpendicular lines and planes and use this information to complete proofs and exercises.		
5.1	Apply definitions of perpendicular lines and planes.	8	98.78
Goal 6	The learner will recognize parallel lines and planes and use this knowledge to complete proofs and exercises.		
6.1	Identify parallel lines and planes, and skew lines.	8	99.70
6.2	Identify corresponding angles and alternate interior angles which are formed when two parallel lines are cut by a transversal.	8	99.09
6.3	State conditions under which lines are parallel.	8	98.17
6.4	State which angles are congruent when two parallel lines are cut by a transversal.	8	98.78
6.5	Identify which angles are supplementary when two parallel lines are cut by a transversal.	8	97.57

[a]Percentage of North Carolina Geometry teachers rating the objective as basic to instruction in Geometry.

Goal/Obj	Description	No. Items 1989	% of Teachers Reporting as Basic[a]
Goal 7	The learner will identify polygons and complete proofs and exercises related to them.		
7.1	Classify a triangle according to its sides.	8	99.39
7.2	Classify a triangle according to its angles.	8	99.39
7.3	Classify a polygon according to the number of its sides or angles.	8	98.47
7.4	Classify a convex polygon according to the measure of its angles.	4	74.77
7.5	Apply the fact that the sum of the measures of the angles of a triangle is 180.	8	99.69
7.6	Find the measures of the exterior angles of a triangle.	8	96.32
7.7	Find the measures of the interior and exterior angles of a convex polygon.	8	90.21
7.8	Apply the characteristics of various quadrilaterals.	8	92.00
Goal 8	The learner will identify congruent triangles and complete proofs and exercises related to them.		
8.1	List the corresponding parts of two congruent triangles.	8	99.39
8.2	Use various postulates and theorems to prove two triangles are congruent and their corresponding parts are congruent.	8	96.95
8.3	Identify the altitudes and medians of triangles.	8	96.94
8.4	Apply the theorem about the segment joining the midpoints of two sides of a triangle.	8	87.16
8.5	Apply the theorem about the intersection of the medians of a triangle.	4	57.19

[a]Percentage of North Carolina Geometry teachers rating the objective as basic to instruction in Geometry.

Goal/Obj	Description	No. Items 1989	% of Teachers Reporting as Basic[a]
Goal 9	The learner will demonstrate when two polygons are similar and develop proofs and solve exercises related to them.		
9.1	Identify regular polygons and determine the measures of the angles.	8	98.78
9.2	Solve a proportion.	8	99.03
9.3	Use proportions to solve geometric problems.	8	97.08
9.4	Find the geometric mean of two numbers.	8	93.20
9.5	Determine whether or not two polygons are similar.	8	94.16
9.6	Prove two triangles are similar.	8	93.19
9.7	Apply properties of similar triangles to find corresponding proportional sides.	8	94.89
9.8	Apply theorems which involve dividing segments proportionally.	8	79.56
Goal 10	The learner will state some of the characteristics of a right triangle and solve exercises related to them.		
10.1	State two relationships that exist in a right triangle.	8	95.57
10.2	Use the Pythagorean Theorem and its converse to find the lengths of the sides of a right triangle or a quadrilateral.	8	99.51
10.3	Use the relationships that exist in special right triangles to solve problems.	8	95.63
Goal 11	The learner will list some characteristics of a circle and develop proofs and solve exercises related to them.		
11.1	Use the definitions of a circle and the lines and segments related to it.	8	98.29
11.2	Recognize polygons inscribed in or circumscribed about a circle.	8	91.50

[a]Percentage of North Carolina Geometry teachers rating the objective as basic to instruction in Geometry.

Goal/Obj	Description	No. Items 1989	% of Teachers Reporting as Basic[a]
11.3	Apply the properties involving arcs and angles of circles.	8	95.87
11.4	Apply the theorems about the chords of a circle.	8	94.17
11.5	Apply the theorems that relate to the tangents, secants, and radii of a circle.	8	90.29
Goal 12	The learner will find the perimeter, area, and volume of geometric figures.		
12.1	Find the perimeter of a geometric figure.	8	99.03
12.2	Compute the area of a triangle, parallelogram, trapezoid, and rectangle.	8	97.82
12.3	Find the ratio of both the areas and the perimeters of similar triangles.	4	66.91
12.4	Compute the apothem, radius, and area of special regular polygons.	4	72.57
12.6	Compute arc lengths and the areas of sectors of a circle.	4	74.27
12.7	Identify and describe space figures.	4	54.00
12.8	Compute the lateral area, total area, and volume of a right prism or pyramid.	4	57.87
12.9	Compute the lateral area and volume of a right circular cylinder or cone.	4	57.42
Goal 14	The learner will investigate some of the properties of coordinate geometry.		
14.1	Write the coordinates for a point in the coordinate plane.	8	78.10
14.2	Write equations for vertical and horizontal lines in the coordinate plane.	4	71.53
14.3	Use the distance formula to solve problems.	8	75.37
14.4	Use the midpoint formula to find the coordinates of the midpoint or endpoint of a segment.	8	75.43
14.7	Write an equation for a line which is parallel or perpendicular to a given line.	4	56.45

[a]Percentage of North Carolina Geometry teachers rating the objective as basic to instruction in Geometry.

Appendix B
Goals and Objectives Rejected for Use

Goal/Obj	Description	% of Teachers Reporting as Basic[a]
Goal 1	The learner will state the characteristics of sets of points.	
1.1	Identify and name sets of points, such as line, ray, segment, and plane.	99.70
1.2	Draw representations of points, lines, and planes.	99.39
1.4	Find the coordinate of a point on a line.	97.87
1.6	Identify congruent segments.	99.70
1.8	Use a protractor to find the measure of an angle.	87.20
Goal 3	The learner will develop geometric proofs.	
3.5	Write a proof using the two-column format.	96.32
3.6	Write an indirect proof.	59.20
Goal 4	The learner will use some of the properties of angles and lines to develop proofs and solve exercises.	
4.1	Use three letters, a number, or a single letter to name an angle.	100.00
4.7	Recognize congruent angles.	99.08
Goal 10	The learner will state some of the characteristics of a right triangle and solve exercises related to them.	
10.4	Use a table and/or calculator to apply the definitions of sine, cosine, and tangent to solve right triangles.	76.94

[a]Percentage of North Carolina Geometry teachers rating the objective as basic to instruction in Geometry.

Goal/Obj	Description	% of Teachers Reporting as Basic[a]
Goal 12	The learner will find the perimeter, area, and volume of geometric figures.	
12.5	Compute the circumference and area of a circle.	96.61
Goal 13	The learner will complete a geometric construction and describe the locus of a point or points.	
13.1	Construct a segment congruent to a given segment.	74.94
13.2	Construct an angle congruent to a given angle.	73.61
13.3	Construct the bisector of an angle.	74.33
13.4	Construct a line perpendicular to a line through a point on the line.	73.61
13.5	Construct a line perpendicular to a line through a point not on the line.	73.54
13.6	Construct the perpendicular bisector of a segment.	72.64
13.7	Construct a line parallel to a line through a given point.	68.93
13.8	Construct the tangents to a circle from a point outside the circle.	48.43
13.9	Circumscribe a circle about a triangle.	49.27
13.10	Inscribe a circle inside a triangle.	48.18
13.11	Divide a segment into a given number of congruent segments.	45.32
13.12	Given three segments, construct a fourth segment such that the lengths of the four segments are proportional.	26.65
13.13	Construct a segment whose length is the geometric mean between the lengths of two given segments.	24.02
13.14	Construct quadrilaterals which meet certain criteria.	27.21
13.15	Construct a circle through three non-collinear points.	23.04

[a]Percentage of North Carolina Geometry teachers rating the objective as basic to instruction in Geometry.

Goal/Obj	Description	% of Teachers Reporting as Basic[a]
Goal 14	The learner will investigate some of the properties of coordinate geometry.	
14.5	Find the slope of the line given two points on the line.	65.94
14.6	Find the slope and y–intercept of a line.	63.26
14.8	Write the equation and draw the graph of line when given either two points on the line, one point and the slope of the line, or the slope and y–intercept of the line.	55.96
14.9	Use coordinate geometry to prove some of the properties of polygons.	25.37
14.10	Write an equation of a circle given its center and radius length.	41.99
14.11	Find the center and radius length of a circle given an equation.	41.50

[a]Percentage of North Carolina Geometry teachers rating the objective as basic to instruction in Geometry.

Appendix 16

END

U.S. Dept. of Education

Office of Education
Research and
Improvement (OERI)

ERIC

Date Filmed

March 21,1991