DOCUMENT RESUME

ED 436 567 TM 030 429

TITLE	The High School Competency Test. October 1993 Technical Report.
INSTITUTION	Florida State Dept. of Education, Tallahassee. Assessment, Testing, and Evaluation Section.
PUB DATE	1993-00-00
NOTE	35p.
PUB TYPE	Reports - Research (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; Achievement Tests; Communication (Thought Transfer); *Competence; High School Students; High Schools; Mathematics Tests; *Psychometrics; Reading Tests; *School Districts; Schools; State Programs; *Test Results; Testing Programs
IDENTIFIERS	*Florida Statewide Assessment Program

ABSTRACT

Florida's High School Competency Test (HSCT) is a multiple-choice test of the application of basic academic skills to everyday life situations. The test is divided into mathematics and communications sections, with the communications section including skills in reading and writing. Passing the HSCT is a high school graduation requirement. This technical report presents background information about the HSCT and statistical data from the October 1993 administration to 11th graders. Information is provided to facilitate an evaluation of the psychometric adequacy of those results. The descriptive statistics and item discrimination indices are reported with a discussion of the validity and reliability of HSCT scores. Scoring and equating scores are discussed. Appendixes contain a summary of the skills within the HSCT and a chart of frequency and percentage distributions for the HSCT scores. (Contains 10 tables, 1 figure, and 10 references.) (SLD)

The High School Competency Test

October 1993 Technical Report

Florida Statewide Assessment Program
Student Assessment Services Section
Bureau of School Improvement and Instruction
Division of Public Schools
Department of Education
Florida Education Center, Suite 701
Tallahassee, Florida 32399-0400

Table of Contents

	Page
List of Tables	iii
List of Figures	iv
Description of the Florida High School Competency Test	1
Background	1
Item Specifications and Item Development	2
Preventing Test Item Bias	5
Construction of Test Forms	5
HSCT Scoring and Equating Scores	5
Test Booklet Format	6
Administration Procedures	6
Psychometric Characteristics of the HSCT Scores	7
Descriptive Statistics	7
Validity	9
Reliability	11
KR-20 Reliability Estimates	11
Brennan-Kane Index of Dependability	11
Threshold Loss Agreement Index	12
Reliability Summary	14
Item Difficulties	14
Discriminating Between High and Low Achievers	15
Student Performance on the HSCT	17
Summary of Item and Skill Statistics for the HSCT	18
References	26
Appendices	27
Appendix 1 Summary of the Skills Within the HSCT	28
Appendix 2 Frequency and Percentage Distributions for the HSCT Scores	29

List of Tables

Table		Page
1	Descriptive Statistics for the High School Competency Test by Section	8
2	Brennan-Kane Indices of Dependability for Classification of Achievement on the HSCT Skills	13
3	Summary of Item Difficulty Indices by Section	15
4	Discrimination Indices for the HSCT Items	17
5	Summary of Item and Skill Statistics for Scores on the HSCT Mathematics Section	19
6	Summary of Item and Skill Statistics for Scores on the HSCT Communications-Reading Section	22
7	Summary of Item and Skill Statistics for Scores on the HSCT Communications-Writing Section	25
8	Communications and Mathematics Skills Tested by the HSCT	28
9	Frequency Distributions and Percentages for the HSCT Communications Section	29
10	Frequency Distributions and Percentages for the HSCT Mathematics Section	30

List of Figures

Figure		Page
1	Item Specification Example	3

Description of the Florida High School Competency Test

Florida's High School Competency Test (HSCT) is a multiple-choice test of the application of basic academic skills to everyday life situations. The HSCT is divided into two sections: mathematics and communications. The communications section includes skills in both reading and writing. Passing the HSCT is a high school graduation requirement; Florida students are required to achieve passing scores on both sections of the HSCT to qualify for a regular high school diploma.

This technical report presents background information about the HSCT and statistical data from the October 1993 administration of the HSCT to eleventh-grade students. Information is provided to facilitate an evaluation of the psychometric adequacy of the October 1993 HSCT results. The descriptive statistics and item discrimination indices are reported with a discussion of the validity and reliability of HSCT scores. Additional results may be found in the State, District, and Regional Report of Statewide Assessment Results, October 1993, High School Competency Test, Grade Eleven (Department of Education, 1994).

Background

The HSCT is designed to measure Minimum Student Performance Standards (MSPS) in reading, writing, and mathematics. The MSPS define instructional outcomes in the form of broad standards and component minimum student performance skills. The skills tested on the October 1993 HSCT are a part of the MSPS adopted by the State Board of Education (SBE) in 1979 and are in effect for the school years 1985-1986 through 1993-1994. These skills are presented in Table 8 (p. 28). The MSPS were developed by the Florida Department of Education (DOE) with extensive involvement of state curriculum specialists, teachers, and district-level administrators. Citizens and students throughout the state were involved in the review and critique of the MSPS to ensure that the standards were clearly stated and that the standards were appropriate as minimal skills to be attained by all Florida students.

The 1976 Educational Accountability Act mandated the development of a test of MSPS that high school students would be required to pass in order to receive a regular high school diploma. The first version of this test was administered in 1977 following the adoption of the initial set of MSPS by the SBE. Beginning in 1978, the test was named the State Student Assessment Test-Part II (SSAT-II). (Students were not required to achieve passing scores on the State Student Assessment Test-Part I [SSAT-I], although they were required eventually to master all standards tested by the SSAT-I, as determined by local school districts, to qualify for a regular high school diploma.) The 1990 Florida Legislature revised Florida's assessment program, eliminating administration of the SSAT-I and changing the name of the SSAT-II to the HSCT.

School districts must provide remediation for students who do not pass one or both sections of the HSCT, and students are required to retake the sections of the test that they have not yet passed. In the spring of 1981, the time for students initially to take the SSAT-II was changed from the fall of grade eleven to the spring of grade ten to enable students to retake the

test four additional times, if necessary, before completion of the twelfth grade. In 1992, the time for taking the HSCT was returned to the fall of the student's eleventh-grade year.

Item Specifications and Item Development

Specifications for creating test items for each skill tested on the HSCT were developed by DOE with assistance from state university and district personnel. The item specifications describe what types of questions may be asked and the manner in which incorrect answers may be formulated. Each specification also provides one or more sample items. Personnel from each district in the state had the opportunity to critique these specifications before they were finalized. An example of an item specification is presented in Figure 1 (p. 3). The item specifications are available to all districts for use in developing instructional materials to assist students in achieving the skills.

Items for the HSCT were developed by state university personnel, district personnel, and private test developers under contracts or grants directed by DOE. Item developers adhered strictly to the item specifications to control the content and difficulty levels of the items. All items were reviewed extensively throughout the developmental process by Florida educators, curricular specialists, and measurement specialists to ensure that the items followed the specifications and to eliminate possible bias (see p. 5) in the items.

After new items had been reviewed extensively, they were administered to a sample of the population in either of two ways: in combination with operational test items on the HSCT test, or by themselves in separate test booklets. For the October 1993 administration, field-test items were included in the test booklets. The data from the field test and from the actual HSCT were analyzed using classical and Rasch item response theory (IRT) procedures. The primary use of the Rasch IRT values was to calibrate the items to a single difficulty scale. Therefore, all items in the bank had comparable Rasch difficulty values, although data for items administered on separate experimental forms were considered to be approximate. Besides Rasch difficulty values, the bank includes information concerning the percentage of students correctly answering each item, the item point-biserial discrimination indices, and the item response analyses (frequency with which each option is chosen). If all of the item statistics meet DOE standards, the item may be used on a future form of the HSCT.

The HSCT items are maintained by DOE staff in a bank that includes the item text, item responses, and any graphics. Data used for equating test scores are also maintained and updated in the bank after each administration.

Figure 1. Item Specification Example

GRADE 11

SECTION HSCT Mathematics

SKILL 141: Solve problems involving the perimeter or area of a rectangular region using metric or customary units.

STIMULUS ATTRIBUTES

Format

1. Description of a realistic situation requiring computation of the area or perimeter of a rectangular region.
2. Items may or may not require conversions. If conversions are required, they are within the system being used (no conversion between metric and customary units) and they may be to a larger or smaller unit.
3. A conversion may be required from a standard unit to a nonconventional unit that represents the packaging of an item; e.g., "If there are 20 one-foot-square tiles in one box, how many boxes of the tiles are necessary to cover the area?"
4. The dimensions of the rectangular region should be given in only one unit.
5. With customary units the only fraction allowed in computation is one-half. Items should not require multiplying or dividing fractions.
6. With metric units, computation should involve no more than three digits to the right or left of the decimal point.
7. The unit of the correct response should be explicitly stated in the stem if a conversion is required.
8. If an abbreviation or symbol for a unit of measurement is used in the graphic, the narrative must contain the word and the abbreviation or symbol. (The knowledge of the meaning of the abbreviation or symbol is not being tested.)
9. Items may use the terms "area" or "perimeter."
10. Permissible units are mm, cm, m, km, inches, feet, yards, and miles.
11. Graphics or pictures are appropriate (e.g., pictorial support depicting the rectangle and its dimensions).
12. An appropriate conversion table (including square measure) is required for items that require conversion.

RESPONSE ATTRIBUTES

1. Correct response.
2. Solutions to problems involving customary units to be whole numbers.
3. Distracters directed at:
   a. Incorrect operation;
   b. Failure to distinguish between perimeter and area;
   c. Finding only part of the perimeter (e.g., adding length and width);
   d. Not converting or converting when not necessary;
   e. Use of incorrect conversion factor (conversion table must be given);
   f. Common computational errors.

(figure continues)

(figure 1 continued)

Grade 11 Mathematics Skill 141

SAMPLE ITEMS

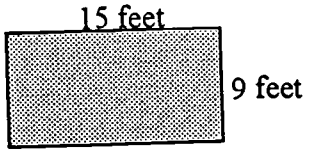
For questions 1 and 2, refer to the following table, as needed.

CONVERSION TABLE

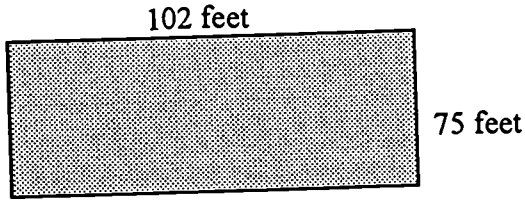
1 foot = 12 inches
1 yard = 3 feet
1 mile = 5280 feet

1 square yard = 144 square inches
1 square yard = 9 square feet

1. How many square yards of carpet are needed to cover a floor that is 9 feet long and 15 feet wide?



A. 15 square yards (1)
B. 45 square yards (3e)
C. 48 square yards (3b)
D. 135 square yards (3d)



2. Mr. Smith wants to fence in the lot shown in the figure above. How many feet of fence will he need?

F. 118 feet (3d)
G. 177 feet (3c)
H. 244 feet (3f)
I. 354 feet (1)

Preventing Test Item Bias

In addition to the strict guidelines in the item specifications, strict procedures were followed for reviewing the content of items to ensure that no item contained material that demeaned the gender, race, ethnicity, religion, or geographical region of any examinee. Steps were taken to ensure that no item had language that gave an advantage to or placed at a disadvantage members of any group.

The review of items for objectionable content and bias involved several steps. Each item-writing contractor was required to have draft items examined by educators qualified to review test items for gender, racial, ethnic, and other types of bias. After the items were critiqued and modified by the contractor, the DOE conducted an additional review of the items. This review was conducted by an independent group of educators from school districts throughout the state of Florida. The group was composed of teachers, principals, and curriculum experts who represented the different gender, racial, and ethnic groups. The item reviewers examined each item for objectionable content or potential bias. Items that might be objectionable to any student or biased for or against any group of students were modified or discarded.

Construction of Test Forms

The October 1993 HSCT was developed from the spring 1989 SSAT-I form in the same manner that new forms of the SSAT-II have been developed since 1977. To develop a new form of the HSCT, about 30% to 40% of the items from a previous form of the HSCT were replaced. Replacement items were selected to maintain the comparability of the content and the difficulty values of the original items. Both classical and IRT data were used in selecting items for the new forms of the test.

HSCT Scoring and Equating Scores

Each section of the HSCT contained a total of 75 operational items. Five items were used to measure each of the 15 mathematics skills and each of the 15 communications skills. The communications section included 11 reading skills and 4 writing skills. Students were assigned a mastery score or a nonmastery score for each of the skills tested. To master a skill, a student must answer correctly at least 4 out of the 5 items used to test the skill. Each student-level report indicated skills that the student achieved and the total number of skills achieved. Skill mastery rates were reported at the school, district, regional, and state levels.

Passing scores were based on the number of items that must be answered correctly on each section of each form of the test. The procedures used to develop items and construct test forms helped to ensure that a comparable passing standard was applied to students from year to year. The strict guidelines in the item specifications and the item replacement procedures assisted in maintaining similar difficulty levels of the items for each skill and the content equivalence of the test forms. Passing scores on the initial forms of the SSAT-II were based simply on the percentage of items answered correctly. Beginning in March 1983, Rasch IRT common-item equating procedures were used to establish a passing scale score to adjust for variation in the

difficulty of test forms. Use of this equating procedure resulted in small changes in the percentage of correctly answered items needed to pass different forms of the test.

For March 1983, a passing score of 700 was set to be equivalent to the passing score on the October 1978 test. In March 1984, the first SSAT-II form was administered that measured the 1985-86 through 1993-94 MSPS. A new passing standard was applied to this administration. The new standard was established by the State Board of Education following the recommendation of a committee consisting of Florida citizens and educators. Beginning in 1985, the scale score of 700 for each succeeding form of the HSCT has been set equivalent to the passing score of the March 1984 test using a common-item Rasch equating procedure. Item difficulty values and ability estimates continued to be placed on the 1978 base/scale through Rasch procedures.

After the October 1993 HSCT raw scores were equated to the 1984 SSAT-II scale, each student's equated ability logit, X, was converted to the October 1993 scale using the following linear transformations:

$$\text{Communications Scale Score} = 25(X - 2.090) + 700$$

$$\text{Mathematics Scale Score} = 25(X - 1.295) + 700$$

The constants 2.090 and 1.295 were the passing logits for the communications and mathematics sections, respectively. The 25 and 700 were constants selected to determine the scale characteristics.

To obtain a passing score of 700 on the October 1993 HSCT, a student needed to answer correctly 53 out of the 75 items on the mathematics section and 62 out of the 75 items on the communications section. Passing rates for each section of the HSCT were reported at the student, school, district, regional, and state levels.

Test Booklet Format

A single HSCT booklet was used, and answers were recorded on machine-scorable answer sheets. The mathematics and communications sections of the booklet were individually sealed to enhance test security.

Administration Procedures

The HSCT was administered by the DOE with the cooperation of a district coordinator in each of the state's 67 school districts. District coordinators of assessment designated and trained school coordinators, scheduled testing dates, administered makeup tests, maintained control of the materials, and accounted for the test booklets.

School coordinators were responsible for controlling all HSCT materials within their schools, training test administrators, and supervising test administration. In most instances, test administrators were either classroom teachers or guidance counselors.

The DOE provided manuals for all personnel involved in conducting the assessment. District and school coordinators received a common manual; test administrators received a separate manual. In addition, administration procedures for field-test items were provided to personnel administering those forms.

For the October 1993 HSCT administration, districts were requested to test all students, with the following exceptions: (1) students who have been enrolled in an English-speaking school less than two consecutive years; (2) students who have a temporary physical disability or temporary emotional problem; and (3) students with an active individual educational plan (IEP) who are classified according to SBE Rule 6A-6.331, FAC, as either Educable Mentally Handicapped, Trainable Mentally Handicapped, Hearing Impaired, Specific Learning Disabled, Emotionally Handicapped, Profoundly Handicapped, or Physically Impaired with Impaired Communication Abilities. Students in these exceptional categories who participate in 12 hours or less of special programming per week were encouraged to take the HSCT, although aggregated test results do not include the scores of these exceptional students. The test is untimed so that all students have the opportunity to answer all of the test items. The testing time was approximately 180 minutes.

Psychometric Characteristics of the HSCT Scores

Descriptive Statistics

Four hundred sixty-five schools and 102,402 eleventh-grade students participated in the October 1993 HSCT administration. The raw score measures of central tendency and dispersion statistics are presented in Table 1. All of the data except the number of schools and number of students are based upon a 5% random sample of examinees drawn from the population of eleventh graders who took the HSCT, excluding the students classified as exceptional students (as previously discussed).

Table 1

Descriptive Statistics for the High School Competency Test by Section

	Mathematics	Communications
Number of Schools	465	465
Number of Students	102,402	100,548
Number of Items	75	75
Raw Score Mean	60.19	68.74
Raw Score Q1	53	67
Raw Score Median	64	71
Raw Score Q3	70	73
Raw Score Mode	71	74
Raw Score Standard Deviation	12.06	7.83
Raw Score Range	7-75	4-75
Standard Error of Measurement	2.99	2.16

Note. Except for "Number of Schools" and "Number of Students," the statistical values were computed from a 5% random sample (N = 4991) of students drawn from the population of eleventh graders. These and subsequent analyses exclude exempted exceptional students.

Validity

The American Psychological Association's (APA) Standards for Educational and Psychological Testing (1985) addressed the concept of validity in testing:

> Validity is the most important consideration in test evaluation. The concept refers to the appropriateness, meaningfulness, and usefulness of the specific inferences made from test scores. Test validation is the process of accumulating evidence to support any particular inference. Validity, however, is a unitary concept. Although evidence may be accumulated in many ways, validity always refers to the degree to which that evidence supports the inferences that are made from the scores. The inferences regarding specific uses of a test are validated, not the test itself. (p. 9)

As an illustrative example, suppose the owners of a swimming pool decided to require bathers to pass a swimming test before permitting them to enter the deep end of the pool. The test consisted of swimming, unaided, for a distance of 20 yards. Test validation evidence was gathered by comparing pool accident records over a period of one year for pools that required the swimming test and pools that did not require the test. The results of the assessment confirmed that fewer accidents occurred at pools where bathers were required to pass the swimming test before entering the deeper water. Based on the results of the validation study, the pool owners would be justified in employing the swimming test to improve pool safety. On the other hand, if the pool owners decided to use the test results to determine who would be a potential competition swimmer, they probably would be making invalid interpretations of the information provided. The test had not been intended or validated for that purpose.

Generally, achievement tests are used for either (1) making predictions about students or (2) describing students' performance (Mehrens & Lehmann, 1991). The second purpose is most relevant for the Florida HSCT. The purpose of the HSCT is to document whether or not students have achieved the minimum student performance skills outlined in the Educational Accountability Act of 1990. To ensure that the test scores will allow interpretations appropriate for this purpose, it is necessary that the content of the test be carefully matched to the specified skills. Evidence of content-related validity is of primary importance for the HSCT. The APA (1985) stated:

> Content-related evidence of validity is a central concern during test development. . . . Expert professional judgment should play an integral part in developing the definition of what is to be measured, such as describing the universe of content, generating or selecting the content sample, and specifying the item format and scoring system. (p. 11)

In expanding on content-related validity evidence, some psychometricians have become interested in the concepts of curricular and instructional validity evidence, especially within the context of required minimum competency tests. Curricular validity evidence is related to the match between test content and curricular materials. Instructional validity evidence is related to the match between the skills measured by the test and the skills taught by the schools. Curricular

and instructional sources of validity evidence were addressed by the courts in the Debra P. v. Turlington court case. Florida's evidence in this case consisted of documentation from local school districts that the minimum student performance skills are being taught.

Unfortunately, as Brown (1976, p. 123) has noted, there are no well-established or satisfactory numerical indices to indicate the match of item content to objectives, test content to curricular materials, or skills measured to skills taught. Reliance must be placed on human judgment. Accordingly, the content validity of the Florida HSCT was determined by judging the extent to which test construction plans and procedures could reasonably be assumed to ensure validity. The general procedures used in test development were as follows:

1. The Minimum Student Performance Standards (MSPS) were developed with the involvement of instructional specialists.

2. The standards and skills were deemed acceptable. Educators and citizens were involved in this process.

3. Item specifications were written for each performance skill (see Figure 1).

4. Test items were written according to the guidelines provided by the item specifications.

5. The draft items were reviewed by instructional specialists and practicing teachers. Revisions were made when necessary.

6. The test items were subjected to final editing, as necessary.

The MSPS and their associated skills assessed in the October 1993 HSCT were selected from those adopted by the State Board of Education in 1979 (the "1985" MSPS). For both the 1977 and 1985 MSPS, the standards and associated skills were patterned after and developed under almost identical procedures used to develop the 1976 "milestone objectives." Drafts of goal statements were developed by the Florida Department of Education, then extensively reviewed and critiqued by district educators throughout the State, by citizens, and by students. Criteria used in these reviews included the clarity with which the goals were expressed and their appropriateness as minimal goals and subgoals to be attained by all Florida students. The final step in the process was adoption by the State Board of Education after a public hearing.

The development of item specifications was accomplished by contracting with the School Board of Dade County in 1976-1977 and Educational Testing Service in 1977. An example of an item specification is shown in Figure 1 (p. 3). Specifications were written for each performance skill to be included in the tests. Item specifications were critiqued by practicing Florida educators and made available to all Florida school districts. Since then, the specifications have been revised and updated to match the 1985 MSPS, based on suggestions from school districts, contractors, and Statewide Assessment Program staff members.

Steps four through six (see p. 10) were accomplished through test development contracts with testing agencies and Florida universities. In the past, some commercially produced items were leased for use on early forms of the SSAT-II, but the vast majority of HSCT items have been produced especially for use in Florida's tests.

All test-development contracts have proceeded in the following manner. Draft items were written according to the test specifications, then subjected to several reviews and pilot-tested. Next, items were field-tested to provide estimates of Florida students' performance on the items. After a final review, surviving items were placed in the item bank for future use. In all cases, items were reviewed by Florida educators, curricular specialists, and measurement specialists. All items were reviewed for cultural, ethnic, language, and gender bias.

Reliability

Reliability refers to the consistency or stability of test scores. If, with repeated administrations, a test can produce stable scores for each examinee, under the condition that the examinee is not affected by the examination process, the test can be called reliable. A test is also reliable if the items consistently measure the trait, ability, or construct being assessed. In this sense, it becomes quite clear why reliability is a necessary but not a sufficient condition for establishing test validity.

Two indices of the consistency of test scores, the KR-20 (Kuder-Richardson formula 20) estimate of internal consistency reliability (Kuder & Richardson, 1937) and the Brennan-Kane index of dependability (Brennan & Kane, 1977) were used to assess the reliability of the HSCT results.

An index designed specifically for measuring the reliability of criterion-referenced tests was also employed. A threshold loss agreement index (agreement coefficient, ρ_o) was used to estimate the reliability of mastery-nonmastery classifications (Berk, 1984).

KR-20 Reliability Estimates.

The KR-20 reliability coefficient can be considered an estimate of the correlation between the scores from the test and those from a parallel form of the test. The square root of the KR-20 coefficient is an estimate of the correlation between scores on the test and true scores on the trait measured by the test. The KR-20 reliability coefficients were 0.94 and 0.92 for the mathematics and communications skills section of the HSCT, respectively. These reliability estimates are comparable to the estimates reported for standardized norm-referenced achievement tests and provide strong evidence that the total scores from the assessment tests were highly consistent measures.

Brennan-Kane Index of Dependability.

Since a primary purpose of the HSCT is to provide information to help determine whether or not students have achieved specific skills, a second type of reliability estimate, more sensitive

to this demand of criterion-referenced tests, was computed. The Brennan-Kane (B-K) index of dependability was computed to estimate the consistency of test scores in classifying examinees as achievers or nonachievers of the minimum student performance skills. The B-K index, like the KR-20, ranges from zero to one with values approaching unity indicating greater dependability. Thus, the larger the index, the more appropriate the conclusion that any randomly selected examinee's test score is correct in classifying that examinee as a master or nonmaster of the goal in question.

Table 2 (p. 13) summarizes the B-K indices for the October 1993 skill scores. In general, these indices were satisfactory. For the mathematics section of the HSCT, the B-K indices ranged from 0.55 to 0.87 with 60% of the B-K indices greater than 0.60. For the communications section of the HSCT, the B-K indices ranged from 0.25 to 0.84 with 73% greater than 0.60.

The magnitudes of the B-K indices for the individual skills within the mathematics and communications sections indicated some variation in the confidence with which achievement status could be inferred. Since each skill involved a total of five items, these results were not unexpected. In general, the magnitude of the B-K indices for most mathematics and communications skills indicated that they were dependable measures of examinee achievement status.

Threshold Loss Agreement Index.

This threshold loss agreement index examines the proportion of students correctly classified as masters or nonmasters across classically parallel test forms. Subkoviak's (1988) method (agreement coefficient, ρ_o) generates an estimate of the proportion of students correctly classified as masters or nonmasters based on a single test administration. The agreement coefficient index ranges from 0.50 to 1.0, with values of one indicating greater overall consistency. The agreement coefficient was 0.90 for each section of the HSCT. Subkoviak noted that a value of 0.85 would be an acceptable value for making serious decisions (e.g., high school graduation) about students. Therefore, the value of 0.90 indicates a highly acceptable value for making mastery-nonmastery decisions about students for each section of the HSCT.

Table 2

Brennan-Kane Indices of Dependability for Classification of Achievement on the HSCT Skills

Skills	B-K Index
Communications Skills	
12. Determine the stated main idea.	.67
13. Find specific information in a selection.	.80
15. Identify the stated cause or effect.	.79
16. Follow written directions.	.84
19. Identify the implied main idea.	.49
21. Identify a conclusion or generalization.	.63
23. Distinguish between facts and opinions.	.74
25. Use pictures, maps, and signs.	.49
26. Use diagrams, tables, graphs, or schedules.	.25
27. Use indexes, tables of contents, or dictionaries.	.78
29. Identify appropriate sources of information.	.78
32. Include necessary information in messages.	.80
41. Include necessary information in letters.	.60
46. Complete common forms.	.68
47. Complete a money order or check and its stub.	.84
Mathematics Skills	
116. Solve problems involving averages.	.66
117. Solve problems involving whole numbers.	.79
120. Solve problems involving proper fractions.	.87
121. Solve problems involving decimals or percents.	.57
130. Determine equivalent amounts of money.	.78
131. Solve problems involving comparison shopping.	.58
132. Solve problems by finding simple interest.	.59
133. Solve purchase problems involving sales tax.	.55
134. Solve purchase problems involving discounts.	.62
140. Determine the elapsed time between two events.	.64
141. Solve problems involving the perimeter or area of a rectangle	.69
142. Solve problems involving length, width, or height.	.59
143. Solve problems involving capacity.	.70
144. Solve problems involving mass or weight.	.57
149. Determine relationships in graphs or tables.	.74

Note. These indices were computed from a 5% random sample of examinees.

Reliability Summary.

Reliability coefficients derived by different methods often depend on different assumptions about the nature or purpose of a test and will often produce different results. Brennan-Kane dependability indices and KR-20 internal consistency values will not necessarily be identical. The relationship between the magnitude of B-K values and KR-20 values is, in part, a function of the mastery or nonmastery cutoff score determined for a test or subtest. Furthermore, reliability or dependability indices are influenced by such basic factors as the length of test and the varied abilities of the group being tested. These factors affect the variability of the test scores and the item and test difficulty indices. In general, if longer tests are given to students with heterogeneous abilities with regard to the skills being assessed, reliability indices will be maximized. Since the HSCT assesses minimum competencies, the variability of abilities for the skills is highly restricted. This is clearly reflected in the difficulty indices associated with items, which are answered correctly by 90-99% of the examinees. The reliabilities demonstrated by the test scores are all the more impressive given these limitations of reliability estimates derived for norm-referenced tests, but applied to minimum competency tests.

Low reliability coefficients for a set of items measuring a single performance skill are not surprising and are not necessarily a cause for concern. Nevertheless, local district personnel should be cautious in interpreting test results with lower reliabilities. To make decisions about students' performance on individual skills, educators should consider collecting additional performance data, especially if the skill mastery status will be used to make decisions with serious consequences.

The overall consistency in the classification of students as masters or nonmasters was computed for each section of the HSCT. As with the KR-20 and the B-K index, the ρ_o index is sensitive to the test length and score variability. An additional factor affecting ρ_o is the distance of the cutoff score from the mean score. While these factors affect the ρ_o index, the values of ρ_o indicate that each section of the HSCT consistently classifies students as masters or nonmasters.

Item Difficulties

Item difficulty is computed as the percentage of examinees who answered a question correctly. Thus, item difficulty is an indirect index that ranges from 0 (no examinee answered the item correctly) to 100 (every examinee answered the item correctly). Table 3 (p. 15) summarizes the distribution of item difficulties on the HSCT. Most of the item difficulties are above 80 with few below 70. The communications items are generally easier than the mathematics items.

Table 3

Summary of Item Difficulty Indices by Section

Item Difficulty	Mathematics	Communications
90 to 99	17	54
80 to 89	26	18
70 to 79	20	2
60 to 69	7	1
50 to 59	4	0
40 to 49	0	0
30 to 39	1	0
20 to 29	0	0
10 to 19	0	0
0 to 9	0	0
Total	75	75

Note. These indices were computed from a 5% random sample of examinees.

Discriminating Between High and Low Achievers

In testing, discrimination generally refers to the relationship between the examinees' performance on an item and their performance on a section of the exam or on the entire exam. An item is said to be discriminating to the extent that performance on the item is predictive of overall performance. Thus, correlational techniques are usually used to assess the relationships between item and overall performance.

Three discrimination indices were used to assess the discrimination power of the HSCT: point biserial, corrected point biserial, and phi coefficients of correlation. The point biserial, corrected point biserial, and phi coefficients are all statistically equivalent to the Pearson product moment correlation between item performance and overall performance. The differences among these indices result from distinct ways of conceptualizing item performance and overall performance. Point biserial and corrected point biserial coefficients represent correlations between item performance and performance on a particular skill or section of the test. Item performance is treated as a dichotomous 0-1 variable (i.e., a variable that can have only two

values, such as zero if an item is answered incorrectly, and one if an item is answered correctly). Skill or section performance is a continuous variable (i.e., a variable that can take on any value across a continuum). The corrected point biserial correlation differs from the point biserial correlation only in that the item in question is not included in the particular overall performance score, thus avoiding a potential confounding between the correlation of item with overall performance. The phi coefficient represents correlations between item performance, again treated as a dichotomous 0-1 variable, and overall section performance, dichotomized into a 0-1 (fail or pass) variable. The discrimination indices for the HSCT are shown in Table 4 (p. 17).

Item discrimination is reflected by the magnitude of the point biserial, corrected point biserial, and phi coefficients, where larger values imply greater item discrimination. However, criteria for evaluating the magnitude of these indices have not been clearly specified for criterion-referenced tests. Criteria have been proposed for norm-referenced tests (Ebel, 1979, p. 227; Mehrens & Lehmann, 1991, p. 167). Whether these criteria can be applied to criterion-referenced data is uncertain, primarily because of the curvilinear relationship between item discrimination and item difficulty. As item difficulty approaches an extreme value (i.e., zero or one), item discrimination must approach zero since if all people get an item correct (item difficulty of one) or if all examinees miss an item (item difficulty of zero), the correlation between item and test performance must be zero. On the other hand, item discrimination is maximized as item difficulty approaches the middle value of 0.5. Thus, on criterion-referenced tests, where item difficulties tend to be in the 90-99% range, it is clear that the range of values for item discrimination indices is limited. A similar argument holds for the case in which the discrimination index assesses the relationship between standard performance and overall performance: when the distribution of masters to nonmasters is highly skewed, the sensitivity of the phi coefficient is reduced.

However, these item discrimination indices are useful if the intent is to locate negatively discriminating items or to contrast different types of discrimination indices computed for a single item. For example, when comparing point biserial with corrected point biserial correlations when the discrimination index involves a relatively small number of total items, as is the case when the index involves item to section correlations, one should be aware that the magnitude of change between the point biserial and corrected point biserial reflects the relative importance of the item in discriminating performance. As the difference between the point biserial and corrected point biserial becomes larger, item discriminating power (i.e., the ability of the item to differentiate between individuals of differing ability on a particular standard) increases. On the other hand, as the difference becomes smaller, item discriminating power decreases. If this difference is negative, it is clear that the item is counterproductive in discriminating between high and low standard performance, since examinees with less ability correctly answer that item more often than examinees with more ability. This pattern was not found for any of the items for which item to section discrimination indices were computed.

Table 4

Discrimination Indices for the HSCT Items

Range of Item Discrimination Indices	Point Biserial[a]		Corrected Point Biserial[b]		Corrected Point Biserial[a]		Phi[a]	
	Math	Comm.	Math	Comm.	Math	Comm.	Math	Comm.
.90 to .99								
.80 to .89	3						1	
.70 to .79	12	4			3		4	
.60 to .69	32	40			1		22	6
.50 to .59	20	18	14	2	5		18	33
.40 to .49	7	9	29	34	18	12	19	20
.30 to .39	1	3	18	22	25	29	7	13
.20 to .29		1	10	14	18	22	4	3
.10 to .19			4	3	5	12		
.00 to .09								
Total	75	75	75	75	75	75	75	75

Note. These indices were computed from a 5% random sample of examinees.
[a]Computed between each item and the skill.
[b]Computed between each item and the section.

Student Performance on the HSCT

Student performance on the HSCT is described in Tables 1, 5, 6, and 7 and at length in the statistical report titled State, District, and Regional Report of Statewide Assessment Results, October 1993, High School Competency Test, Grade Eleven (Department of Education, 1994). These results show that students performed at a high level on these tests, as would be expected for minimum competency tests. The mean percentage of items answered correctly on the mathematics section was 80%. The mean percentage of items answered correctly on the communications section was 92%.

Tables 9 and 10 report the frequency, percentage distributions, and cumulative percentage distributions for the raw and scale scores.

Summary of Item and Skill Statistics
for the HSCT

Summaries of item and skill statistics for the HSCT are given in Tables 5-7. The following example shows the statistics that are included for each skill and item and provides a key for their interpretation.

Characteristics of Skill		Characteristics of Items					
1. a. b. c.	ITEM NO.	P-VALUE	STD DEV	PHI I,SKL	PBIS I,SKL	C-PBIS I,SKL	C-PBIS I,SEC
G111 (4 ITEMS) COUNT = 5086	16	0.95	0.22	0.58	0.70	0.44	0.29
2. MEAN # ITEMS ACHIEVED = 3.80	17	0.95	0.21	0.63	0.73	0.49	0.28
3. STANDARD DEVIATION = 0.63	18	0.95	0.21	0.68	0.74	0.53	0.27
4. ATTAINMENT RATE = 0.95	19	0.95	0.23	0.65	0.75	0.50	0.32
5. R(PPM) WITH SECTION = 0.70							
6. BRENNAN-KANE INDEX = 0.89							

1a. SKILL LETTER and NUMBER
1b. NUMBER of items measuring the skill.
1c. COUNT denotes the number of students in the sample.
2. MEAN denotes the average number of items correctly answered on the skill for the students in the sample.
3. STANDARD DEVIATION denotes the variability, dispersion, or spread of the scores on the skill around the mean score.
4. ATTAINMENT RATE denotes the proportion of students in the sample achieving the skill. It may also be expressed as a percentage if multiplied by 100.
5. R(PPM) WITH SECTION denotes the Pearson product moment coefficient of correlation between the score on the skill and the overall section (mathematics or communications).
6. BRENNAN-KANE INDEX denotes the reliability or dependability of the mastery or nonmastery classification for the skill.
7. P-VALUE denotes the proportion of students in the sample who answered the item correctly. It may be expressed as a percentage if multiplied by 100.
8. STD DEV (STANDARD DEVIATION) denotes the variability, dispersion, or spread of the item score around the item p-value.
9. PHI I,SKL denotes the phi coefficient of correlation between the item, which is scored zero if incorrectly answered and one if correctly answered, and the skill, which is scored zero if not achieved and one if achieved.
10. PBIS I,SKL denotes the point biserial coefficient of correlation between the item, which is scored zero if answered incorrectly and one if answered correctly, and the skill total score.
11. C-PBIS I,SKL denotes the corrected point biserial coefficient of correlation between the item, which is scored zero if answered incorrectly and one if answered correctly, and an adjusted skill total score. The adjustment involves subtracting one from the total score if the item in question was answered correctly.
12. C-PBIS I,SEC denotes the corrected point biserial correlation between the item and the total score on the section.

Table 5

Summary of Item and Skill Statistics for Scores on the HSCT Mathematics Section

Characteristics of Skill		Characteristics of Items						
Mathematics		ITEM NO.	P-VALUE	STD DEV	PHI I,SKL	PBIS I,SKL	C-PBIS I,SKL	C-PBIS I,SEC
M116 (5 Items) Count	= 4991	21	0.89	0.31	0.45	0.56	0.29	0.36
Mean # Items Achieved	= 4.43	22	0.88	0.33	0.57	0.65	0.38	0.39
Standard Deviation	= 0.98	23	0.85	0.36	0.50	0.62	0.32	0.34
Attainment Rate	= 0.87	24	0.93	0.26	0.45	0.55	0.33	0.39
R(PPM) With Section	= 0.66	25	0.87	0.33	0.64	0.69	0.44	0.46
Brennan-Kane Index	= 0.66							
M117 (5 Items) Count	= 4991	26	0.94	0.23	0.59	0.65	0.42	0.40
Mean # Items Achieved	= 4.68	27	0.93	0.26	0.47	0.62	0.35	0.30
Standard Deviation	= 0.78	28	0.92	0.27	0.68	0.73	0.49	0.48
Attainment Rate	= 0.93	29	0.98	0.13	0.35	0.48	0.33	0.24
R(PPM) With Section	= 0.64	30	0.90	0.30	0.61	0.71	0.42	0.50
Brennan-Kane Index	= 0.79							
M120 (5 Items) Count	= 4991	31	0.83	0.38	0.70	0.79	0.67	0.51
Mean # Items Achieved	= 4.07	32	0.76	0.43	0.63	0.72	0.54	0.48
Standard Deviation	= 1.58	33	0.80	0.40	0.82	0.85	0.74	0.52
Attainment Rate	= 0.79	34	0.84	0.37	0.79	0.87	0.79	0.50
R(PPM) With Section	= 0.64	35	0.85	0.36	0.78	0.87	0.79	0.50
Brennan-Kane Index	= 0.87							
M121 (5 Items) Count	= 4991	36	0.82	0.38	0.56	0.61	0.33	0.46
Mean # Items Achieved	= 4.14	37	0.79	0.41	0.63	0.68	0.40	0.49
Standard Deviation	= 1.14	38	0.79	0.41	0.36	0.53	0.20	0.32
Attainment Rate	= 0.77	39	0.86	0.35	0.55	0.61	0.36	0.41
R(PPM) With Section	= 0.74	40	0.88	0.33	0.55	0.61	0.37	0.46
Brennan-Kane Index	= 0.57							
M130 (5 Items) Count	= 4991	01	0.96	0.20	0.37	0.49	0.17	0.22
Mean # Items Achieved	= 4.75	02	0.95	0.22	0.51	0.58	0.26	0.31
Standard Deviation	= 0.60	03	0.96	0.20	0.42	0.52	0.23	0.26
Attainment Rate	= 0.96	04	0.94	0.24	0.38	0.56	0.20	0.17
R(PPM) With Section	= 0.44	05	0.94	0.24	0.43	0.58	0.22	0.20
Brennan-Kane Index	= 0.78							

(table continues)

(Table 5 continued)

Characteristics of Skill		Characteristics of Items						
Mathematics		ITEM NO.	P-VALUE	STD DEV	PHI I,SKL	PBIS I,SKL	C-PBIS I,SKL	C-PBIS I,SEC
M131 (5 Items) Count	= 4991	06	0.67	0.47	0.68	0.74	0.46	0.56
Mean # Items Achieved	= 4.02	07	0.82	0.38	0.52	0.62	0.36	0.42
Standard Deviation	= 1.21	08	0.90	0.30	0.35	0.46	0.23	0.29
Attainment Rate	= 0.72	09	0.89	0.31	0.43	0.52	0.30	0.35
R(PPM) With Section	= 0.74	10	0.73	0.44	0.67	0.73	0.46	0.50
Brennan-Kane Index	= 0.58							
M132 (5 Items) Count	= 4991	51	0.72	0.45	0.61	0.65	0.38	0.43
Mean # Items Achieved	= 3.77	52	0.76	0.43	0.59	0.68	0.44	0.54
Standard Deviation	= 1.31	53	0.76	0.43	0.43	0.57	0.28	0.30
Attainment Rate	= 0.65	54	0.72	0.45	0.57	0.62	0.34	0.44
R(PPM) With Section	= 0.73	55	0.81	0.39	0.44	0.53	0.27	0.37
Brennan-Kane Index	= 0.59							
M133 (5 Items) Count	= 4991	16	0.87	0.33	0.36	0.46	0.21	0.25
Mean # Items Achieved	= 3.93	17	0.79	0.41	0.64	0.69	0.44	0.56
Standard Deviation	= 1.22	18	0.73	0.45	0.47	0.58	0.25	0.35
Attainment Rate	= 0.71	19	0.80	0.40	0.63	0.68	0.44	0.52
R(PPM) With Section	= 0.73	20	0.74	0.44	0.48	0.59	0.27	0.36
Brennan-Kane Index	= 0.55							
M134 (5 Items) Count	= 4991	41	0.85	0.36	0.46	0.57	0.35	0.42
Mean # Items Achieved	= 3.77	42	0.73	0.45	0.48	0.59	0.30	0.30
Standard Deviation	= 1.35	43	0.75	0.43	0.62	0.67	0.43	0.48
Attainment Rate	= 0.65	44	0.69	0.46	0.62	0.68	0.42	0.42
R(PPM) With Section	= 0.69	45	0.75	0.43	0.57	0.65	0.40	0.41
Brennan-Kane Index	= 0.62							
M140 (5 Items) Count	= 4991	46	0.78	0.42	0.67	0.74	0.49	0.44
Mean # Items Achieved	= 4.20	47	0.86	0.35	0.27	0.42	0.14	0.16
Standard Deviation	= 1.16	48	0.86	0.35	0.67	0.71	0.50	0.39
Attainment Rate	= 0.79	49	0.88	0.32	0.43	0.53	0.28	0.43
R(PPM) With Section	= 0.64	50	0.82	0.38	0.72	0.75	0.54	0.49
Brennan-Kane Index	= 0.64							

(table continues)

(Table 5 continued)

Characteristics of Skill		Characteristics of Items						
Mathematics		ITEM NO.	P-VALUE	STD DEV	PHI I,SKL	PBIS I,SKL	C-PBIS I,SKL	C-PBIS I,SEC
M141 (5 Items) Count	= 4991	56	0.75	0.43	0.40	0.63	0.38	0.49
Mean # Items Achieved	= 2.82	57	0.51	0.50	0.44	0.56	0.24	0.29
Standard Deviation	= 1.41	58	0.69	0.46	0.42	0.61	0.34	0.41
Attainment Rate	= 0.34	59	0.36	0.48	0.55	0.57	0.27	0.29
R(PPM) With Section	= 0.68	60	0.51	0.50	0.59	0.61	0.31	0.38
Brennan-Kane Index	= 0.69							
M142 (5 Items) Count	= 4991	61	0.92	0.27	0.27	0.42	0.23	0.33
Mean # Items Achieved	= 3.38	62	0.71	0.45	0.49	0.61	0.33	0.41
Standard Deviation	= 1.36	63	0.67	0.47	0.56	0.64	0.36	0.42
Attainment Rate	= 0.52	64	0.54	0.50	0.60	0.67	0.37	0.47
R(PPM) With Section	= 0.73	65	0.54	0.50	0.62	0.67	0.38	0.43
Brennan-Kane Index	= 0.59							
M143 (5 Items) Count	= 4991	66	0.68	0.46	0.61	0.68	0.44	0.49
Mean # Items Achieved	= 3.76	67	0.72	0.45	0.67	0.73	0.52	0.58
Standard Deviation	= 1.45	68	0.66	0.47	0.66	0.72	0.49	0.55
Attainment Rate	= 0.67	69	0.85	0.35	0.48	0.61	0.42	0.49
R(PPM) With Section	= 0.81	70	0.84	0.36	0.55	0.68	0.51	0.53
Brennan-Kane Index	= 0.70							
M144 (5 Items) Count	= 4991	71	0.83	0.37	0.51	0.60	0.37	0.48
Mean # Items Achieved	= 3.74	72	0.71	0.45	0.52	0.61	0.31	0.42
Standard Deviation	= 1.30	73	0.75	0.43	0.60	0.66	0.40	0.51
Attainment Rate	= 0.65	74	0.81	0.39	0.48	0.57	0.31	0.36
R(PPM) With Section	= 0.73	75	0.64	0.48	0.50	0.60	0.27	0.31
Brennan-Kane Index	= 0.57							
M149 (5 Items) Count	= 4991	11	0.83	0.38	0.27	0.76	0.14	0.47
Mean # Items Achieved	= 4.74	12	0.97	0.16	0.31	0.41	0.14	0.14
Standard Deviation	= 0.57	13	0.99	0.10	0.29	0.35	0.18	0.13
Attainment Rate	= 0.97	14	0.97	0.17	0.62	0.55	0.28	0.24
R(PPM) With Section	= 0.55	15	0.98	0.16	0.64	0.55	0.31	0.26
Brennan-Kane Index	= 0.74							

Table 6

Summary of Item and Skill Statistics for Scores on the HSCT Communications-Reading Section

Characteristics of Skill		Characteristics of Items						
Communications Reading		ITEM NO.	P-VALUE	STD DEV	PHI I,SKL	PBIS I,SKL	C-PBIS I,SKL	C-PBIS I,SEC
R12 (5 Items) Count	= 4991	04	0.90	0.30	0.52	0.61	0.34	0.41
Mean # Items Achieved	= 4.47	07	0.89	0.31	0.54	0.61	0.34	0.39
Standard Deviation	= 0.95	18	0.85	0.35	0.58	0.68	0.38	0.43
Attainment Rate	= 0.87	21	0.96	0.20	0.45	0.55	0.39	0.54
R(PPM) With Section	= 0.74	23	0.87	0.34	0.58	0.67	0.40	0.46
Brennan-Kane Index	= 0.67							
R13 (5 Items) Count	= 4991	01	0.97	0.17	0.43	0.53	0.27	0.35
Mean # Items Achieved	= 4.77	02	0.99	0.10	0.35	0.40	0.25	0.28
Standard Deviation	= 0.59	19	0.93	0.26	0.57	0.69	0.34	0.48
Attainment Rate	= 0.96	32	0.98	0.13	0.47	0.49	0.29	0.39
R(PPM) With Section	= 0.71	70	0.90	0.30	0.48	0.72	0.28	0.45
Brennan-Kane Index	= 0.80							
R15 (5 Items) Count	= 4991	05	0.97	0.18	0.40	0.53	0.33	0.40
Mean # Items Achieved	= 4.68	08	0.95	0.23	0.53	0.60	0.36	0.46
Standard Deviation	= 0.77	38	0.93	0.25	0.54	0.64	0.38	0.48
Attainment Rate	= 0.93	43	0.94	0.25	0.60	0.67	0.43	0.48
R(PPM) With Section	= 0.78	69	0.90	0.30	0.56	0.70	0.39	0.50
Brennan-Kane Index	= 0.79							
R16 (5 Items) Count	= 4991	20	0.98	0.14	0.43	0.52	0.32	0.37
Mean # Items Achieved	= 4.78	22	0.92	0.27	0.46	0.67	0.31	0.40
Standard Deviation	= 0.63	50	0.96	0.20	0.58	0.66	0.41	0.48
Attainment Rate	= 0.96	52	0.98	0.15	0.59	0.64	0.46	0.48
R(PPM) With Section	= 0.71	74	0.94	0.23	0.54	0.65	0.35	0.43
Brennan-Kane Index	= 0.84							
R19 (5 Items) Count	= 4991	06	0.97	0.17	0.38	0.48	0.32	0.46
Mean # Items Achieved	= 4.34	37	0.91	0.29	0.51	0.57	0.30	0.47
Standard Deviation	= 0.93	57	0.92	0.27	0.52	0.56	0.31	0.49
Attainment Rate	= 0.85	67	0.72	0.45	0.44	0.65	0.21	0.29
R(PPM) With Section	= 0.73	71	0.83	0.38	0.57	0.63	0.28	0.38
Brennan-Kane Index	= 0.49							

(table continues)

(Table 6 continued)

Characteristics of Skill		Characteristics of Items						
Communications Reading		ITEM NO.	P-VALUE	STD DEV	PHI I,SKL	PBIS I,SKL	C-PBIS I,SKL	C-PBIS I,SEC
R21 (5 Items) Count	= 4991	39	0.89	0.32	0.50	0.60	0.35	0.43
Mean # Items Achieved	= 4.31	56	0.89	0.32	0.56	0.62	0.38	0.47
Standard Deviation	= 1.07	59	0.93	0.25	0.47	0.57	0.37	0.46
Attainment Rate	= 0.83	68	0.78	0.41	0.56	0.68	0.36	0.39
R(PPM) With Section	= 0.75	75	0.82	0.39	0.62	0.69	0.41	0.47
Brennan-Kane Index	= 0.63							
R23 (5 Items) Count	= 4991	15	0.89	0.32	0.58	0.66	0.41	0.48
Mean # Items Achieved	= 4.48	41	0.91	0.28	0.63	0.69	0.49	0.47
Standard Deviation	= 1.01	42	0.92	0.27	0.61	0.68	0.49	0.49
Attainment Rate	= 0.87	61	0.92	0.27	0.56	0.64	0.43	0.44
R(PPM) With Section	= 0.73	65	0.84	0.37	0.57	0.69	0.41	0.41
Brennan-Kane Index	= 0.74							
R25 (5 items) Count	= 4991	09	0.97	0.17	0.28	0.35	0.15	0.24
Mean # Items Achieved	= 4.45	10	0.80	0.40	0.38	0.59	0.13	0.18
Standard Deviation	= 0.82	11	0.91	0.29	0.41	0.50	0.17	0.22
Attainment Rate	= 0.88	24	0.85	0.35	0.58	0.64	0.27	0.37
R(PPM) With Section	= 0.57	25	0.92	0.28	0.51	0.55	0.25	0.37
Brennan-Kane Index	= 0.49							
R26 (5 items) Count	= 4991	26	0.94	0.24	0.39	0.45	0.17	0.23
Mean # Items Achieved	= 4.32	27	0.98	0.14	0.29	0.36	0.20	0.31
Standard Deviation	= 0.81	28	0.96	0.19	0.36	0.44	0.22	0.35
Attainment Rate	= 0.86	49	0.84	0.37	0.60	0.60	0.17	0.32
R(PPM) With Section	= 0.57	51	0.61	0.49	0.41	0.71	0.16	0.25
Brennan-Kane Index	= 0.25							
R27 (5 Items) Count	= 4991	12	0.99	0.10	0.21	0.28	0.12	0.17
Mean # Items Achieved	= 4.75	34	0.89	0.32	0.51	0.73	0.30	0.38
Standard Deviation	= 0.61	47	0.97	0.17	0.47	0.54	0.29	0.34
Attainment Rate	= 0.95	48	0.98	0.14	0.48	0.50	0.30	0.32
R(PPM) With Section	= 0.62	58	0.92	0.27	0.57	0.69	0.32	0.41
Brennan-Kane Index	= 0.78							

(table continues)

(Table 6 continued)

Characteristics of Skill		Characteristics of Items						
Communications Reading		ITEM NO.	P-VALUE	STD DEV	PHI I,SKL	PBIS I,SKL	C-PBIS I,SKL	C-PBIS I,SEC
R29 (5 Items) Count	= 4991	03	0.98	0.12	0.30	0.38	0.20	0.27
Mean # Items Achieved	= 4.73	40	0.89	0.31	0.46	0.68	0.26	0.39
Standard Deviation	= 0.65	54	0.93	0.26	0.45	0.61	0.27	0.33
Attainment Rate	= 0.95	60	0.97	0.17	0.54	0.58	0.36	0.46
R(PPM) With Section	= 0.68	62	0.95	0.21	0.62	0.64	0.38	0.46
Brennan-Kane Index	= 0.78							

Table 7

Summary of Item and Skill Statistics for Scores on the HSCT Communications-Writing Section

Characteristics of Skill		Characteristics of Items						
Communications Writing		ITEM NO.	P-VALUE	STD DEV	PHI I,SKL	PBIS I,SKL	C-PBIS I,SKL	C-PBIS I,SEC
W32 (5 Items) Count	= 4991	35	0.95	0.22	0.55	0.62	0.37	0.42
Mean # Items Achieved	= 4.72	36	0.92	0.27	0.45	0.61	0.29	0.26
Standard Deviation	= 0.71	66	0.97	0.16	0.50	0.59	0.42	0.49
Attainment Rate	= 0.95	72	0.91	0.28	0.50	0.66	0.33	0.42
R(PPM) With Section	= 0.67	73	0.96	0.20	0.56	0.63	0.42	0.45
Brennan-Kane Index	= 0.80							
W41 (5 Items) Count	= 4991	16	0.83	0.38	0.47	0.62	0.25	0.28
Mean # Items Achieved	= 4.45	17	0.86	0.35	0.54	0.64	0.31	0.35
Standard Deviation	= 0.90	33	0.96	0.20	0.34	0.45	0.25	0.38
Attainment Rate	= 0.88	53	0.89	0.31	0.51	0.60	0.30	0.43
R(PPM) With Section	= 0.69	55	0.91	0.29	0.52	0.60	0.33	0.47
Brennan-Kane Index	= 0.60							
W46 (5 Items) Count	= 4991	44	0.99	0.11	0.38	0.42	0.25	0.31
Mean # Items Achieved	= 4.68	45	0.94	0.24	0.39	0.52	0.16	0.24
Standard Deviation	= 0.63	46	0.96	0.20	0.36	0.46	0.15	0.20
Attainment Rate	= 0.95	63	0.89	0.31	0.41	0.61	0.15	0.21
(R(PPM) With Section	= 0.49	64	0.91	0.29	0.43	0.60	0.16	0.25
Brennan-Kane Index	= 0.68							
W47 (5 items) Count	= 4991	13	0.98	0.15	0.39	0.50	0.24	0.29
Mean # items Achieved	= 4.82	14	0.97	0.18	0.34	0.49	0.16	0.17
Standard Deviation	= 0.51	29	0.97	0.18	0.37	0.54	0.22	0.35
Attainment Rate	= 0.97	30	0.98	0.15	0.52	0.55	0.29	0.31
R(PPM) With Section	= 0.56	31	0.93	0.25	0.47	0.67	0.25	0.35
Brennan-Kane Index	= 0.84							

References

American Psychological Association. (1985). Standards for educational and psychological testing. Washington, DC: Author.

Berk, R. A. (1984). A guide to criterion-referenced test construction. Baltimore, MD: John Hopkins University Press.

Brennan, R. L., & Kane, M. T. (1977). An index of dependability for mastery tests. Journal of Educational Measurement, 14, 277-289.

Brown, F. G. (1976). Principles of educational and psychological testing (2nd ed.). New York: Holt, Rinehart and Winston.

Ebel, R. L. (1979). Essentials of educational measurement (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. Psychometrika, 2, 151-160.

Legg, S. M. (1993). Final report: Score calibration for the High School Competency Test 1993. Tallahassee, FL: Florida Department of Education.

Mehrens, W. A., & Lehmann, I. J. (1991). Measurement and evaluation in education and psychology (3rd ed.). New York: Holt, Rinehart and Winston.

Student Assessment Services Section, Division of Public Schools, Florida Department of Education (1994). State, district, and regional report of statewide assessment results, October 1993, High School Competency Test, grade eleven. Tallahassee, FL: Florida Department of Education.

Subkoviak, M. J. (1988). A practitioner's guide to computation and interpretation of reliability indices for mastery tests. Journal of Educational Measurement, 25, 47-55.

Appendices

Appendix 1

Summary of the Skills Within the HSCT

Table 8

Communications and Mathematics Skills Tested by the HSCT

Skills
Communications
12. Determine the stated main idea.
13. Find specific information in a selection.
15. Identify the stated cause or effect.
16. Follow written directions.
19. Identify the implied main idea.
21. Identify a conclusion or generalization.
23. Distinguish between facts and opinions.
25. Use pictures, maps, or signs.
26. Use diagrams, tables, graphs, or schedules.
27. Use indexes, tables of contents, or dictionaries.
29. Identify appropriate sources of information.
32. Include necessary information in messages.
41. Include necessary information in letters.
46. Complete common forms.
47. Complete a money order or check and its stub.
Mathematics
116. Solve problems involving averages.
117. Solve problems involving whole numbers.
120. Solve problems involving proper fractions.
121. Solve problems involving decimals or percents.
130. Determine equivalent amounts of money.
131. Solve problems involving comparison shopping.
132. Solve problems by finding simple interest.
133. Solve purchase problems involving sales tax.
134. Solve purchase problems involving discounts.
140. Determine the elapsed time between two events.
141. Solve problems involving the perimeter or area of a rectangle.
142. Solve problems involving length, width, or height.
143. Solve problems involving capacity.
144. Solve problems involving mass or weight.
149. Determine relationships in graphs or tables.

Note. Five items were used to test each skill.

Appendix 2

Frequency and Percentage Distributions for the HSCT Scores

Table 9

Frequency Distributions and Percentages for the HSCT Communications Section

Raw Score	Scale Score	Freq[a]	Percent[b]	Cumulative Percent[c]	Raw Score	Scale Score	Freq	Percent	Cumulative Percent
75	797	463	9.3	100.0	37	655	4	0.1	1.4
74	774	700	14.0	90.7	36	654	7	0.1	1.4
73	756	694	13.9	76.7	35	652	8	0.2	1.2
72	745	510	10.2	62.8	34	651	5	0.1	1.1
71	737	448	9.0	52.6	33	649	3	0.1	1.0
70	731	392	7.9	43.6	32	648	4	0.1	0.9
69	726	285	5.7	35.7	31	646	6	0.1	0.8
68	721	219	4.4	30.0	30	644	1	0.0	0.7
67	717	156	3.1	25.6	29	643	4	0.1	0.7
66	713	147	2.9	22.5	28	641	4	0.1	0.6
65	710	129	2.6	19.6	27	639	1	0.0	0.5
64	707	94	1.9	17.0	26	637	4	0.1	0.5
63	704	98	2.0	15.1	25	636	2	0.0	0.4
62	702	86	1.7	13.1	24	634	3	0.1	0.4
61	699	69	1.4	11.4	23	632	1	0.0	0.3
60	697	66	1.3	10.0	22	630	2	0.0	0.3
59	694	40	0.8	8.7	21	628	5	0.1	0.3
58	692	37	0.7	7.9	20	626	0	0.0	0.2
57	690	29	0.6	7.2	19	624	2	0.0	0.2
56	688	32	0.6	6.6	18	622	2	0.0	0.1
55	686	25	0.5	6.0	17	620	1	0.0	0.1
54	684	31	0.6	5.4	16	618	2	0.0	0.1
53	682	18	0.4	4.8	15	616	0	0.0	0.0
52	680	18	0.4	4.5	14	613	0	0.0	0.0
51	679	12	0.2	4.1	13	611	0	0.0	0.0
50	677	28	0.6	3.9	12	608	0	0.0	0.0
49	675	12	0.2	3.3	11	605	0	0.0	0.0
48	673	12	0.2	3.1	10	602	0	0.0	0.0
47	672	7	0.1	2.8	9	599	0	0.0	0.0
46	670	10	0.2	2.7	8	595	0	0.0	0.0
45	668	10	0.2	2.5	7	591	0	0.0	0.0
44	667	8	0.2	2.3	6	587	0	0.0	0.0
43	665	7	0.1	2.1	5	582	0	0.0	0.0
42	663	6	0.1	2.0	4	575	1	0.0	0.0
41	662	8	0.2	1.9	3	568	0	0.0	0.0
40	660	2	0.0	1.7	2	557	0	0.0	0.0
39	659	6	0.1	1.7	1	539	0	0.0	0.0
38	657	5	0.1	1.5	0	516	0	0.0	0.0

Note. [a] Freq (Frequency) is the number of students obtaining that score.
[b] Percent is the percentage of students at that score.
[c] Cumulative percent is the percentage of students at that score and below.

(Appendix 2 continued)

Frequency and Percentage Distributions for the HSCT Scores

Table 10

Frequency Distributions and Percentages for the HSCT Mathematics Section

Raw Score	Scale Score	Freq[a]	Percent[b]	Cumulative Percent[c]	Raw Score	Scale Score	Freq	Percent	Cumulative Percent
75	817	77	1.5	100.0	37	673	37	0.7	6.3
74	793	145	2.9	98.5	36	672	34	0.7	5.5
73	775	219	4.4	95.6	35	670	32	0.6	4.9
72	764	294	5.9	91.2	34	668	28	0.6	4.2
71	757	304	6.1	85.3	33	667	23	0.5	3.7
70	750	271	5.4	79.2	32	665	28	0.6	3.2
69	745	245	4.9	73.8	31	663	17	0.3	2.6
68	740	243	4.9	68.8	30	661	16	0.3	2.3
67	736	238	4.8	64.0	29	660	7	0.1	2.0
66	733	193	3.9	59.2	28	658	17	0.3	1.8
65	729	166	3.3	55.3	27	656	11	0.2	1.5
64	726	152	3.0	52.0	26	654	7	0.1	1.3
63	723	143	2.9	49.0	25	653	7	0.1	1.1
62	720	135	2.7	46.1	24	651	9	0.2	1.0
61	718	135	2.7	43.4	23	649	7	0.1	0.8
60	715	119	2.4	40.7	22	647	6	0.1	0.7
59	713	115	2.3	38.3	21	645	4	0.1	0.6
58	711	122	2.4	36.0	20	643	3	0.1	0.5
57	709	117	2.3	33.6	19	641	5	0.1	0.4
56	707	97	1.9	31.2	18	638	2	0.0	0.3
55	705	92	1.8	29.3	17	636	4	0.1	0.3
54	703	105	2.1	27.4	16	634	3	0.1	0.2
53	701	88	1.8	25.3	15	631	0	0.0	0.1
52	699	85	1.7	23.6	14	629	3	0.1	0.1
51	697	84	1.7	21.9	13	626	1	0.0	0.1
50	695	77	1.5	20.2	12	623	2	0.0	0.1
49	693	71	1.4	18.6	11	620	0	0.0	0.0
48	692	57	1.1	17.2	10	616	0	0.0	0.0
47	690	63	1.3	16.1	9	613	0	0.0	0.0
46	688	62	1.2	14.8	8	609	0	0.0	0.0
45	686	58	1.2	13.6	7	605	1	0.0	0.0
44	685	53	1.1	12.4	6	600	0	0.0	0.0
43	683	54	1.1	11.3	5	594	0	0.0	0.0
42	681	47	0.9	10.3	4	588	0	0.0	0.0
41	680	42	0.8	9.3	3	579	0	0.0	0.0
40	678	37	0.7	8.5	2	568	0	0.0	0.0
39	677	34	0.7	7.7	1	550	0	0.0	0.0
38	675	38	0.8	7.1	0	526	0	0.0	0.0

Note. [a] Freq (Frequency) is the number of students obtaining that score.
[b] Percent is the percentage of students at that score.
[c] Cumulative percent is the percentage of students at that score and below.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

REPRODUCTION BASIS

This document is covered by a signed "Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").