DOCUMENT RESUME

ED 327 575 TM 016 022

AUTHOR Beard, Jacob G.; And Others

TITLE The Prediction of State Student Assessment Test Scores from Scores on the Districts' Standardized Norm-Referenced Tests.

INSTITUTION Florida State Univ., Tallahassee. Coll. of Education.

SPONS AGENCY Florida State Dept. of Education, Tallahassee. Assessment, Testing, and Evaluation Section.

PUB DATE 21 May 90

CONTRACT FSDE-371-9095-89006

NOTE 64p.

PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC03 Plus Postage.

DESCRIPTORS Educational Assessment; *Elementary School Students; Elementary Secondary Education; Grade 3; Grade 5; Grade 8; Grade 10; *Norm Referenced Tests; *Predictive Measurement; *School Districts; Scores; Secondary School Students; *Standardized Tests; State Programs; Testing Programs; Test Results; *Test Use

IDENTIFIERS *Florida State Student Assessment Test

ABSTRACT

A study was conducted to examine the feasibility of using scores on the districts' standardized norm-referenced tests (NRTs) to predict performance on Florida's State Student Assessment Tests (SSAT) at grades 3, 5, and 8, and on the SSAT-I and SSAT-II given in grade 10. Specific questions about the use of district scores were addressed in a study of three districts involving 3,043 students in grade 3, 3,048 students in grade 5, 3,018 students in grade 8, and 2,046 students in grade 10. Findings indicate that students' scores from NRTs could be used to exempt students from the SSAT, the SSAT-I, and the SSAT-II. The NRT scores identified approximately 75% of students whose SSAT performance was unsatisfactory. The effectiveness of different NRTs used by the districts was similar, although different cutoff points were necessary and the results varied across grade level. Ten tables contain data from the study. An appendix provides a 15-page table of test scores and 15 scatterplots of the relationships between tests. (SLD)

THE PREDICTION OF STATE STUDENT ASSESSMENT TEST SCORES FROM SCORES ON THE DISTRICTS' STANDARDIZED NORM-REFERENCED TESTS

Prepared By

JACOB G. BEARD, RICHARD KEVAN,
RAPIN POSRIE, and VALERIE B. SHERIDAN

EDUCATIONAL RESEARCH AND EVALUATION SERVICES
COLLEGE OF EDUCATION, FLORIDA STATE UNIVERSITY

MAY 21, 1990

Supported by Grant Number 371-9095-89006

From the

ASSESSMENT, TESTING, AND EVALUATION SECTION
FLORIDA DEPARTMENT OF EDUCATION
TALLAHASSEE, FLORIDA

The Prediction of State Student Assessment Test Scores From Scores On the Districts' Standardized Norm-Referenced Tests

The purpose of this study was to examine the feasibility of using scores on standardized norm-referenced tests (NRT) administered by the districts to predict performance of students on Florida's State Student Assessment Tests (SSAT) at grades 3, 5, and 8, and on the SSAT-I and SSAT-II at grade 10.

The SSAT and SSAT-I were designed to test students for mastery of minimum competencies in communications and mathematics basic skills. The SSAT-II tests students for their minimum competency in the application of basic skills at the 10th grade level. Therefore, these tests are easy for students who exhibit above average achievement, which is shown by the median percentage of students complying with established minimum performance standards for the 1988-89 school year. These percentages were 94, 89, 84, and 89 at grade levels 3, 5, 8, and 10 respectively. The SSAT-II Communications subtest was passed by 85 percent of the students in 1988-89, and the Mathematics subtest by 76 percent.

These minimum competency tests might be considered redundant for a large proportion of the students who demonstrate acceptable achievement on their district's norm-referenced testing program. If this redundancy exists, it might be possible and practical to exempt some students from taking the SSAT on the basis of their norm-referenced test scores. The purpose of this study was to investigate relationships between the NRT scores and SSAT performance, and general issues associated with exempting the SSAT based on NRT scores.

Some subsidiary issues were investigated in the study. One issue had to do with whether equivalent predictions of SSAT performance could be made from different norm-referenced tests. Several different norm-referenced tests are in use in Florida at any given time, and districts may periodically change from one NRT to another. Therefore, it was important to investigate similarities and differences among NRT in predicting the SSAT scores.

A second issue was related to the proportion of the student population that could reasonably be exempt from taking the SSAT. It was assumed that school administrators would be reluctant to adopt a strategy in which roughly equal proportions of students would, and would not be exempt. The logistics required to attend to half the student population of a school while administering the SSAT to the remaining half would be inordinately cumbersome. Therefore, the proportion of students that might be exempt was of interest and would probably have implications for the likelihood of any such strategy being adopted.

A third issue involved the demographics of the group that would not exempt the SSAT. It would be particularly objectionable to institute a procedure which tended to segregate those exempt from and those tested by the SSAT according to racial/ethnic classification.

A fourth issue related to the current practice of using the percentage of students mastering the standards tested in the SSAT as an indicator of school and district achievement. Currently, the SSAT scores are the only objective index of achievement that is common across all schools and districts in the state. If an exemption procedure were implemented, how might a single composite score index for schools and districts be structured that would indicate SSAT performance, assuming that the majority of students would have only an NRT score and the remainder of the students would have both NRT and SSAT scores?

In order to investigate the major purpose and subsidiary problems the following specific questions were addressed in the study.

1. How well do NRT scores predict mastery of mathematics and communications standards assessed by the SSAT for the following grade levels, subjects, and test administrations?

a. NRT test scores obtained in the spring of grades 2, 4, and 7 used to predict SSAT performance in grades 3, 5, and 8 the following fall

b. NRT test scores obtained in grade 9 or in the fall of grade 10 used to predict SSAT-I performance in the spring of grade 10

c. NRT test scores obtained in grade 9 or in the fall of grade 10 used to predict scores on the mathematics and communications sections of the SSAT-II in the spring of grade 10

2. Can equivalent predictions of SSAT performance be made using scores from different NRT used by Florida school districts?

3. What proportions of students could reasonably be exempt from SSAT testing at each grade level?

4. How would a group of students likely to be selected for SSAT testing on the basis of NRT scores differ from the population of students as a whole, with reference to race, sex, geographic location, or other relevant characteristics?

5. How might a single composite score index for schools and districts be developed to indicate SSAT performance, assuming that the majority of students would have only an NRT score and the remainder of students would have both NRT and SSAT scores?

Method

Sampling Tests, Districts, and Students

It was proposed that data from three districts for each of three NRT be included in the study. We proceeded with the task of selecting a sample by preparing a computer file of names of norm-referenced tests used by districts. This file was based on the document entitled "Survey of Norm-Referenced Testing Programs in the Florida Public School Districts," provided to us by the SSAT program staff. This file was sorted on the "Test" field, which brought together all of the districts using each particular test. It was found that one test was used by 33 districts, a second was used by 13 districts, and a third test was used by 9 districts. Following these three tests was another that would account for a large number of students, but that was used in only two districts. Another test was used by three medium-sized districts; however, tapes of scores were not available for one of these districts. Therefore, the three most widely used NRT were chosen for use in the study; and hereafter these tests will be referred to by the pseudonyms: Test A, Test B, and Test C.

We next selected districts within tests. Three districts which varied in geographic location, and which had indicated that tapes of data were available, were chosen for each test.

Each district chosen was asked to provide data for at least 300 students at each of grades 3, 5, 8, and 10. This would provide initial samples of 900 students for each test. This sample size was large enough to allow for attrition in the matching process while leaving a sufficient sample for dependable results. In fact, the final sample sizes for grades 3, 5, and 8 were larger than proposed because some districts supplied more data than requested and because data from two districts became available after replacement districts were contacted. When extra data were available, the amount of data from some districts was randomly reduced to provide a sample of approximately 1,000 cases. The data reduction process was designed to improve the sample's representation of districts and of racial/ethnic groups.

Data proved to be more difficult to get at the tenth than at other grade levels. For Test B at the tenth grade, matched data were available for 197 students from one district, 499 from another, and for only 57 students from another. For Test C at the tenth grade, only one district was able to provide tenth grade scores and the analyses are based on that district's scores. The number of cases used in the analyses are shown in Table 1.

The mean scores, gender, and racial/ethnic classification of the students in each sample are also shown in Table 1. The samples for Test B have the fewest minority students and those for Test C have the most. The sample for Test C, tenth grade, has the smallest number of students and 91 percent of them are black. Caution should be used in comparing the results of different tests because of these differences in racial/ethnic composition of the samples.

Table 1

Characteristics of the Samples by Test and Grade Level

		Grade Level			
		3	5	8	10
Test A					
Number		1031	1023	1014	1006
Mean Total	NRT NCE	61.91	57.76	58.86	60.64
	SSAT Raw	119.93	175.84	198.86	140.21
Percentage	Male	46	48	50	48
	Female	54	52	50	52
	White	69	57	63	64
	Black	23	35	28	29
	Hispanic	6	7	6	4
Test B					
Number		1010	1022	1004	753
Mean Total	NRT NCE	57.59	57.00	56.31	53.73
	SSAT Raw	120.86	182.74	208.47	141.36
Percentage	Male	50	48	49	47
	Female	50	52	51	53
	White	73	75	77	74
	Black	25	21	19	21
	Hispanic	2	2	3	4
Test C					
Number		1002	1003	1000	287
Mean Total	NRT NCE	55.31	54.07	49.89	48.99
	SSAT Raw	118.58	175.23	196.37	128.48
Percentage	Male	51	47	46	44
	Female	48	53	54	56
	White	47	38	52	9
	Black	35	42	32	91
	Hispanic	12	18	15	--
Mean State SSAT Raw Score		120.04	178.96	199.73	139.17

Note. There are small percentages of other racial/ethnic groups.

Procedures

The study required that the school districts provide their students' NRT scores and that the Department of Education provide the SSAT scores. Letters requesting the districts' assistance were sent from the Director of the Division of Public Schools in the Florida Department of Education to the superintendents of the districts selected to participate in the study. All district superintendents who were contacted agreed to participate. Following their agreement to participate, a letter requesting the data was sent from the principal investigator to the test coordinators in these districts. Simultaneously, a request was sent to the Department of Education asking for the SSAT data for the districts selected. Only one selected district was unable to supply data. Two districts were initially unable to supply data, but went to considerable lengths to eventually provide useful data.

Upon receipt, the NRT data tapes were converted to a form usable by the Florida State University mainframe computer, and the relevant identification and score data were saved in files. These data records were then matched with the SSAT records supplied by the Department of Education. Some of the files were matched by Florida identification numbers; however, most of the matching was done using the first eight letters of the students' last names and the first six letters of their first names. Names, rather than identification numbers, proved to be a more effective means of matching the district records with state records. Many districts did not attempt to include identification numbers at the lower grade levels, and the numbers available at the upper grade levels contained many coding errors. Successful matches for districts varied from approximately 55 to 75 percent of the cases.

For districts providing large amounts of data, the records were matched on the mainframe computer and downloaded to personal computers. For other districts the data were downloaded to personal computers and matched using the SYSTAT statistical analysis package. The use of personal computers offered the advantage of being able to easily examine the matched data files to insure the correctness of the matching procedure. All statistical analyses were done using the SYSTAT package on IBM compatible personal computers.

Some of the districts used out-of-level forms in testing some of their students, and one district used out-of-level norms for transforming the raw scores of the out-of-level forms to standardized scores. These cases were removed from the files before analyzing the data.

Analyses

Norm-Referenced Subtests Included in the Study. There was variability in the selection of NRT subtests administered by the districts, both within and between tests. Subtests that were not common across districts using a particular test were not included in the analyses. The subtests that were common across districts and which were included in the analyses are shown in Table 2.

After several initial analyses (described later in the paper) were done to determine the most appropriate level of score aggregation for the main analyses, the mean of reading, language, and spelling subscores were computed to generate a "verbal" subscore, and the mean of mathematics scores were computed to form a "mathematics" subscore. The equally weighted means of these verbal and mathematics scores were then computed to generate a total score for each student for each test. These total scores were related to performance on the SSAT in the analyses.

Normal Curve Equivalent Scores. It was decided that Normal Curve Equivalent (NCE) scores would be used in the analyses in order to facilitate comparisons among tests. NCE scores are area transformations of national percentile ranks to scores having a mean of 50 and a standard deviation of 21.06. Although national percentile ranks were requested, some of the districts included NCE scores in the data submitted. For the remainder, NCE scores were computed by using an inverse normal cumulative distribution function to find the normal deviate below which lies the proportion of the normal distribution indicated by the percentile rank. This normal deviate was multiplied by 21.06 (the standard deviation of NCE scores) and added to 50 (the mean of NCE scores), yielding NCE scores. All analyses were done using NCE scores.

All NCE scores less than one were coded as missing values and excluded from the analyses. As a result the number of cases for the different subtests varied slightly.

Choosing an Index of Performance on the SSAT. It was necessary to choose a metric for indicating performance on the SSAT. The obvious choices were "number of items answered correctly" or "number of standards mastered." Discussions with Department of Education personnel suggested that the metric chosen should relate to the number of standards mastered, since performance has been typically reported in those terms. Therefore, when preliminary analyses indicated that the relationships with the NRT scores would be similar for the two types of indices, or scores, a decision was made to base the analyses on "number of standards mastered" on the SSAT. Hereafter, we will use the term "score" to refer to performance (number or percentage of standards mastered) on the SSAT and SSAT-I. The results would not have been substantially different had we chosen to use the number of items answered correctly.

The SSAT-II results are reported to students using an equated score scale on which a score of 700 is passing. These scaled scores are derived from the number of items answered correctly using Rasch procedures. The original "number correct" scores were used in the analyses. The SSAT-II total score consists of the sum of the number of communications and mathematics items answered correctly.

Choice of Score Aggregation Level for Study. Initially, descriptive statistics were computed for each grade level and subtest. These were inspected for aberrant data. Then, intercorrelations among the subtests were computed for each grade level and test. These correlations were examined to determine the interrelationship among the subtest scores, particularly the relationships among the verbal, mathematics, and total scores. These relationships were used in deciding whether to analyze the data separately by mathematics and communications subtests, or by total scores.

The intercorrelations among the NRT scores (verbal, mathematics, total) and the SSAT scores (communications, mathematics, total) are shown in Tables A-2 to A-4 of the Appendix. Selected mean correlations between the NRT and SSAT scores are shown in Table 3. (Computing means of correlations was justifiable in this case because the correlations averaged were highly similar.) These mean correlations show that, for all three tests and four grade levels, the NRT total scores predict the SSAT part scores as well as the NRT part scores. The NRT total scores predict SSAT total scores better than NRT part scores predict SSAT part scores, except for SSAT-II mathematics.

A more detailed analysis of these relationships can be seen in Table 4. This table shows a comparison of predictability using total and part scores for Test A at the eighth grade. These data show that, for an NRT NCE score of 50 which is the most plausible cut-off score, the number of Communications standards mastered is predicted almost as well from the NRT total score as from the NRT verbal score. The number of Mathematics standards mastered is predicted exactly as well from the NRT total scores as from the NRT math scores. Furthermore, the percentage of false positive and false negative predictions are the same or highly similar for the total and part score predictions. False positive errors occurred when students would have been exempt from the SSAT but in fact mastered less than 80 percent of the SSAT standards. False negative errors occurred when students would not have been exempt but in fact mastered more than 80 percent of the standards.

Table 2

Common Subtests by Test and Grade

Test	Subtests	Grade 3	5	8	10
A	Reading Vocabulary	X	X	X	X
	Reading Comprehension	X	X	X	X
	Total Reading	X	X	X	X
	Spelling	X	X	X	X
	Expression	X	X	X	X
	Mechanics	X	X	X	X
	Total Language Arts	X	X	X	X
	Computation	X	X	X	X
	Concepts & Applications	X	X	X	X
	Total Mathematics	X	X	X	X
B	Vocabulary	X	X	X	X
	Reading Comprehension	X	X	X	X
	Total Reading	X	X	X	X
	Spelling	X	X	X	X
	Language Expression	X	X	X	X
	Language Mechanics	X	X	X	X
	Total Language	X	X	X	X
	Computation	X	X	X	X
	Concepts & Applications	X	X	X	X
	Total Mathematics	X	X	X	X
C	Reading Comprehension	X	X	X	X
	Word Study Skills	X	X	X	
	Total Reading	X	X	X	
	Language		X	X	X
	Spelling	X	X	X	X
	Total Language		X	X	X
	Concepts of Numbers	X	X	X	X
	Computation	X	X	X	X
	Applications	X	X	X	X
	Total Mathematics	X	X	X	X

Table 3

Selected Mean Correlations

	vC		mM		tC		tM		tT	
	M	SD	M	SD	M	SD	M	SD	M	SD
Three Tests: Five Levels N = 15	.556	.057	.624	.078	.563	.059	.613	.075	.656	.050
Test A: All Grades N = 5	.562	.050	.618	.092	.571	.054	.618	.084	.660	.048
Test B: All Grades N = 5	.557	.069	.613	.087	.556	.076	.602	.089	.657	.056
Test C: All Grades N = 5	.548	.063	.641	.067	.563	.057	.619	.067	.651	.057
Grade 3: All Tests N = 3	.476	.031	.508	.015	.484	.015	.512	.007	.574	.020
Grade 5: All Tests N = 3	.531	.034	.651	.031	.523	.041	.646	.022	.660	.031
Grade 8: All Tests N = 3	.593	.037	.661	.022	.600	.041	.643	.038	.683	.027
Grade 10 SSAT-I: All Tests N = 3	.604	.051	.594	.071	.602	.035	.565	.060	.661	.018
Grade 10 SSAT-II: All Tests N = 3	.577	.005	.705	.030	.607	.012	.700	.028	.701	.028

Note. v = NRT Verbal
C = SSAT Communications
m = NRT Mathematics
M = SSAT Mathematics
t = NRT Total
T = SSAT Total

Because the NRT total scores predicted the communications and mathematics subtest scores approximately as well as the NRT verbal and mathematics subscores, respectively, and because the NRT total scores predicted the SSAT total scores better than the part scores, it was decided to focus on the use of NRT total scores to predict SSAT total scores. The correctness of this decision was supported by a consideration of the logistical implications of alternative exemption procedures. For example, there would be fewer practical problems in exempting students from the entire SSAT than in exempting some students from the communications part and others from the mathematics part. The perspective implied by this approach is that exempt students are assumed to have adequate basic skills while those not exempted would take the SSAT and demonstrate competence in both communications and mathematics basic skills, or would be given remedial instruction in one or both of these areas.

It was also necessary to establish a criterion for satisfactory performance on the SSAT since there was no inherent "passing" score. Extensive discussions were held with DOE personnel in deciding that mastering 80 percent of the standards would constitute satisfactory performance on the SSAT and SSAT-I. The 80 percent value was a relatively stringent criterion of successful performance in view of the fact that 90, 78, 74, and 72 percent of the students in the samples of this study met this criterion of success; i. e., mastered more than 80 percent of the standards at grade levels 3, 5, 8, and 10 respectively, while standards of median difficulty were mastered by 94, 89, 84, and 89 percent of the students, statewide, at grades 3, 5, 8, and 10 respectively. The latter values are sometimes used as indicators of success rates at these grade levels.

Passing scores for the SSAT-II Communications and Mathematics subtests have been set by the Florida Board of Education. For the 1988-89 administration these official passing scores were equivalent to number correct scores of 62 for Communications and 54 for Mathematics. For the purposes of this study, these two raw scores were added together to yield an index of satisfactory performance for the total SSAT-II of 116.

Analysis of Relationships Between NRT and SSAT Scores. The major purpose of the data analysis was to locate a cutting score on the NRT tests which best predicted satisfactory performance (as defined for the study) on the SSAT. "Best prediction" was defined as a prediction which maximized the number of hits (correct predictions) and which yielded fewer false positive than false negative errors. False positive errors were viewed as more critical than false negative because these students could have deficiencies in basic skills which might go undetected and for which they might not receive remedial instruction.

Scatterplots of the NRT and SSAT scores were made for each test and grade level combination. A line of best fit was plotted for each scatterplot using the LOWESS smoothing technique. This curve-fitting technique found the average y-value for the 20 percent of cases closest to each x-value and plotted these values as a line of best fit.

For example, a scatterplot of the relationship between NRT and SSAT total scores for test A at the eighth grade level is shown in Figure 1. (Scatterplots for each combination of Test and Grade Level are shown in the Appendix.) The line of relationship fitted through the LOWESS smoothing technique has been superimposed on the scatterplot.

These scatterplots were found to be distinctly nonlinear. The SSAT scores initially increased as the NRT scores increased, but the rate of increase diminished substantially in the upper regions of the NRT score scales. This pattern reflects the relatively low ceiling of the SSAT tests. That is, students in the upper two-thirds of the achievement distribution tend to master most of the SSAT standards.

The shapes of these scatterplots ruled out reliance upon statistical procedures which assume linearity, normality, and homoscedasticity. Instead, the scatterplots were inspected to find possible cutting scores on the NRT score scale which would optimize the prediction of successful SSAT performance according to the criteria given above; i. e., yielded the maximum percentage of "hits," or correct predictions, with fewer false positive than false negative errors.

Three cutting scores were chosen for each scatterplot. First, a score was chosen which appeared to best meet the criteria given above. Then scores five NCE points below and five points above the "best" score were also selected for inclusion in the analyses. The scores were dichotomized at the selected points on the NRT NCE scale and at the passing score on the SSAT scale. The percentage of hits, false positives, and false negatives were then calculated for each of the potential cutting scores.

Additionally, the racial composition of each hypothetically exempt and non-exempt group was calculated.

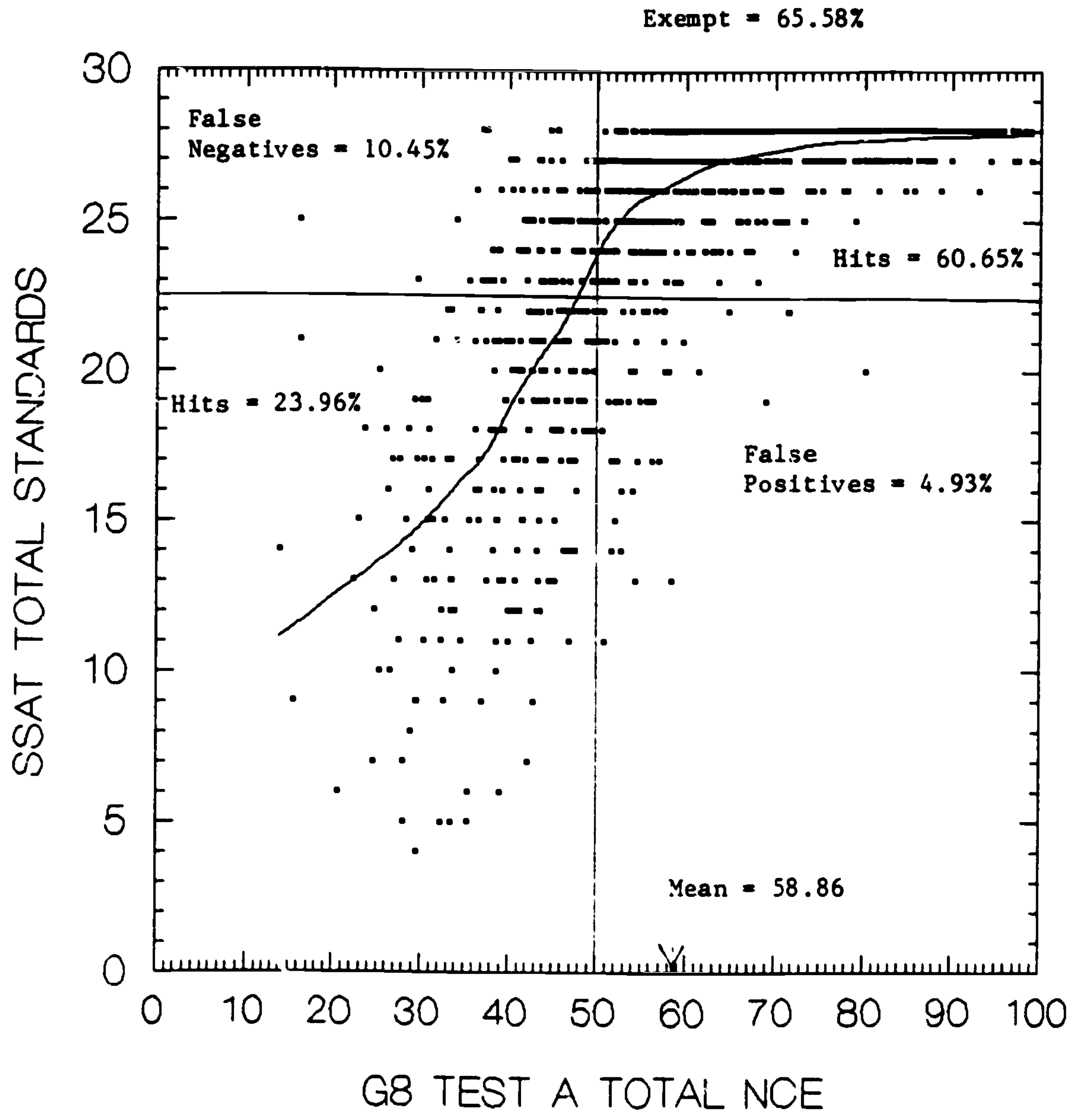


Figure 1. An Example of a Scatterplot of Norm-Referenced and SSAT Scores

Results

Selected means of correlations between the NRT and SSAT scores are summarized in Table 3. Several patterns are apparent in these data. (These correlations understate the magnitude of these relationships because the relationships are curvilinear. However their relative sizes are meaningful.)

First, it can be seen that the correlations between NRT and SSAT total scores are higher than those involving part scores. Also, the correlations involving mathematics scores are higher than those involving communications scores.

Second, the mean correlations, across grades, for Tests A, B, and C are similar, indicating that the three different NRT tests predict SSAT performance to about the same degree. The largest difference in mean correlation among the three tests for "total-total" correlations was (.009). This indicates that different NRT would have similar effectiveness in exempting students from taking the SSAT.

Third, the predictability of SSAT scores from NRT scores increases from the third to the eighth grade. From the eighth to the tenth grade the predictability for communications is similar, while the predictability for mathematics is less at the tenth than at the eighth grade. At the tenth grade the predictability for the SSAT-II is greater than for the SSAT-I.

The mean correlations shown in Table 3 generally indicate the pattern of predictability among the different tests, grade levels, and subject matter areas. However, the nature of the relationships can be seen more precisely in scatterplots of the paired variables.

The scatterplot in Figure 1 is typical of all grade levels and tests in general form , and shows that the number of SSAT standards mastered increased rapidly as the NRT NCE scores increased from the lowest scores up to a point near the mean (for this set of data) of 59. Beyond that point almost all students exceeded the cut-off point of 22.4 standards (80%) mastered. However, large numbers of students having NRT scores between 40 and 59 also mastered more than 80 percent of the standards. By drawing a vertical line at a hypothetical NRT cut-off score of 50 and a horizontal line at the SSAT cut-off point of 22.4 on the SSAT scale, the general nature and magnitude of prediction errors can be observed. The resulting upper right and lower left quadrants contain correct predictions, or "hits". The lower right quadrant contains "false positive" predictions; i. e., students who were predicted to master the SSAT but did not. The upper left quadrant contains "false negatives"; i. e., students who were predicted to score below the SSAT cut-off point, but in fact exceeded it.

Scatterplots, similar to that in Figure 1, for every test and grade level are shown in Figures A-1 through A-15 of the Appendix.

The scatterplots are excellent devices for showing the relationship between the paired variables, and were useful in choosing tentative cut-offs for the NRT scores. However, the large number of cases in the data sets caused many data points to be superimposed on others, making it difficult to ascertain the proportion of hits, false positive, and false negative cases. Therefore, additional analyses were done in order to determine more precisely the percentage of these errors and the percentage of correct classifications made. These analyses included the dichotomization of the NRT scores at hypothetical cut-off scores and the SSAT performance at the preselected "passing" criteria. The exact percentage of cases in each of the quadrants was then computed. The analysis was done for three NRT cut-off scores in the region expected to provide optimal results. "Optimal result" was defined as one having a near-maximum percentage of "hits" and fewer false positive than false negative errors. The rationale for this definition was that it would be more serious to erroneously label a true non-master as a master than to label a true master as a non-master.

These data are shown in Table 5 and summarized in Table 6. Three NRT cut-off scores are given for each test and grade level combination in Table 5. Generally, the middle cut-off score meets the criteria described above, and will be referred to as the "proposed" NRT cut-off score. These data are useful in observing the effects of different NRT cut-off scores on the number and direction of errors of prediction.

Table 6 includes data only for the recommended cut-off scores, which facilitates interpretation of these data. The results shown in Table 6 are best interpreted by examining the data for grade 3 separately from that for grades 5 through 10. For grade 3, an NRT cut-off score of 40 is proposed for all three tests, and the mean results represent those for the three tests quite well. For this combined group, 90 percent of the students met the established SSAT passing criterion. Overall, 84 percent of the students would have been exempted by the proposed cut-off score. Eighty-eight percent of the exempt-nonexempt decisions would have been correct; and, of the 12 percent errors, 3 percent would have been false positive and 9 percent would have been false negative errors.

Overall, greater accuracy could have been obtained by predicting that all students would pass the SSAT. However, an unidentified 10 percent of this third grade group would have been deficient by SSAT standards. Seven of this 10 percent would have been correctly identified using the proposed NRT cut-off score. Therefore, using the exemption procedure, only 16 percent of the students would have been required to take the SSAT and only 3 of the 10 percent of SSAT deficient students would not have been identified.

The results for grades 5 through 10 are similar except for the anomalous data for test C at the tenth grade (see Table 1). Taking the means for grade eight as an example, the SSAT pass rate was 74 percent. The proposed NRT cut-off score would have exempted 67 percent of the students, leaving 33 percent to take the SSAT. Only 5 of the 26 percent SSAT deficient students would have remained unidentified.

Table 5

Accuracy in Predicting Performance on the SSAT Using Different Cut-Off Scores on the NRT

Test	Grade	Cut-Off Score	Percentage Exempt	Hits	False Positive	False Negative
A	3	35	90	90	5	5
	3	40	86	89	4	8
	3	45	79	84	2	13
A	5	40	82	81	14	5
	5	45	72	82	9	9
	5	50	66	80	4	16
A	8	45	78	84	12	5
	8	50	66	85	5	10
	8	55	55	80	2	18
A	SSAT-I	40	85	85	11	4
	SSAT-I	45	78	85	7	8
	SSAT-I	50	67	82	3	15
A	SSAT-II	40	85	87	7	7
	SSAT-II	45	78	84	4	12
	SSAT-II	50	67	78	2	20
B	3	35	91	92	3	4
	3	40	85	88	2	10
	3	45	77	82	1	17
B	5	35	93	89	9	2
	5	40	87	90	5	5
	5	45	78	86	3	12
B	8	35	89	86	9	4
	8	40	83	86	7	7
	8	45	74	83	4	13
B	SSAT-I	35	84	84	10	6
	SSAT-I	40	75	84	6	10
	SSAT-I	45	63	79	2	18
B	SSAT-II	35	84	84	8	7
	SSAT-II	40	75	82	5	13
	SSAT-II	45	63	75	3	22

(table continues)

(Table 5 continued)

Test	Grade	Cut-Off Score	Percentage Exempt	Hits	False Positive	False Negative
C	3	35	90	89	6	5
	3	40	82	86	4	10
	3	45	71	79	2	19
C	5	40	79	82	12	5
	5	45	69	83	7	10
	5	50	59	79	3	18
C	8	35	86	77	21	2
	8	40	76	80	14	5
	8	45	60	80	6	13
C	SSAT-I	45	61	75	13	12
	SSAT-I	50	45	75	6	19
	SSAT-I	55	32	69	2	29
C	SSAT-II	40	75	78	15	7
	SSAT-II	45	61	77	8	15
	SSAT-II	50	45	72	3	25

Table 6

Accuracy in Predicting SSAT Performance From Recommended NRT Cut-Off Scores and Percentage of Students Exempt by Racial/Ethnic Group

Test	Grade	NRT Cut Score	Percentage: SSAT Pass	Exempt: All	Exempt: White	Exempt: Black	Exempt: Hisp.	Hits	False Positive	False Negative
Test A	3	40	90	86	89	72	93	89	4	8
Test B	3	40	92	85	90	69	83	88	2	10
Test C	3	40	88	82	93	67	79	86	4	10
MEAN			90	84	91	69	85	88	3	9
Test A	5	45	73	72	79	60	72	82	9	9
Test B	5	40	87	87	92	72	74	90	5	5
Test C	5	45	73	69	81	50	73	83	7	10
Mean			78	76	84	61	73	85	7	8
Test A	8	50	71	66	77	39	54	85	5	10
Test B	8	45	83	74	79	53	79	83	4	13
Test C	8	45	67	60	81	44	53	80	6	13
Mean			74	67	79	45	62	83	5	12
Test A	SSAT-I	45	79	78	86	58	76	85	7	8
Test B	SSAT-I	40	79	75	83	47	63	84	6	10
Test C	SSAT-I	50	59	45	68	43	na	75	6	19
Mean			72	66	79	49	70	81	6	12
Test A	SSAT-II	45	85	78	86	58	76	84	4	12
Test B	SSAT-II	40	83	75	83	47	63	82	5	13
Test C	SSAT-II	45	67	61	76	59	--	77	8	15
Mean			78	71	82	55	46	81	6	13

Raising the NRT cut-off scores would exempt fewer students and decrease the percentage of false positive (unidentified SSAT deficient) students. The effect of raising or lowering the NRT cut-off can be seen in Table 5 where results are given for three cut-off scores for each grade level by test combination.

Lowering the NRT cut-off scores would exempt more students, leaving fewer to take the SSAT, but would increase the percentage of false positive (unidentified SSAT deficient) students. On the other hand, the percentage of students having to take the SSAT when they have already mastered the SSAT standards would decrease.

The NRT NCE scores of 40 and 45 were the most frequently occurring "proposed" cut-off scores. The only exceptions were two cases in which 50 was chosen. These cut-off scores amount to percentile ranks, for these samples, of 16 at the third grade level and approximately 30 at grades 5, 8, and 10.

An NCE score of 40 is equivalent to a national percentile rank of 32. For these Florida samples an NCE score of 40 usually functioned like a somewhat lower state percentile rank; i. e., fewer than 32 percent scored lower than an NCE of 40, suggesting that these scores are greater than those in the national norm groups. This conclusion is supported by the mean NCE scores for these groups which typically was greater than 50.

Theoretically, the achievement level of the samples used in this study should be biased upward since the records of transient students who moved from one school to another between the school years would not have been matched and such students would not have been included in these samples. However, Table 1 shows that the mean SSAT scores for the samples used in the study are near the mean statewide scores given in the Technical Report for 1988-89 (1989).

Inspection of these data does not permit conclusions that the norms of any of the three tests are generally "easier" than the others. However, these data might be used to research this question further.

The proposed cut-off scores were most consistent across grade levels for Test B. For Test A the cut-off scores increased with grade levels three through eight, and for Test C there was no discernible pattern.

Taylor and Russell (1939) presented some interesting approaches for showing the practical effectiveness of using selection instruments. One of these approaches was to compare the percentage of "successful examinees" in the selected group with the percentage in the unselected group. According to our definition of success on the SSAT, and using the eighth grade Test A scores as an example, 71 percent of the students in the total group were successful. Of those students selected as probably being successful by the NRT, 92 percent were in fact successful; of the students below the NRT cut-off score, thirty percent were successful.

These data are shown for all tests and grade levels in Table 7. For example, at the fifth grade level an average of 78 percent of the students in the total sample, 91 percent of the students scoring at or above the NRT cut-off, and 35 percent of the student scoring below the NRT cut-off score passed the SSAT. The selection procedure can be considered valuable to the extent that the values in the next-to-last column are large and those in the last column are small, relative to those for the total group.

The NRT are relatively accurate in identifying large percentages of students who are likely to master the SSAT. On the other hand, of those students identified as likely to fail, substantial percentages also mastered the SSAT. This characteristic of the results is in part a function of the criterion used in choosing an NRT cut-off score. That is, an NRT cut-off score was deliberately chosen which produced fewer false positive than false negative errors.

Table 7

Percentage Passing the SSAT for: The Total Group, Those Above the NRT Cut-Off, and Those Below the NRT Cut-off

Test	Grade	NRT Cut Score	Percentage Passing the SSAT of		
			Total Group	NRT ≥ Cut-Off	NRT < Cut-Off
Test A	3	40	90	96	54
Test B	3	40	92	97	64
Test C	3	40	88	96	56
Mean			90	96	58
Test A	5	45	73	88	34
Test B	5	40	87	94	38
Test C	5	45	73	90	33
Mean			78	91	35
Test A	8	50	71	92	30
Test B	8	45	83	95	42
Test C	8	45	67	89	33
Mean			74	92	38
Test A	10	45	79	91	37
Test B	10	40	79	93	40
Test C	10	50	59	88	35
Mean			72	91	37
Test A	10-II	45	85	95	54
Test B	10-II	40	83	93	51
Test C	10-II	45	67	87	37
Mean			78	92	47

Note. For the SSAT and SSAT-I passing was 80 percent of the standards. For the SSAT-II, passing was a number correct score of 116, which was the sum of passing number correct scores on Communications (62) and Mathematics (54) for March of 1989.

Racial Cha acteristics of Non-Exempt Students

The mean test scores of minority groups are typically lower than those of the majority group in the state of Florida. In view of this fact, an exemption process could have the effect of segregating exempt and non-exempt students along racial/ethnic lines. The data were examined to determine the potential effects of proposed exemption procedures on the racial/ethnic composition of exempt and non-exempt groups.

The percentage of students in each of four racial/ethnic groups who would have been exempt from the SSAT are shown in Table 6. These percentages are shown for "All" students in the sample, and for white, black, and Hispanic students. The mean percentages for the combined three tests at each grade level are also shown. These data show that substantial percentages of all groups would be exempt; however, larger percentages of white than Hispanic, and Hispanic than black students, would be exempt. Larger percentages of all groups would be exempt at the lower grade levels than at the upper.

Table 8 shows the percentage of the major racial groups not exempt by grade level. These percentages do not describe the make-up of any particular non-exempt school group which might take the SSAT, since that would depend on the racial make-up of the particular school.

Table 9 gives the Sexual and Racial/Ethnic composition of the total grade-level samples and the non-exempt subgroups of each sample. Generally, the non-exempt groups include slightly greater percentages of male students than the total samples. The non-exempt groups also include smaller percentages of white and larger percentages of black students than the total samples. The larger percentage of black students in the total sample and in the non-exempt groups for SSAT-I and SSAT-II was caused, at least in part, by the fact that the tenth grade sample for Test C, although small, included 91 percent black students.

The percentages of Hispanic students were slightly greater in the non-exempt groups than the total samples only at grades 5 and 8.

An example of the differences in composition of one of the samples and its non-exempt subgroup is shown in Table 10. The percentages of white, black, Hispanic, and Asian students, by sex, in the total sample for Test A at the tenth grade level are shown in the left side of the table. The corresponding percentages for the non-exempt students are shown in the right side. Male students were slightly more prevalent in the non-exempt group than in the total sample. Black students were almost twice as prevalent, and white students were substantially less prevalent in the non-exempt group than in the total sample. Had this particular group been the test population for a school, the non-exempt group would have had approximately 50 percent more black than white students.

Table 8

<u>Percentage of Major Racial Groups Not Exempting the SSAT</u>

Grade	Percentage of Racial Group			
	All	White	Black	Hispanic
3	16	9	31	15
5	23	16	39	27
8	33	21	55	38
10 SSAT-I	34	21	51	30
10 SSAT-II	29	18	45	30

Table 9

<u>The Composition of Total and Non-Exempt Groups by Grade Level</u>

Sexual or Racial/Ethnic Classification	Grade Level									
	3		5		8		SSAT-I[a]		SSAT-II[a]	
	Tot.	NE	Tot.	NE	Tot.	NE	Tot.	NE	Tot.	NE
Male	49	55	47	51	48	52	47	48	47	50
Female	51	44	53	48	51	48	53	52	53	50
White	63	37	61	39	59	38	60	33	60	36
Black	27	54	29	50	30	49	35	62	35	59
Hispanic	7	7	8	9	9	12	4	4	4	4
Asian	1	1	1	1	2	-	1	-	1	-

[a]The sample for Test C was anomalous in its composition (see Table 1).

<u>Note</u>. Tot. = Total sample
NE = Non-Exempt students

Table 10

An Example of Constituency of Total and Non-Exempt Groups: Test A Tenth Grade

Racial/Ethnic Classification	Percentage Constituency			
	Total Group		Non-Exempt Group	
	Male	Female	Male	Female
White	32	32	21	18
Black	14	15	29	27
Hispanic	2	3	1	4
Asian	1	1	0	1

Reporting School and District Achievement If an Exemption Procedure were Used

Florida Statutes require that the composite student performance of a school or program be reported. Since 1977 the composite score for grade 3, 5, 8, and 10 (SSAT-I) has been the average percentage of students mastering each standard within reading, writing, and mathematics as measured through the SSAT. Composite scores for reading, writing, and mathematics are reported for schools, districts, regions, and the State. Composite scores below 80 indicate deficient programs.

Composite scores for the SSAT-II consist of the percentage of students passing the Communications and Mathematics parts of the SSAT-II.

How might a single composite score index for schools and districts be developed to indicate SSAT performance, assuming that the majority of students would have only an NRT score and the remaining students would have both NRT and SSAT scores? The nature of this composite score would differ depending on assumptions made about its purpose. New composite score indices are recommended based on the following two assumptions.

Assumption A. The composite score is to indicate percentage mastering minimum standards. (1). In this case the problem is to structure a composite score which indicates the percentage of students who have mastered the minimum standards. An index could be created which estimates this percentage in a reasonably direct way. First, the percentage of standards mastered by the non-exempt group taking the SSAT would be multiplied by the percentage of the school or district's student population in that group. Next, a constant consisting of a mean percentage of standards mastered by those exempting the SSAT would be multiplied by the percentage of students exempting the SSAT. Next, these two products would be summed to yield an estimate of the percentage of standards mastered by the entire school or district.

The computations are represented by the following formula.

$$(A \times B) + (C \times D) = E$$

where:
- A = mean percentage standards mastered for non-exempt group
- B = percentage of students not exempting SSAT
- C = statewide mean percentage standards mastered for exempt group (a constant)
- D = percentage of students exempting SSAT
- E = estimate of percentage of standards mastered

For example, in our sample data for Test A at grade 8:

A = 68.83 = mean percentage standards mastered for non-exempt group
B = 34.41 = percentage not exempting SSAT
C = 65.58 = percentage exempting SSAT
D = 94.15 = (constant) mean percentage standards mastered for exempt students

$$[(68.83)(34.41) + (94.15)(65.58)] / 100 = 85.43$$

Our estimate of the percentage of standards mastered would be 85.43. However, in this case the constant value (D) was a calculated value rather than an estimate, and our index value of 85.43 is very close to the directly computed value of 85.61. This procedure should provide very accurate estimates of the percentage of standards mastered since it incorporates only one estimated value, "D", which could be estimated with a great deal of precision.

(2). An alternative composite score could be derived by using 100% as the "mean percentage standards mastered for exempt" students in the formula given above. This would yield a composite score which is about 4% higher for the above example and would always yield a composite score that is slightly higher than the one currently used. A rationale for using this estimate is that students who score above the NRT cut-off score have demonstrated minimum achievement in an alternate way. An advantage of this procedure is that it does not require an estimate of percentage standards mastered by the exempt group.

Assumption B. The composite score is to indicate achievement level. In this case the composite score is intended to indicate the overall level of academic achievement in the particular school or district. The composite score could be derived by computing the mean NCE or T score of designated NRT subtests for the entire school or district. This mean score would indicate the school or district's overall achievement more accurately than the current composite score, or than the index described in the preceding paragraphs.

The current scores are subject to ceiling effects which may limit, to some unknown degree, conclusions about a school's overall achievement level. The current composite score indicates success in raising students' achievement level to the minimums defined by the SSAT. Therefore, schools which emphasize basic skills would be expected to do well. On the other hand, schools which emphasize higher order knowledge and skills might not score as well using the current composite index as some schools having lower overall achievement

In order to compare mean scores from different NRT, it would be necessary to equate their score scales. Adjusting constants could be derived for placing means from different NRT onto a common scale.

Another potential problem with a composite score that relies exclusively on the districts' NRT scores is that a school or district might not choose to administer an NRT for the grade level being assessed. In that case, administration of the SSAT to all their students might be required.

Conclusions

Conclusions regarding the questions addressed in this study are given in the following paragraphs.

Question 1. How well do norm-referenced test scores predict mastery of mathematics and communications standards assessed by the SSAT for the following grade levels, subjects, and test administrations?

a. norm-referenced test scores obtained in the Spring used to predict SSAT performance in grades 3, 5, and 8 the following fall

The mean Pearson product-moment correlations between total scores on the NRT and performance on the SSAT for Tests A, B, and C combined were .57, .66, and .68 at grades 3, 5, and 8 respectively. These correlations understate the actual relationships between the variables because they are not linear (see Figure 1). When the relationship was plotted and optimum cut-off points on the NRT were defined, it was found that: at the third grade, 84 percent of the students would be exempt, 88 percent of the predictions of SSAT success were correct, and 3 percent of the errors were false positive; i. e., predicted to pass the SSAT but actually failed.

At the fifth grade, 77 percent of the students would have been exempt, 85 percent of the predictions were correct, and there were 7 percent false positive errors.

At the eighth grade, 67 percent of the students would have been exempt, 83 percent of the predictions were correct, and there were 5 percent false positive errors.

b. norm-referenced test scores obtained in grade 9 or in the Fall of grade 10 used to predict SSAT-I performance in the Spring of grade 10

For the SSAT-I at grade 10, the correlation between NRT scores and performance on the SSAT was .66. Again, this relationship was nonlinear. The optimal NRT cut-off score would have exempted 66 percent of the students, 81 percent of the predictions would have been correct, and there would have been six percent false positive errors.

c. norm-referenced test scores obtained in grade 9 or in the Fall of grade 10 used to predict scores on the mathematics and communications sections of the SSAT-II in the Spring of grade 10

The correlation between the NRT scores and performance on the SSAT-II was .70. Seventy-one percent of the students would have been exempt, 81 percent of the predictions were correct, and there were seven percent false positive errors.

Generally speaking, an NRT cut-off score which would exempt approximately 70 percent of the students from taking the SSAT results in approximately 6 percent false positive errors and approximately 11 percent false negative errors. The results are best at the third grade level where the exemption of about 85 percent of the students would yield fewer errors than at other grade levels. Otherwise, the exemption procedures proposed here would have detected approximately three-fourths of the approximately 25 percent of students considered to be deficient according to SSAT standards.

Question 2. Can equivalent predictions of SSAT performance be made using scores from different norm-referenced tests used by Florida school districts?.

The different norm-referenced tests were found to be similar in their ability to predict the SSAT scores as evidenced by the percentage of "hits" and errors shown in Table 6. However, the NRT cut-off scores found to be optimal for predicting success on the SSAT varied among the different NRT tests by as much as 10 NCE points at one grade level. Differences of this magnitude would require the use of different cut-off scores for different norm-referenced tests, or would require the calibration or equating of the different NRT to a single scale.

Question 3. What proportions of students could reasonably be exempted from SSAT testing at each grade level?

The following conclusions are based on means of values for the three tests included in the study as shown in Table 6. At the third grade level, 84 percent of the students would exempt the SSAT, 77 percent at the fifth grade, and 67 percent would exempt at the eighth grade. At the tenth grade level the sample of students for Test C were anomalous and the following conclusions are based on the results from Test A and Test B only. Seventy-seven percent of the students would have exempted the SSAT-I and the same percentage would have exempted the SSAT-II. These exemption percentages are not precise estimates of statewide exemption percentages because the samples of students used were not randomly chosen from the state population. However, Table 1 shows that the SSAT means of the samples were close to the statewide SSAT means except for Test C at the tenth grade.

Question 4. How would a group of students likely to be selected for SSAT testing on the basis of norm-referenced test scores differ from the population of students as a whole with reference to race, sex, geographic location, or other relevant characteristics.

At the third grade, 486 of 3043 students would not have exempted the SSAT. Thirty-seven percent of the 486 students were white, 54 percent were black, 7 percent were Hispanic, and 1 percent Asian.

At the fifth grade, 723 of 3048 students would not have exempted the SSAT. Of the 723, 39 percent were white, 50 percent were black, 9 percent were Hispanic, and 1 percent were Asian.

At the eighth grade, 1011 of 3018 students would not have exempted the SSAT. Thirty-eight percent of the 1011 students were white, 49 percent were black, 12 percent were Hispanic, and less than one percent were Asian.

For SSAT-I, 575 of 2046 students would not have exempted. Of the 575 students, 33 percent were white, 62 percent were black, 4 percent were Hispanic, and less than one percent were Asian.

For SSAT-II, 528 of 2046 students would not have exempted. Of this group, 36 percent were white, 59 percent were black, 4 percent were Hispanic, and less than one percent were Asian.

An additional analysis was done for the tenth grade because the data for Test C were distinctively non-representative in racial/ethnic composition. The data for Test A and Test B were combined and the data for Test C were not included. For these data, 415 of 1759 students would not have exempted the SSAT-I. Of these, 44 percent were white, 50 percent were black, 5 percent were Hispanic, and less than one percent were Asian. The number of students not exempting and the percentages in the racial/ethnic classifications were identical for SSAT-I and SSAT-II since the NRT cut-off scores were identical for Tests A and B.

The composition of these non-exempt groups can be compared to the composition of the total samples given in Table 1. The non-exempt samples generally contain fewer white and more black and Hispanic students than the total sample.

Question 5. How might a single composite score index for schools and districts be developed to indicate SSAT performance, assuming that the majority of students would have only a norm-referenced score and the balance of students would have both norm referenced and SSAT scores?

Three alternative composite score indices were proposed. Of these three procedures, we recommend that the two procedures described under (A.1.) and (B.) of pages 25 and 26 be used. One of the indices provides an estimate of the mean percentage of SSAT standards mastered, and the other consists of the mean of norm-referenced test scores adjusted for the particular test administered. These two indices assess school achievement from different perspectives and a combination of the two would describe schools' and districts' achievement better than either alone.

Summary of Conclusions

We conclude, based on the results of the study, that students' scores from NRT administered by the districts could be used to exempt the students from the SSAT, the SSAT-I, and the SSAT-II. The NRT scores identified approximately three-fourths of the students whose SSAT performance was considered unsatisfactory. The effectiveness of different NRT were similar, but different cut-off points were necessary, and the results varied across grade level.

The SSAT passing rate for the third grade was so high, 90 percent, that little would be gained by the exemption process in a statistical sense. In fact, one could be correct for 90 percent of the students by merely predicting, without NRT scores, that all students would perform satisfactorily on the SSAT. However, it was possible to use the NRT scores to identify 7 of the 10 percent who did not perform satisfactorily on the SSAT. This identification of students considered to be deficient in basic skills would be educationally important, provided that remedial instruction could be provided.

Reference

Taylor, H. C. & Russell, J. T. (1939). The relationship of validity coefficients to the practical effectiveness of tests in selection. Journal of Applied Psychology, 23, 565-578.

Assessment, Testing, and Evaluation Section, Division of Public Schools, Florida Department of Education. (1989). Technical Report: 1988-89 State Student Assessment Test. Tallahassee, FL: Department of Education.

APPENDIX

Key For Abbreviations Used in Table A-1

NRT		SSAT	
RVOC	- Reading Vocabulary	CIA	- Communication Items
RCOMP	- Reading Comprehension	MIA	- Mathematics Items
RTOT	- Reading Total	TIA	- Total Items
SPELL	- Spelling	CSTA	- Communications Standards
LMECH	- Language Mechanics	MSTA	- Mathematics Standards
LISTEN	- Listening	TSTA	- Total Standards
LEXP	- Language Total	TRSA	- Reading Standards
LTOT	- Language Total	TWSA	- Writing Standards
MCOMP	- Mathematics Computation	COMM	- SSAT-II Communications
MAPP	- Mathematics Applications	MATH	- SSAT-II Mathematics
MCONC	- Mathematics Concepts	TOMTH	- SSAT-II TOTAL
MCA	- Mathematics Concepts and Applications	DIST	- District Identification
MTOT	- Mathematics Total		
VERBAL	- Total Verbal		
ATOTAL	- Mean of Mathematics Total and Total Verbal for Test A		
BTOTAL	- Mean of Mathematics Total and Total Verbal for Test B		
CTOTAL	- Mean of Mathematics Total and Total Verbal for Test C		

Table A-1

Descriptive Statistics for Samples by Test and Grade

TEST A GRADE 3
TOTAL OBSERVATIONS: 1031

	RVOC	RCOMP	RTOT	SPELL	LMECH
N OF CASES	1028	1030	1028	660	1029
MINIMUM	1.000	1.000	1.000	1.000	4.000
MAXIMUM	95.000	98.000	99.000	91.000	98.000
MEAN	57.881	55.993	57.215	53.658	63.092
STANDARD DEV	23.072	24.103	24.277	19.175	24.524

	LEXP	LTOT	MCOMP	MAPP	MTOT
N OF CASES	1030	1029	1030	1030	1030
MINIMUM	1.000	1.000	1.000	1.000	1.000
MAXIMUM	96.000	99.000	97.000	99.000	99.000
MEAN	54.497	60.062	63.557	62.422	66.450
STANDARD DEV	22.183	23.619	25.091	23.733	23.959

	BATT	CIA	MIA	TIA	CSTA
N OF CASES	1027	1031	1031	1031	1015
MINIMUM	1.000	18.000	29.000	64.000	2.000
MAXIMUM	99.000	63.000	67.000	130.000	8.000
MEAN	58.445	58.833	61.101	119.934	7.611
STANDARD DEV	19.230	6.091	5.971	10.653	0.936

	MSTA	TSTA	TRSA	TWSA	SEX
N OF CASES	1031	1015	1017	1027	1031
MINIMUM	2.000	6.000	1.000	1.000	1.000
MAXIMUM	9.000	17.000	4.000	4.000	3.000
MEAN	8.314	15.952	3.857	3.733	1.538
STANDARD DEV	1.132	1.742	0.486	0.630	0.503

	RACE	GRADE	DIST	VERBAL	ATOTAL
N OF CASES	1025	1031	1031	1030	1030
MINIMUM	1.000	3.000	29.000	1.000	1.500
MAXIMUM	4.000	3.000	57.000	99.000	99.000
MEAN	1.392	3.000	42.455	57.376	61.913
STANDARD DEV	0.672	0.000	10.530	21.683	20.468

(table continues)

TEST A GRADE 5
TOTAL OBSERVATIONS: 1023

	RVOC	RCOMP	RTOT	SPELL	LMECH
N OF CASES	1022	1022	1022	680	1021
MINIMUM	1.000	1.000	1.000	1.000	1.000
MAXIMUM	99.000	99.000	99.000	85.000	99.000
MEAN	52.975	52.678	53.556	52.529	59.765
STANDARD DEV	22.629	20.600	21.070	16.229	21.301

	LEXP	LTOT	MCOMP	MAPP	MTOT
N OF CASES	1019	1019	1019	1017	1017
MINIMUM	1.000	1.000	1.000	1.000	1.000
MAXIMUM	99.000	99.000	99.000	99.000	99.000
MEAN	57.355	58.128	60.214	60.554	60.300
STANDARD DEV	22.470	21.316	22.518	22.528	22.247

	BATT	CIA	MIA	TIA	CSTA
N OF CASES	1015	1021	1022	1020	999
MINIMUM	1.000	22.000	32.000	54.000	2.000
MAXIMUM	99.000	101.000	100.000	201.000	11.000
MEAN	56.160	89.713	86.082	175.839	9.910
STANDARD DEV	18.197	11.723	12.872	22.417	1.668

	MSTA	TSTA	TRSA	TWSA	SEX
N OF CASES	1021	999	1010	1004	1023
MINIMUM	2.000	5.000	1.000	1.000	1.000
MAXIMUM	13.000	24.000	5.000	6.000	3.000
MEAN	11.000	20.999	4.410	5.466	1.521
STANDARD DEV	2.333	3.453	0.921	1.053	0.502

	RACE	GRADE	DIST	VERBAL	ATOTAL
N OF CASES	1022	1023	1023	1023	1017
MINIMUM	1.000	5.000	20.000	1.000	3.000
MAXIMUM	5.000	5.000	48.000	99.000	99.000
MEAN	1.532	5.000	37.551	55.112	57.756
STANDARD DEV	0.694	0.000	11.233	19.187	19.419

(table continues)

TEST A GRADE 8
TOTAL OBSERVATIONS: 1014

	RVOC	RCOMP	RTOT	SPELL	LMECH
N OF CASES	1012	1012	1011	1013	1013
MINIMUM	1.000	1.000	1.000	1.000	5.000
MAXIMUM	98.000	99.000	99.000	99.000	99.000
MEAN	54.620	57.170	56.397	57.943	60.799
STANDARD DEV	18.454	19.781	18.993	18.145	19.712

	LEXP	LTOT	MCOMP	MAPP	MTOT
N OF CASES	1012	1011	1012	1011	1011
MINIMUM	1.000	1.000	1.000	1.000	1.000
MAXIMUM	99.000	99.000	99.000	99.000	99.000
MEAN	61.008	61.471	59.764	58.194	59.178
STANDARD DEV	19.999	19.916	20.059	19.672	19.507

	BATT	CIA	MIA	TIA	CSTA
N OF CASES	1013	1007	1013	1007	996
MINIMUM	15.000	43.000	22.000	81.000	2.000
MAXIMUM	99.000	119.000	115.000	234.000	14.000
MEAN	56.971	103.697	95.111	198.862	12.288
STANDARD DEV	14.872	14.509	17.831	30.565	2.516

	MSTA	TSTA	TRSA	TWSA	SEX
N OF CASES	1013	996	999	1004	1014
MINIMUM	1.000	4.000	1.000	1.000	1.000
MAXIMUM	14.000	28.000	6.000	8.000	3.000
MEAN	11.598	23.972	5.314	6.928	1.503
STANDARD DEV	3.007	5.032	1.178	1.642	0.506

	RACE	GRADE	DIST	VERBAL	ATOTAL
N OF CASES	1011	1014	1014	1013	1010
MINIMUM	1.000	8.000	20.000	1.000	7.500
MAXIMUM	4.000	8.000	57.000	99.000	99.000
MEAN	1.489	8.000	46.051	58.589	58.861
STANDARD DEV	0.745	0.000	11.447	16.833	17.039

(table continues)

TEST A GRADE 10
TOTAL OBSERVATIONS: 1006

	RVOC	RCOMP	RTOT	SPELL	LMECH
N OF CASES	1006	1006	1006	579	1005
MINIMUM	2.000	1.000	1.000	10.000	2.000
MAXIMUM	99.000	99.000	99.000	99.000	99.000
MEAN	55.686	55.838	56.069	54.225	61.509
STANDARD DEV	21.225	20.757	20.946	17.185	22.783

	LEXP	LTOT	MCOMP	MAPP	MTOT
N OF CASES	1004	1003	1005	1005	1005
MINIMUM	1.000	1.000	1.000	2.000	1.000
MAXIMUM	99.000	99.000	99.000	99.000	99.000
MEAN	57.658	61.805	62.019	58.272	62.459
STANDARD DEV	19.988	21.952	22.012	22.382	22.302

	BATT	CIA	MIA	TIA	CSTA
N OF CASES	1003	993	995	991	971
MINIMUM	1.000	11.000	19.000	45.000	2.000
MAXIMUM	99.000	80.000	80.000	160.000	9.000
MEAN	58.502	70.155	69.994	140.206	8.127
STANDARD DEV	18.390	8.717	10.519	17.418	1.348

	MSTA	TSTA	TRSA	TWSA	COMM
N OF CASES	988	967	980	980	996
MINIMUM	1.000	6.000	1.000	1.000	17.000
MAXIMUM	8.000	17.000	4.000	5.000	75.000
MEAN	7.160	15.344	3.632	4.458	69.537
STANDARD DEV	1.411	2.277	0.697	0.893	6.698

	MATH	TCMTH	SEX	RACE	GRADE
N OF CASES	998	1001	1006	1005	1006
MINIMUM	10.000	32.000	1.000	1.000	10.000
MAXIMUM	75.000	150.000	2.000	5.000	10.000
MEAN	62.638	131.640	1.516	1.446	10.000
STANDARD DEV	10.817	17.527	0.500	0.684	0.000

	DIST	VERBAL	ATOTAL
N OF CASES	1006	1006	1005
MINIMUM	20.000	2.000	2.000
MAXIMUM	48.000	99.000	99.000
MEAN	35.816	58.787	60.640
STANDARD DEV	10.965	19.547	19.796

(table continues)

TEST B GRADE 3
TOTAL OBSERVATIONS: 1010

	RVOC	RCOMP	RTOT	LMECH	LEXP
N OF CASES	1010	1008	1010	1010	1010
MINIMUM	1.007	1.007	1.000	1.000	1.007
MAXIMUM	99.000	99.000	99.000	99.000	99.000
MEAN	54.908	55.105	54.616	55.905	57.255
STANDARD DEV	20.535	18.643	18.983	20.103	19.187

	LTOT	SPELL	MCOMP	MCA	MTOT
N OF CASES	1010	550	1010	1010	1010
MINIMUM	1.000	1.000	1.000	1.007	1.000
MAXIMUM	99.000	99.000	98.993	99.000	99.000
MEAN	57.386	55.164	58.310	59.169	59.244
STANDARD DEV	19.537	18.826	20.802	18.778	19.155

	CIA	MIA	TIA	CSTA	MSTA
N OF CASES	1010	1010	1010	1004	1010
MINIMUM	17.000	28.000	47.000	2.000	2.000
MAXIMUM	63.000	67.000	130.000	8.000	9.000
MEAN	59.430	61.431	120.860	7.663	8.388
STANDARD DEV	5.254	5.570	9.398	0.810	1.047

	TSTA	TRSA	TWSA	SEX	RACE
N OF CASES	1004	1007	1007	1010	1007
MINIMUM	6.000	1.000	1.000	1.000	1.000
MAXIMUM	17.000	4.000	4.000	3.000	5.000
MEAN	16.062	3.857	3.796	1.504	1.304
STANDARD DEV	1.549	0.452	0.536	0.502	0.547

	GRADE	DIST	VERBAL	BTOTAL
N OF CASES	1010	1010	1010	1010
MINIMUM	3.000	1.000	3.878	9.167
MAXIMUM	3.000	43.000	98.993	97.500
MEAN	3.000	34.014	55.931	57.588
STANDARD DEV	0.000	16.571	18.105	16.911

(table continues)

TEST B GRADE 5
TOTAL OBSERVATIONS: 1022

	RVOC	RCOMP	RTOT	LMECH	LEXP
N OF CASES	1022	782	1021	1021	1022
MINIMUM	1.000	3.000	1.007	1.000	1.000
MAXIMUM	99.000	99.000	99.000	99.000	93.252
MEAN	55.389	53.097	55.187	55.939	53.872
STANDARD DEV	18.152	15.927	17.558	17.602	16.264

	LTOT	SPELL	MCOMP	MCA	MTOT
N OF CASES	1021	574	1022	1021	1021
MINIMUM	1.007	15.000	1.000	1.000	6.748
MAXIMUM	99.000	93.000	99.000	99.000	99.000
MEAN	54.846	52.063	59.408	57.394	59.113
STANDARD DEV	16.628	15.392	17.710	16.950	17.467

	CIA	MIA	TIA	CSTA	MSTA
N OF CASES	1022	1021	1021	1008	1021
MINIMUM	23.000	30.000	61.000	3.000	2.000
MAXIMUM	101.000	100.000	201.000	11.000	13.000
MEAN	92.891	89.852	182.742	10.415	11.697
STANDARD DEV	9.607	10.808	18.755	1.194	1.858

	TSTA	TRSA	TWSA	SEX	RACE
N OF CASES	1007	1014	1013	1022	1020
MINIMUM	8.000	1.000	1.000	1.000	1.000
MAXIMUM	24.000	5.000	6.000	3.000	5.000
MEAN	22.184	4.679	5.710	1.525	1.305
STANDARD DEV	2.555	0.696	0.767	0.502	0.603

	GRADE	DIST	VERBAL	BTOTAL
N OF CASES	1022	1022	1022	1021
MINIMUM	5.000	1.000	3.878	10.567
MAXIMUM	5.000	43.000	99.000	97.558
MEAN	5.000	32.575	54.899	57.007
STANDARD DEV	0.000	17.505	15.560	15.386

(table continues)

TEST B GRADE 8
TOTAL OBSERVATIONS: 1004

	RVOC	RCOMP	RTOT	LMECH	LEXP
N OF CASES	975	975	975	985	987
MINIMUM	1.000	6.000	1.000	1.007	1.007
MAXIMUM	99.000	99.000	99.000	99.000	98.993
MEAN	55.508	55.676	55.695	55.526	56.015
STANDARD DEV	17.222	16.813	16.535	17.228	17.995

	LTOT	SPELL	MCOMP	MCA	MTOT
N OF CASES	985	495	981	982	981
MINIMUM	6.748	9.000	1.000	1.000	1.000
MAXIMUM	99.000	99.000	99.000	99.000	99.000
MEAN	56.343	50.012	56.820	55.902	56.829
STANDARD DEV	17.291	16.144	18.465	17.503	17.981

	CIA	MIA	TIA	CSTA	MSTA
N OF CASES	1000	1000	996	994	1000
MINIMUM	24.000	31.000	55.000	2.000	1.000
MAXIMUM	119.000	115.000	234.000	14.000	14.000
MEAN	107.890	100.498	208.469	13.010	12.406
STANDARD DEV	10.703	13.872	22.453	1.748	2.414

	TSTA	TRSA	TWSA	SEX	RACE
N OF CASES	990	997	996	1004	1001
MINIMUM	7.000	1.000	1.000	1.000	1.000
MAXIMUM	28.000	6.000	8.000	3.000	5.000
MEAN	25.465	5.651	7.343	1.516	1.281
STANDARD DEV	3.656	0.811	1.153	0.502	0.576

	GRADE	DIST	VERBAL	BTOTAL
N OF CASES	1004	1004	1000	977
MINIMUM	8.000	1.000	11.385	10.445
MAXIMUM	8.000	43.000	98.993	98.993
MEAN	8.000	32.398	55.622	56.309
STANDARD DEV	0.000	17.660	15.650	15.772

(table continues)

TEST B GRADE 10
TOTAL OBSERVATIONS: 753

	RVOC	RCOMP	RTOT	LMECH	LEXP
N OF CASES	742	744	742	745	742
MINIMUM	1.000	1.000	1.000	5.000	8.000
MAXIMUM	99.000	99.000	99.000	99.000	99.000
MEAN	51.910	53.642	52.650	53.122	52.443
STANDARD DEV	19.600	19.130	19.475	18.584	18.331

	LTOT	SPELL	MCOMP	MCA	MTOT
N OF CASES	734	491	742	742	742
MINIMUM	9.000	1.000	1.000	1.000	1.000
MAXIMUM	99.000	98.000	99.000	99.000	99.000
MEAN	52.940	49.493	55.612	54.126	55.091
STANDARD DEV	18.308	19.071	18.708	19.604	19.166

	CIA	MIA	TIA	CSTA	MSTA
N OF CASES	745	747	745	737	746
MINIMUM	21.000	25.000	54.000	2.000	1.000
MAXIMUM	80.000	80.000	160.000	9.000	8.000
MEAN	70.258	71.019	141.358	8.037	7.255
STANDARD DEV	8.309	8.809	15.241	1.448	1.275

	TSTA	TRSA	TWSA	COMM	MATH
N OF CASES	737	742	739	745	746
MINIMUM	4.000	1.000	1.000	13.000	21.000
MAXIMUM	17.000	4.000	5.000	75.000	75.000
MEAN	15.319	3.571	4.442	68.103	62.055
STANDARD DEV	2.282	0.738	0.923	8.358	10.363

	TCMTH	SEX	RACE	GRADE	DIST
N OF CASES	746	753	753	753	753
MINIMUM	40.000	1.000	1.000	10.000	1.000
MAXIMUM	150.000	2.000	5.000	10.000	43.000
MEAN	130.067	1.526	1.328	10.000	31.349
STANDARD DEV	17.487	0.500	0.629	0.000	18.079

	VERBAL	BTOTAL
N OF CASES	745	734
MINIMUM	7.000	8.500
MAXIMUM	96.122	96.647
MEAN	52.341	53.729
STANDARD DEV	17.563	17.422

(table continues)

TEST C GRADE 3
TOTAL OBSERVATIONS: 1002

	RCOMP	RTOT	LANG	LISTEN	MCONC
N OF CASES	1000	1000	330	1001	1002
MINIMUM	1.000	1.007	1.007	1.007	6.700
MAXIMUM	99.000	99.000	98.993	99.000	99.000
MEAN	54.126	51.953	47.890	54.235	57.547
STANDARD DEV	18.919	19.160	19.419	19.270	18.500

	MCOMP	MAPP	TMATH	CIA	MIA
N OF CASES	1002	1002	1002	1002	999
MINIMUM	1.007	1.007	1.007	21.000	15.000
MAXIMUM	99.000	99.000	99.000	63.000	67.000
MEAN	59.329	52.459	57.722	58.263	60.498
STANDARD DEV	20.434	19.348	18.754	6.253	5.978

	TIA	CSTA	MSTA	TSTA	TRSA
N OF CASES	1002	1002	999	1002	993
MINIMUM	57.000	1.000	1.000	3.000	1.000
MAXIMUM	130.000	8.000	9.000	17.000	4.000
MEAN	118.581	7.434	8.212	15.622	3.803
STANDARD DEV	11.121	1.173	1.185	2.084	0.544

	TWSA	SEX	RACE	GRADE	VERBAL
N OF CASES	1000	1002	1002	1002	1002
MINIMUM	1.000	1.000	0.000	3.000	4.135
MAXIMUM	4.000	3.000	5.000	3.000	99.000
MEAN	3.673	1.492	1.624	3.000	52.901
STANDARD DEV	0.670	0.512	0.818	0.000	17.159

	CTOTAL
N OF CASES	1002
MINIMUM	13.082
MAXIMUM	99.000
MEAN	55.311
STANDARD DEV	16.651

(table continues)

TEST C GRADE 5
TOTAL OBSERVATIONS: 1003

	RCOMP	RTOT	LANG	LISTEN	MCONC
N OF CASES	1003	1003	998	994	1003
MINIMUM	1.007	1.007	1.000	1.000	1.007
MAXIMUM	99.000	99.000	99.000	99.000	99.000
MEAN	49.185	51.206	51.805	52.004	57.766
STANDARD DEV	17.915	16.178	18.420	19.738	18.527

	MCOMP	MAPP	TMATH	CIA	MIA
N OF CASES	1003	1003	1003	1001	1001
MINIMUM	6.700	1.007	1.007	20.000	32.000
MAXIMUM	99.000	99.000	99.000	101.000	100.000
MEAN	53.324	56.340	56.478	89.925	85.476
STANDARD DEV	17.411	18.788	18.277	11.663	13.361

	TIA	CSTA	MSTA	TSTA	TRSA
N OF CASES	1002	997	1001	996	995
MINIMUM	64.000	1.000	1.000	3.000	1.000
MAXIMUM	201.000	11.000	13.000	24.000	5.000
MEAN	175.226	9.831	10.928	20.785	4.450
STANDARD DEV	23.215	1.871	2.367	3.771	0.903

	TWSA	SEX	RACE	GRADE	VERBAL
N OF CASES	987	1003	1003	1003	1003
MINIMUM	1.000	1.000	0.000	5.000	6.978
MAXIMUM	6.000	3.000	5.000	5.000	98.993
MEAN	5.445	1.538	1.640	5.000	51.658
STANDARD DEV	1.096	0.501	0.768	0.000	15.970

	CTOTAL
N OF CASES	1003
MINIMUM	10.054
MAXIMUM	98.993
MEAN	54.068
STANDARD DEV	16.246

(table continues)

TEST C GRADE 8
TOTAL OBSERVATIONS: 1000

	RCOMP	RTOT	LANG	LISTEN	MCONC
N OF CASES	999	0	756	705	1000
MINIMUM	1.007	.	1.000	1.007	1.007
MAXIMUM	99.000	.	99.000	99.000	99.000
MEAN	46.802	.	51.087	47.386	52.966
STANDARD DEV	17.011	.	16.843	16.375	16.024

	MCOMP	MAPP	TMATH	CIA	MIA
N OF CASES	1000	999	999	992	998
MINIMUM	1.000	1.000	1.007	18.000	18.000
MAXIMUM	99.000	99.000	99.000	119.000	115.000
MEAN	50.365	51.703	52.141	101.618	94.567
STANDARD DEV	16.435	17.579	16.518	15.287	16.769

	TIA	CSTA	MSTA	TSTA	TRSA
N OF CASES	990	967	994	964	980
MINIMUM	52.000	2.000	1.000	3.000	1.000
MAXIMUM	234.000	14.000	14.000	28.000	6.000
MEAN	196.374	12.089	11.555	23.817	5.226
STANDARD DEV	29.259	2.442	2.845	4.510	1.188

	TWSA	SEX	RACE	GRADE	VERBAL
N OF CASES	976	1000	1000	1000	1000
MINIMUM	1.000	1.000	0.000	8.000	1.007
MAXIMUM	8.000	3.000	4.000	8.000	98.993
MEAN	6.792	1.536	1.807	8.000	47.620
STANDARD DEV	1.630	0.503	0.780	0.000	15.679

	CTOTAL
N OF CASES	999
MINIMUM	1.964
MAXIMUM	96.122
MEAN	49.894
STANDARD DEV	15.002

(table continues)

TEST C GRADE 10
TOTAL OBSERVATIONS: 287

	RCOMP	RTOT	LANG	LISTEN	MCONC
N OF CASES	287	0	287	287	287
MINIMUM	1.007	.	1.007	1.007	1.007
MAXIMUM	98 993	.	98.993	98.993	93.252
MEAN	50.657	.	49.078	44.921	50.011
STANDARD DEV	17.804	.	15.930	16.852	18.161

	MCOMP	MAPP	TMATH	CIA	MIA
N OF CASES	287	287	287	282	283
MINIMUM	1.007	1.007	1.007	14.000	8.000
MAXIMUM	93.252	98.993	98.993	80.000	80.000
MEAN	51.985	48.206	49.770	64.475	63.802
STANDARD DEV	17.488	18.211	16.735	11.093	13.924

	TIA	CSTA	MSTA	TSTA	TRSA
N OF CASES	282	263	278	262	270
MINIMUM	42.000	2.000	1.000	5.000	1.000
MAXIMUM	160.000	9.000	8.000	17.000	4.000
MEAN	128.475	7.399	6.547	14.229	3.311
STANDARD DEV	22.867	1.617	1.843	2.552	0.875

	TWSA	COMM	MATH	TCMTH	SEX
N OF CASES	269	273	274	274	287
MINIMUM	1.000	19.000	22.000	33.000	1.000
MAXIMUM	5.000	75.000	75.000	150.000	2.000
MEAN	3.985	65.864	56.420	122.044	1.561
STANDARD DEV	1.156	8.660	11.542	18.973	0.497

	RACE	GRADE	VERBAL	CTOTAL
N OF CASES	287	287	287	287
MINIMUM	1.000	10.000	6.962	3.985
MAXIMUM	3.000	10.000	90.910	94.952
MEAN	1.916	10.000	48.219	48.994
STANDARD DEV	0.290	0.000	14.23[illegible]	14.429

Table A-2

CORRELATIONS AMONG NRT AND SSAT SCORES FOR TEST A

PEARSON CORRELATION MATRIX

	NRT			SSAT		
	VERBAL	MATH	TOTAL	COMM	MATH	TOTAL
GR3						
NRTVERBAL	1.000					
NRTMATH	0.608	1.000				
TOTAL	0.885	0.907	1.000			
SSATCOMM	0.489	0.408	0.499	1.000		
SSATMATH	0.436	0.492	0.519	0.452	1.000	
TOTAL	0.525	0.521	0.585	0.824	0.878	1.000
GR5						
NRTVERBAL	1.000					
NRTMATH	0.765	1.000				
TOTAL	0.930	0.949	1.000			
SSATCOMM	0.557	0.463	0.538	1.000		
SSATMATH	0.570	0.657	0.654	0.569	1.000	
TOTAL	0.625	0.643	0.674	0.849	0.918	1.000
GR8						
NRTVERB	1.000					
NRTMATH	0.761	1.000				
TOTAL	0.929	0.948	1.000			
SSATCOMM	0.631	0.569	0.637	1.000		
SSATMATH	0.597	0.686	0.687	0.714	1.000	
TOTAL	0.656	0.678	0.712	0.914	0.937	1.000
GR10						
NRTVERBAL	1.000					
NRTMATH	0.791	1.000				
TOTAL	0.939	0.953	1.000			
SSATCOMM	0.564	0.542	0.584	1.000		
SSATMATH	0.458	0.551	0.537	0.442	1.000	
TOTAL	0.584	0.633	0.646	0.849	0.850	1.000
GR10 SSAT-II						
NRTVERBAL	1.000					
NRTMATH	0.791	1.000				
TOTAL	0.939	0.953	1.000			
SSATCOMM	0.571	0.559	0.597	1.000		
SSATMATH	0.604	0.704	0.695	0.662	1.000	
TOTAL	0.616	0.671	0.682	0.849	0.92	1.000

Table A-3

CORRELATIONS AMONG NRT AND SSAT SCORES FOR TEST B

PEARSON CORRELATION MATRIX

	NRT			SSAT		
	VERBAL	MATH	TOTAL	COMM	MATH	TOTAL
GR3						
NRTVERBAL	1.000					
NRTMATH	0.648	1.000				
TOTAL	0.902	0.913	1.000			
SSATCOMM	0.498	0.383	0.484	1.000		
SSATMATH	0.419	0.509	0.512	0.418	1.000	
TOTAL	0.529	0.533	0.586	0.799	0.880	1.000
GR5						
NRTVERBAL	1.000					
NRTMATH	0.735	1.000				
TOTAL	0.923	0.939	1.000			
SSATCOMM	0.492	0.401	0.477	1.000		
SSATCOMM	0.536	0.617	0.621	0.525	1.000	
TOTAL	0.568	0.592	0.625	0.820	0.917	1.000
GR8						
NRTVERBAL	1.000					
NRTMATH	0.756	1.000				
TOTAL	0.928	0.946	1.000			
SSATCOMM	0.558	0.485	0.556	1.000		
SSATMATH	0.512	0.646	0.624	0.596	1.000	
TOTAL	0.589	0.637	0.658	0.858	0.924	1.000
GR10						
NRTVERBAL	1.000					
NRTMATH	0.795	1.000				
TOTAL	0.943	0.952	1.000			
SSATCOMM	0.661	0.558	0.643	1.000		
SSATMATH	0.432	0.555	0.524	0.446	1.000	
TOTAL	0.644	0.645	0.681	0.876	0.823	1.000
GR10 SSAT-II						
NRTVERBAL	1.000					
NRTMATH	0.795	1.000				
TOTAL	0.943	0.952	1.000			
SSATCOMM	0.578	0.599	0.621	1.000		
SSATMATH	0.642	0.736	0.730	0.686	1.000	
TOTAL	0.661	0.724	0.733	0.899	0.930	1.000

Table A-4

CORRELATIONS AMONG NRT AND SSAT SCORES FOR TEST C

	PEARSON CORRELATION MATRIX					
	NRT			SSAT		
	VERBAL	MATH	TOTAL	COMM	MATH	TOTAL
GR3						
NRTVERBAL	1.000					
NRTMATH	0.719	1.000				
TOTAL	0.920	0.934	1.000			
SSATCOMM	0.440	0.430	0.469	1.000		
SSATMATH	0.411	0.522	0.506	0.501	1.000	
TOTAL	0.477	0.542	0.5[illegible]1	0.842	0.868	1.000
GR5						
NRTVERBAL	1.000					
NRTMATH	0.799	1.000				
TOTAL	0.941	0.955	1.000			
SSATCOMM	0.543	0.512	0.555	1.000		
SSATMATH	0.570	0.679	0.662	0.610	1.000	
TOTAL	0.617	0.670	0.681	0.872	0.919	1.000
GR8						
NRTVERBAL	1.000					
NRTMATH	0.738	1.000				
TOTAL	0.928	0.936	1.000			
SSATCOMM	0.591	0.543	0.608	1.000		
SSATMATH	0.498	0.652	0.618	0.573	1.000	
TOTAL	0.594	0.670	0.680	0.877	0.896	1.000
GR10						
NRTVERBAL	1.000					
NRTMATH	0.735	1.000				
TOTAL	0.919	0.942	1.000			
SSATCOMM	0.586	0.480	0.580	1.000		
SSATMATH	0.487	0.676	0.634	0.381	1.000	
TOTAL	0.562	0.631	0.656	0.843	0.819	1.000
GR10 SSAT-II						
NRTVERBAL	1.000					
NRTMATH	0.735	1.000				
TOTAL	0.919	0.942	1.000			
SSATCOMM	0.581	0.532	0.603	1.000		
SSATMATH	0.555	0.676	0.674	0.631	1.000	
TOTAL	0.600	0.662	0.688	0.872	0.91	1.000

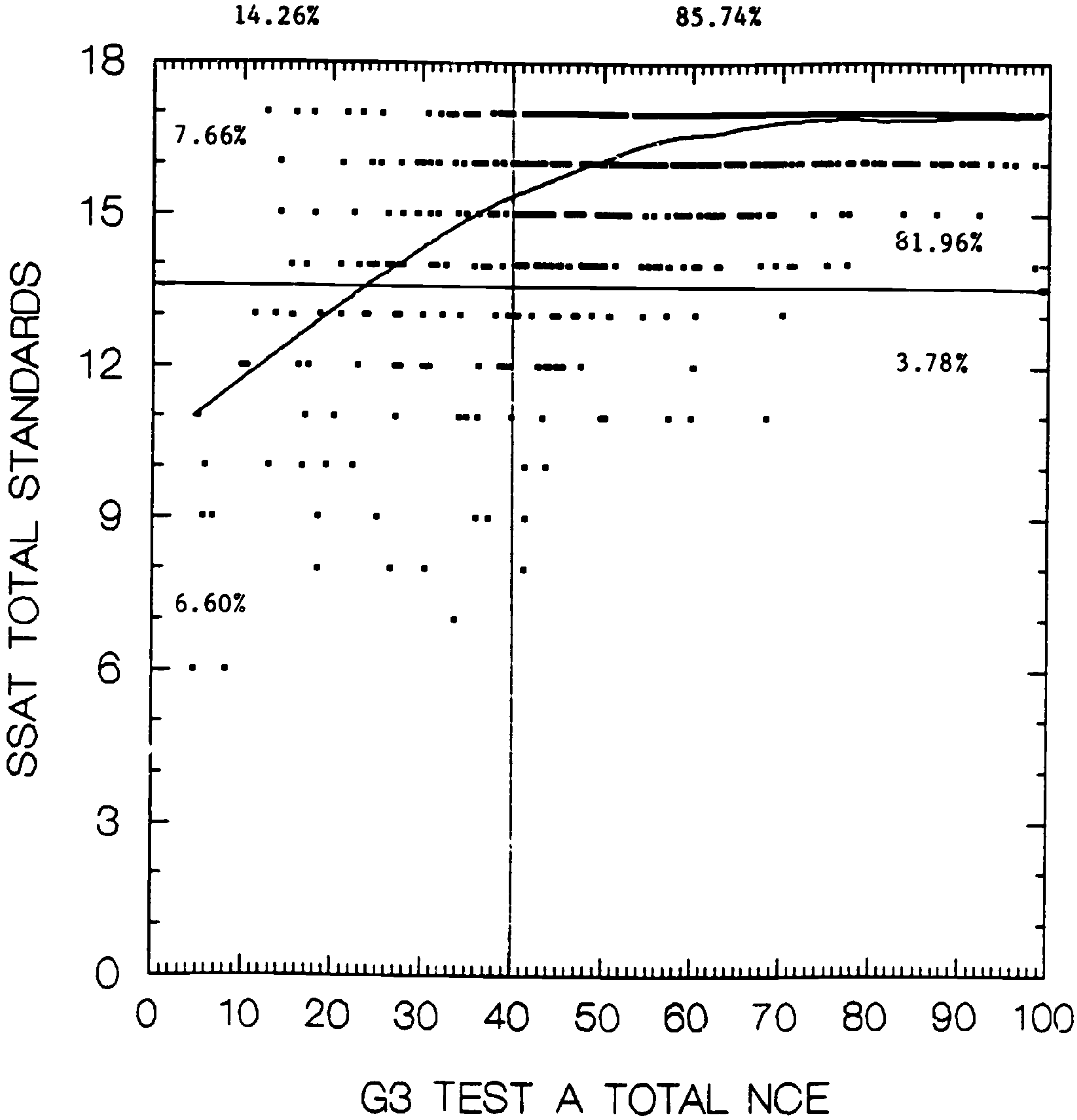


Figure A-1. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.

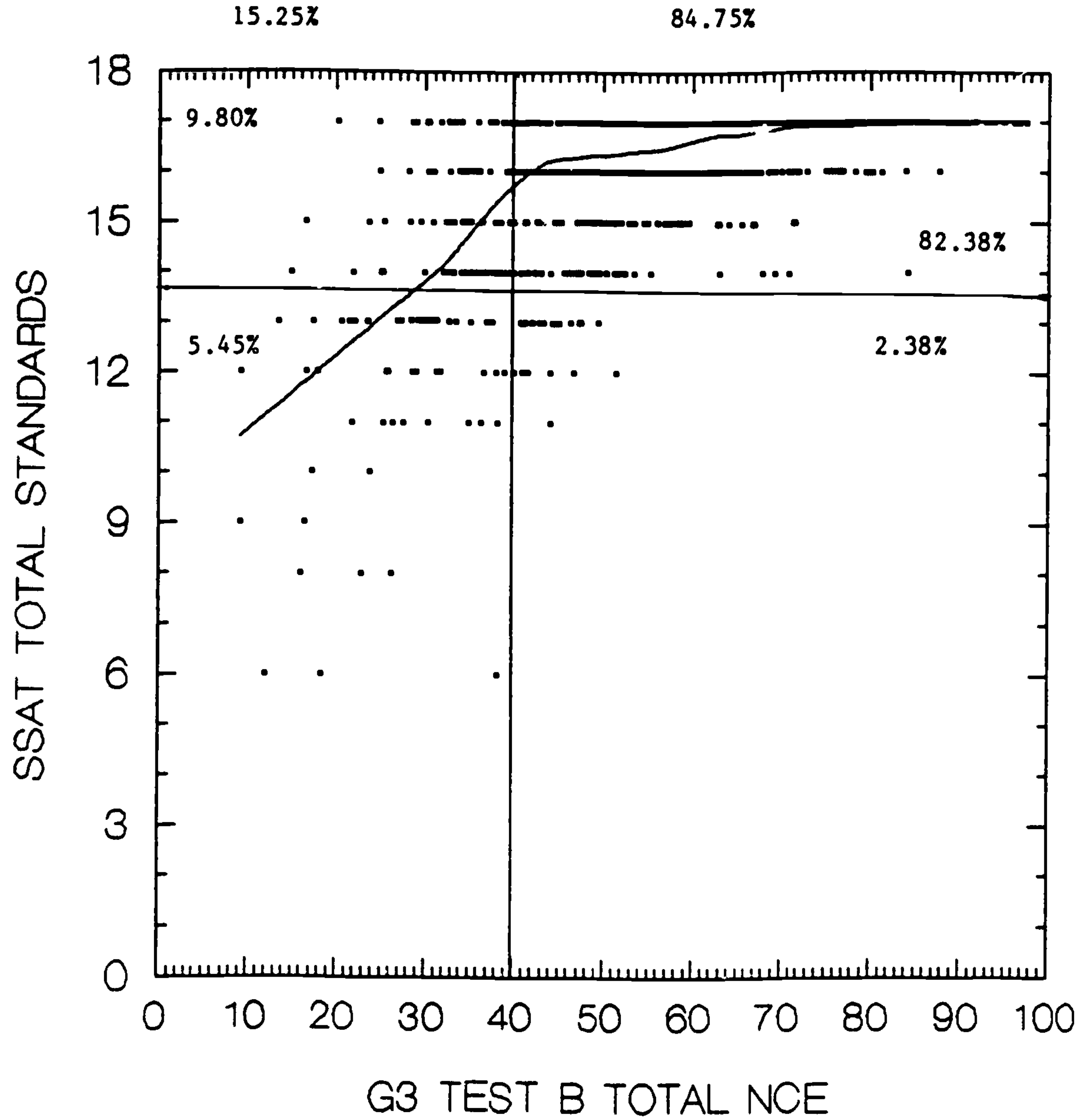


Figure A-2. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.

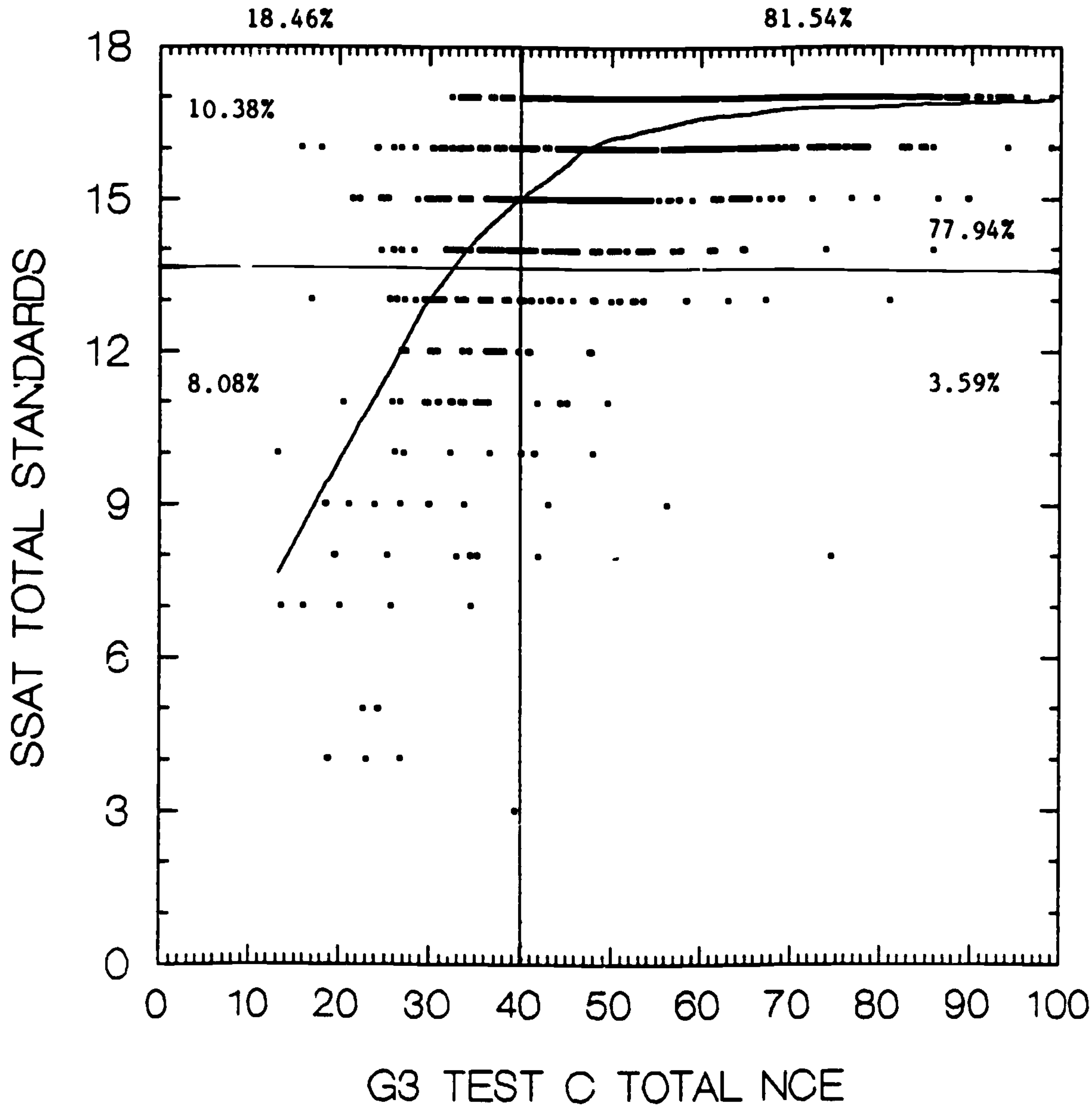


Figure A-3. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.

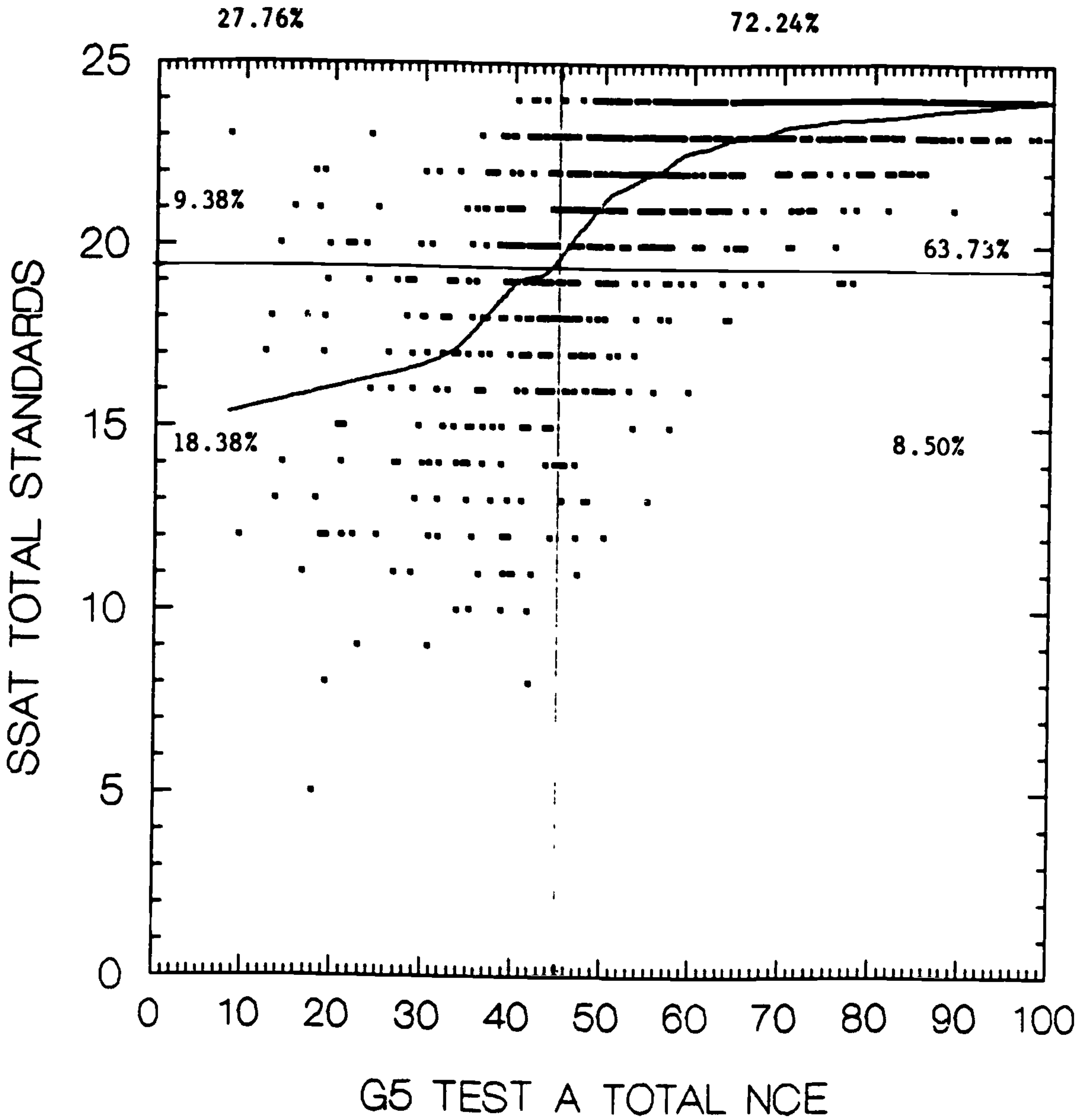


Figure A-4. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.

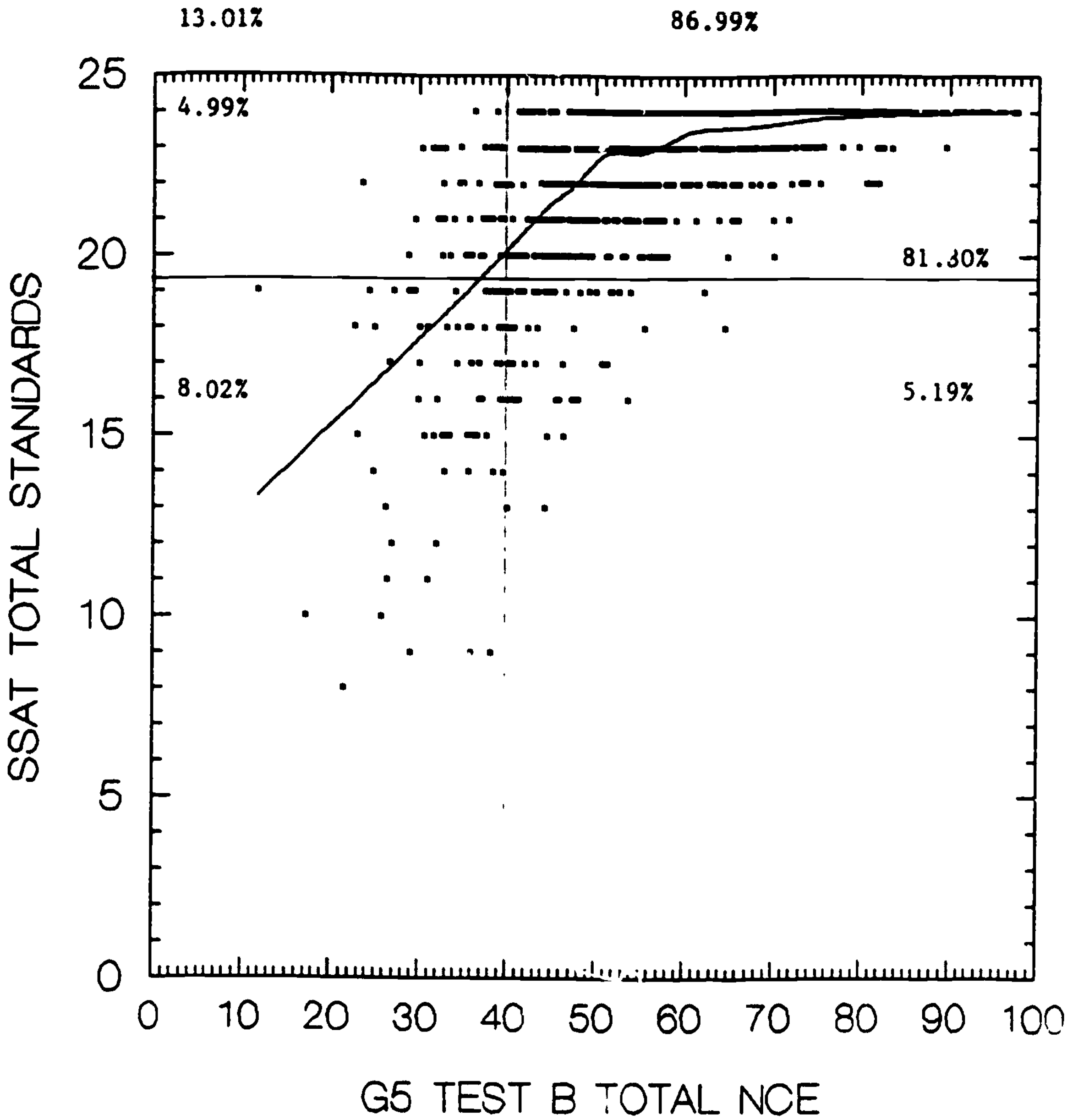


Figure A-5. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.

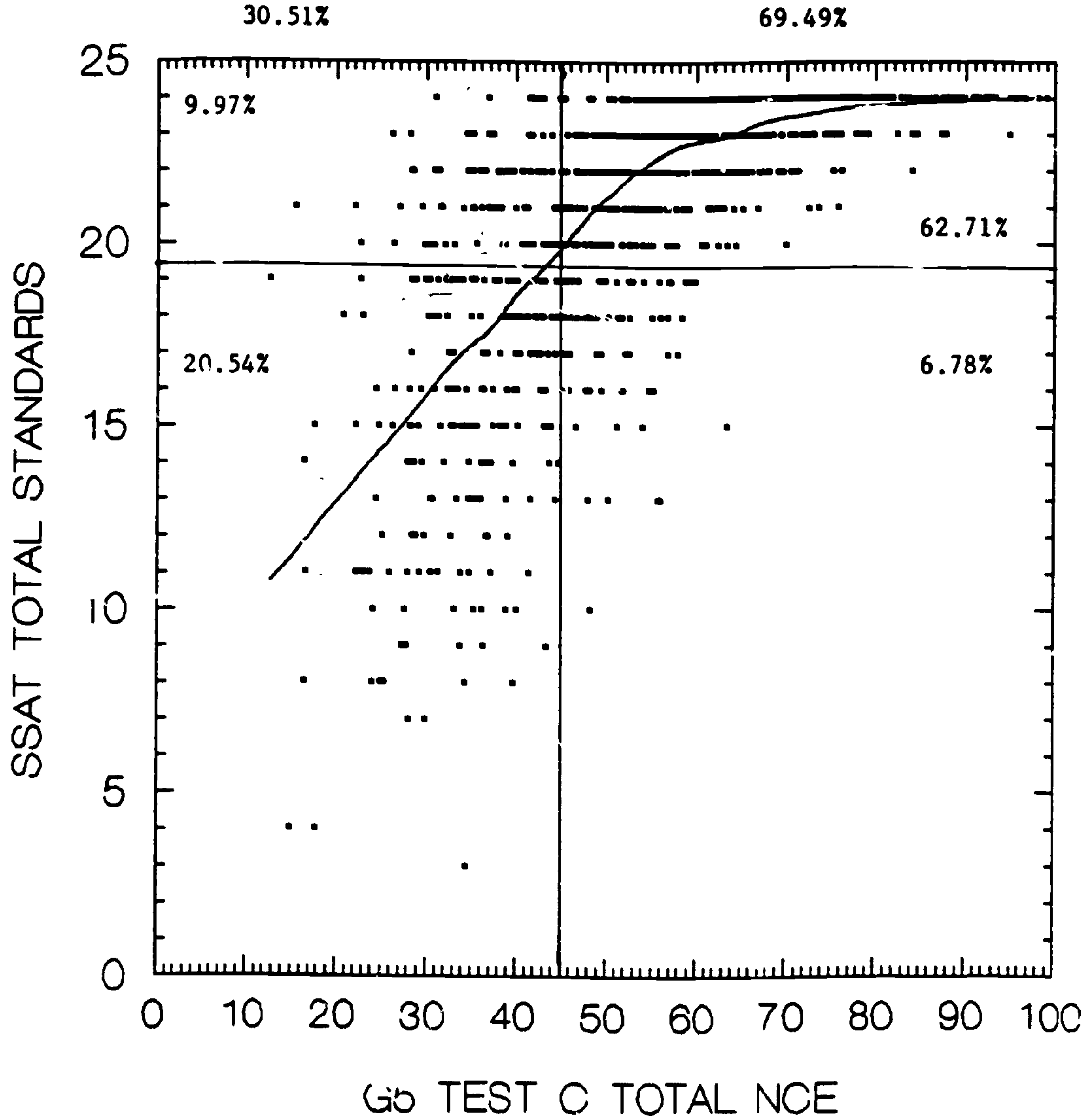


Figure A-6. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.

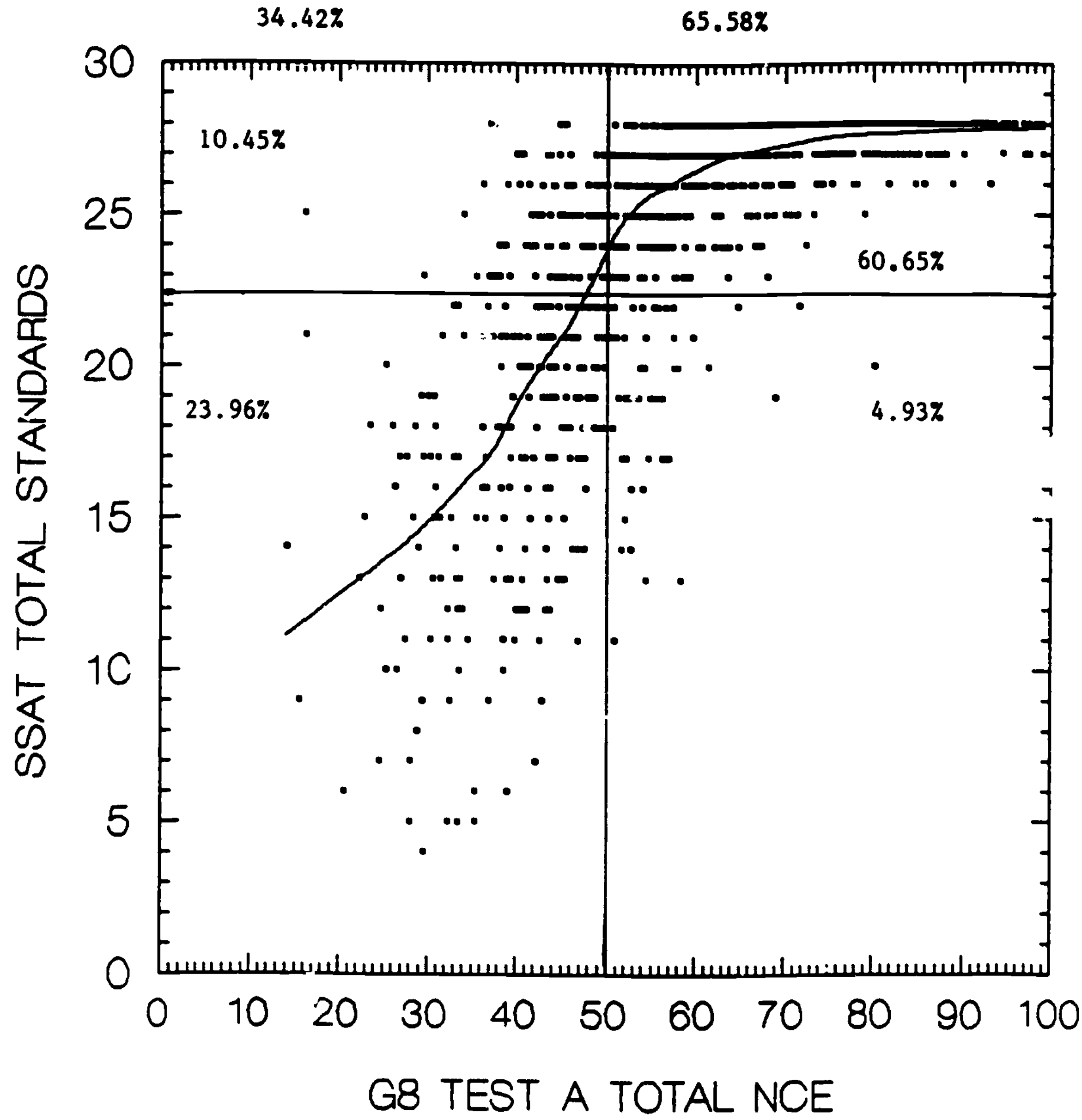


Figure A-7. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.

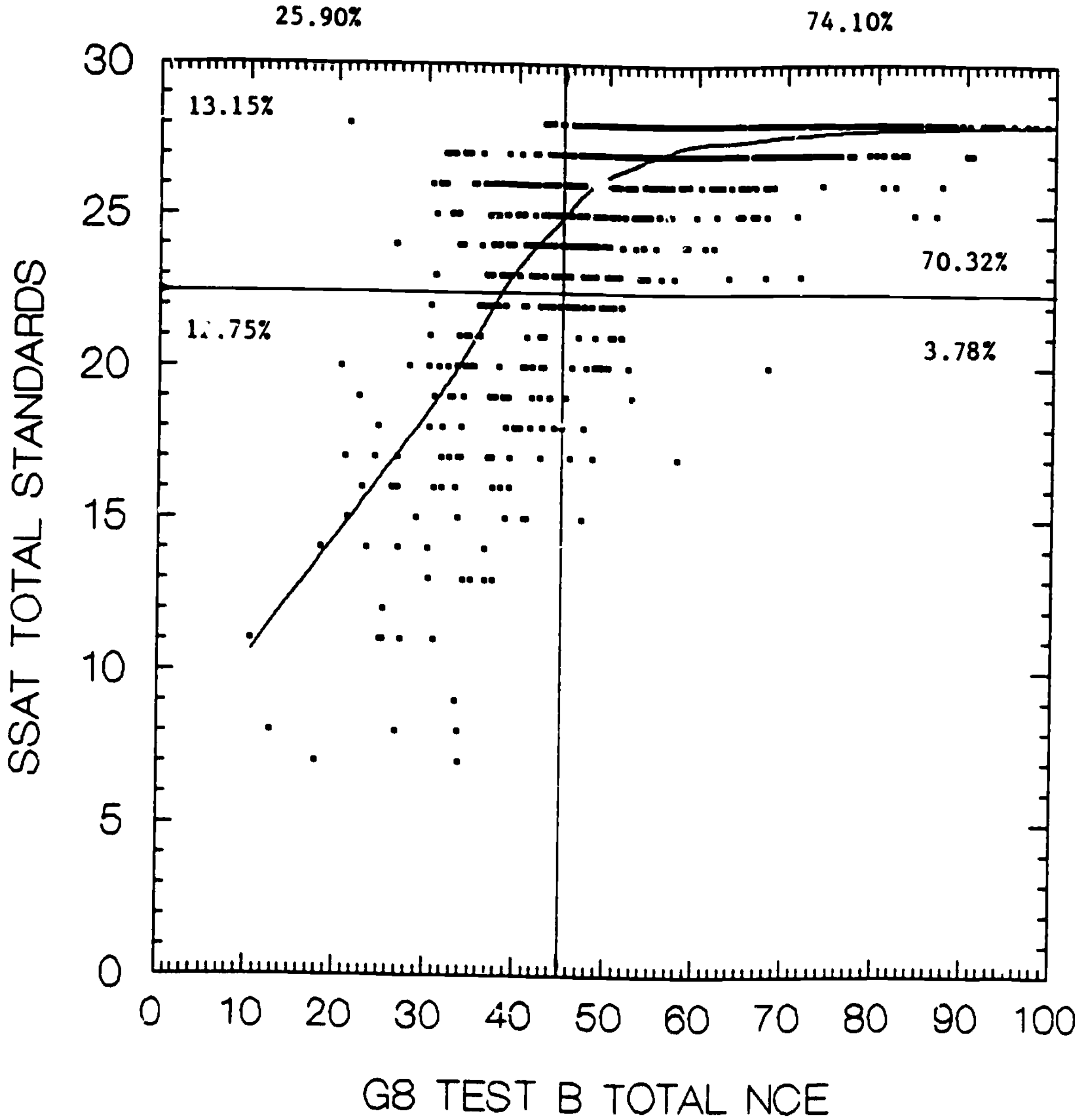


Figure A-8. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.

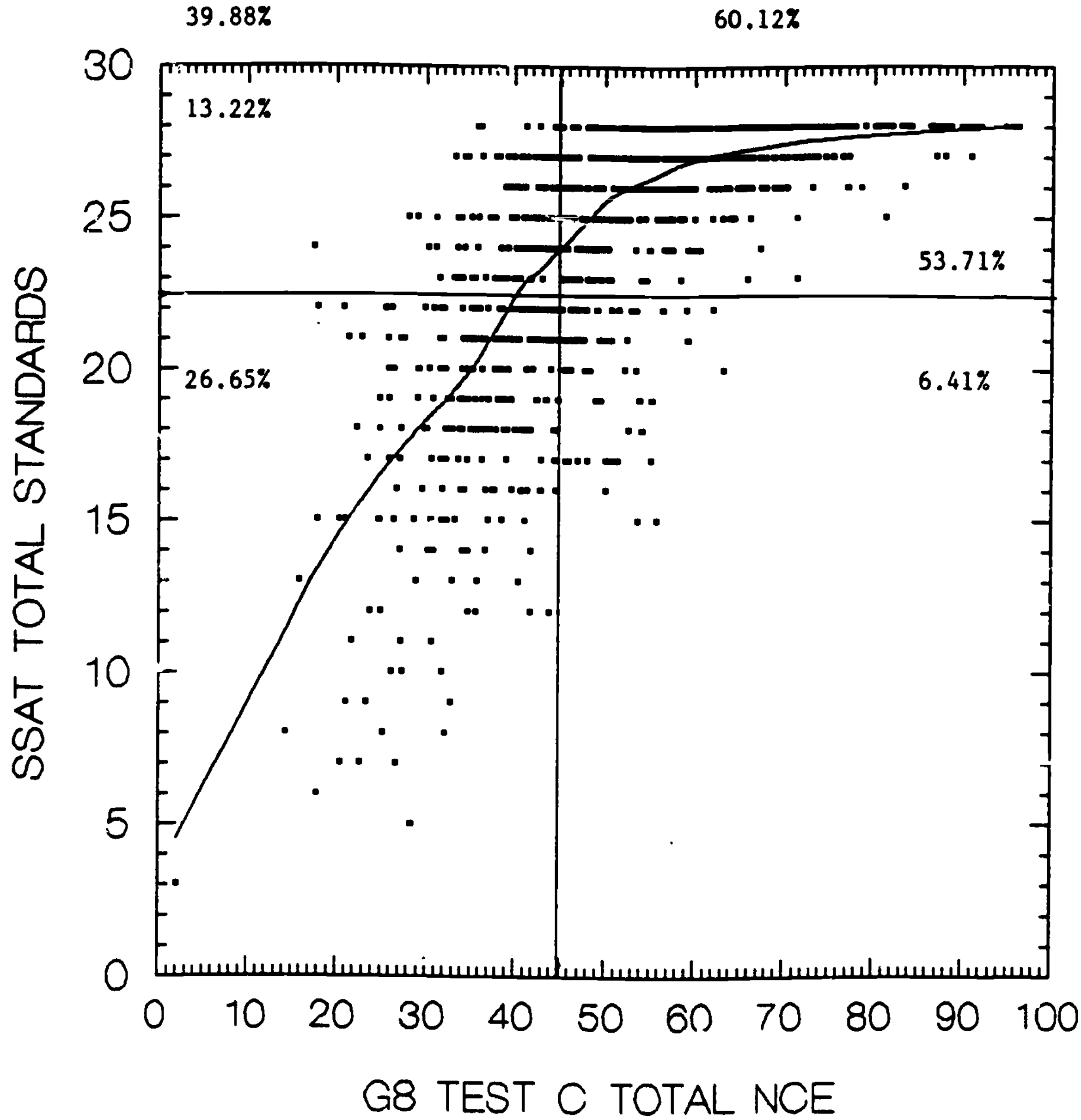


Figure A-9. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.

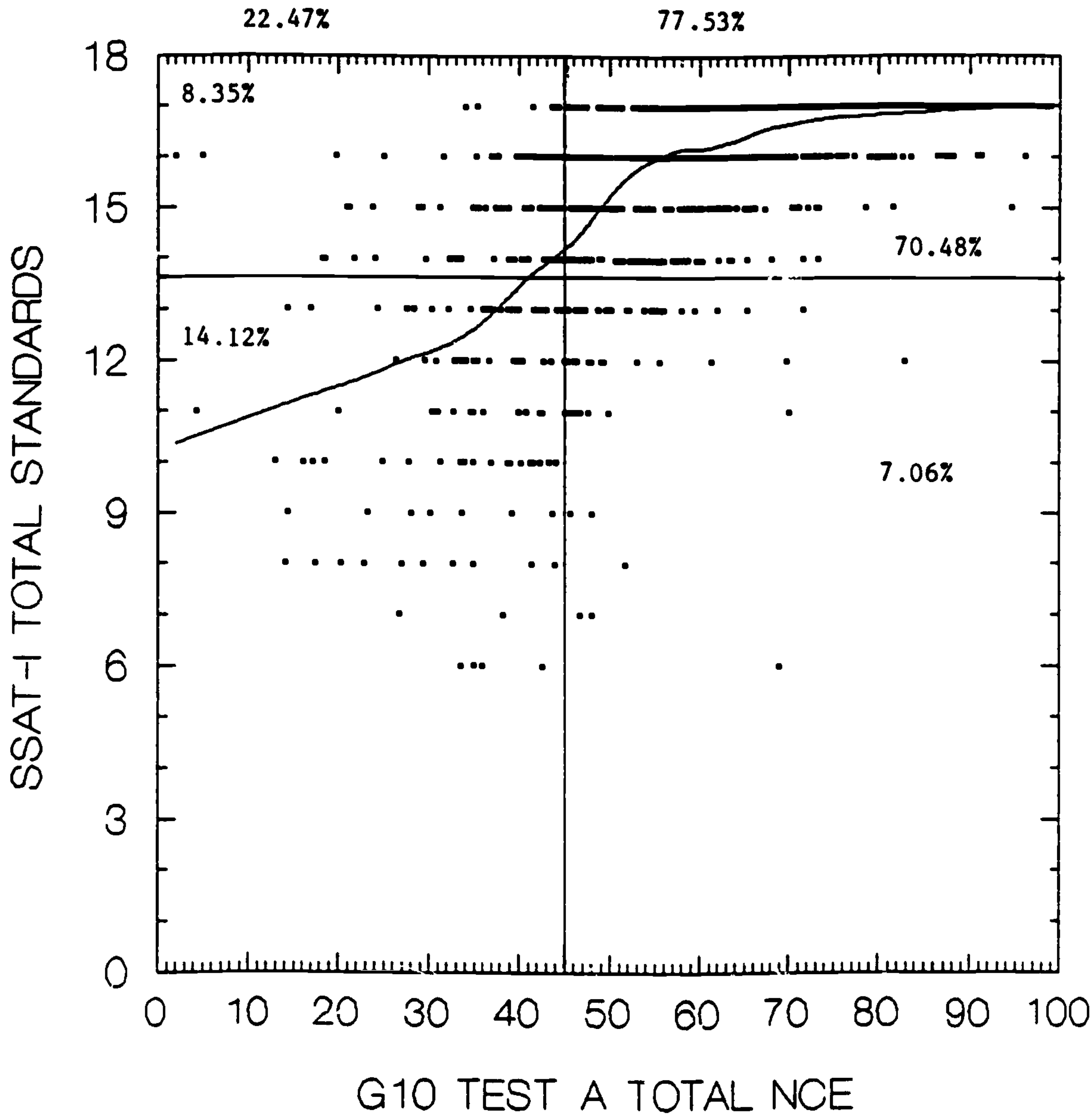


Figure A-10. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.

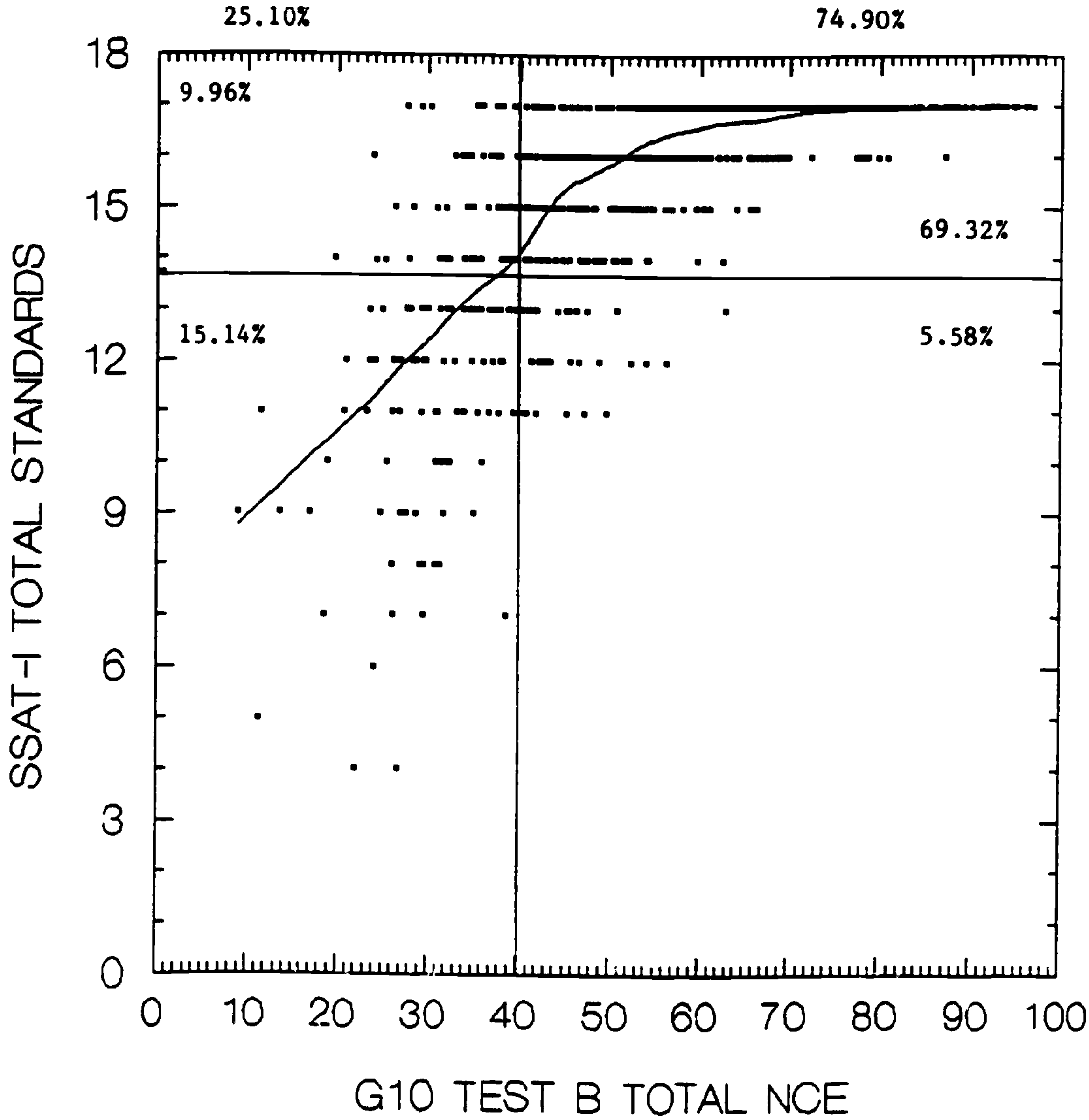


Figure A-11. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.

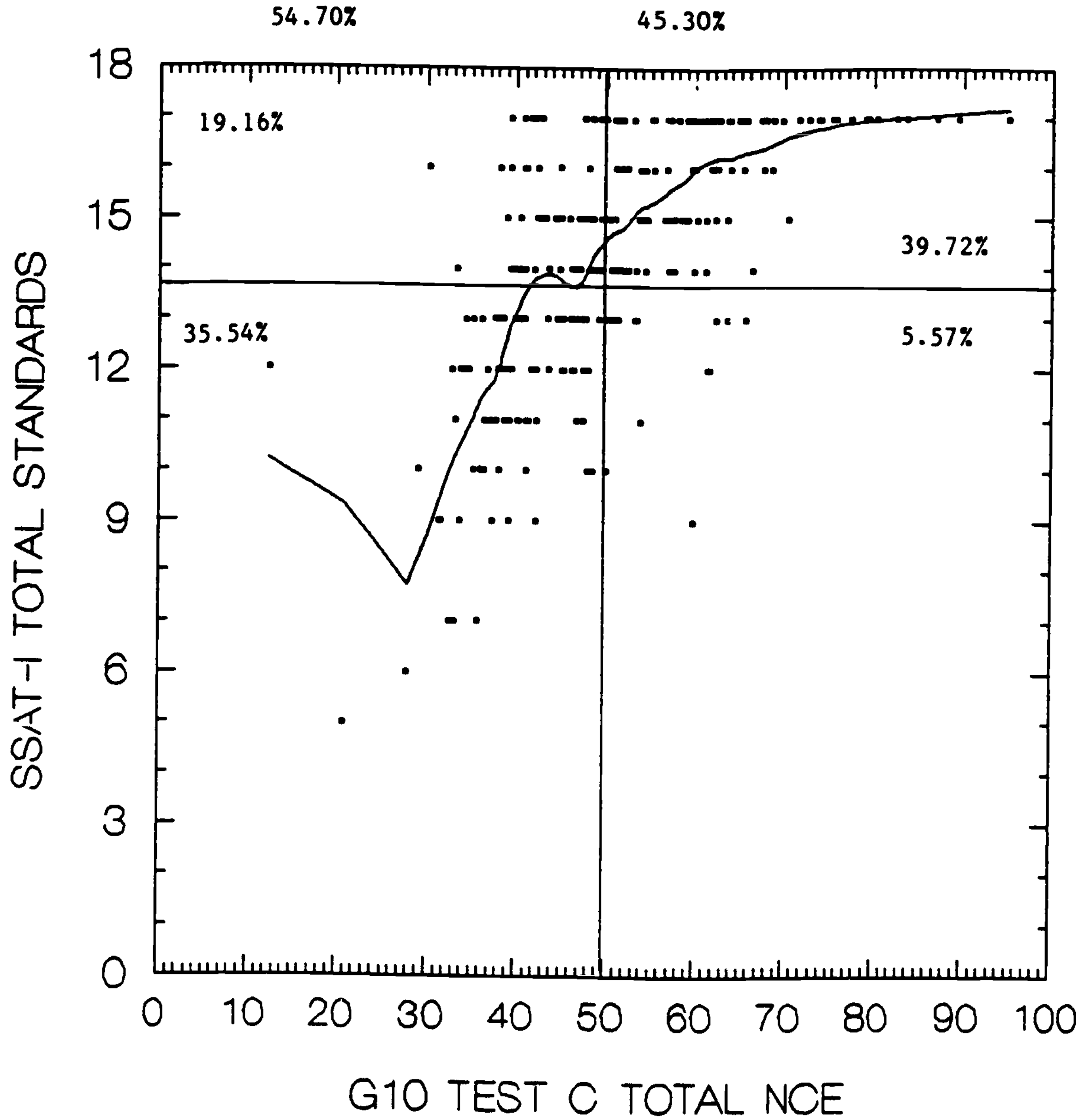


Figure A-12. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.

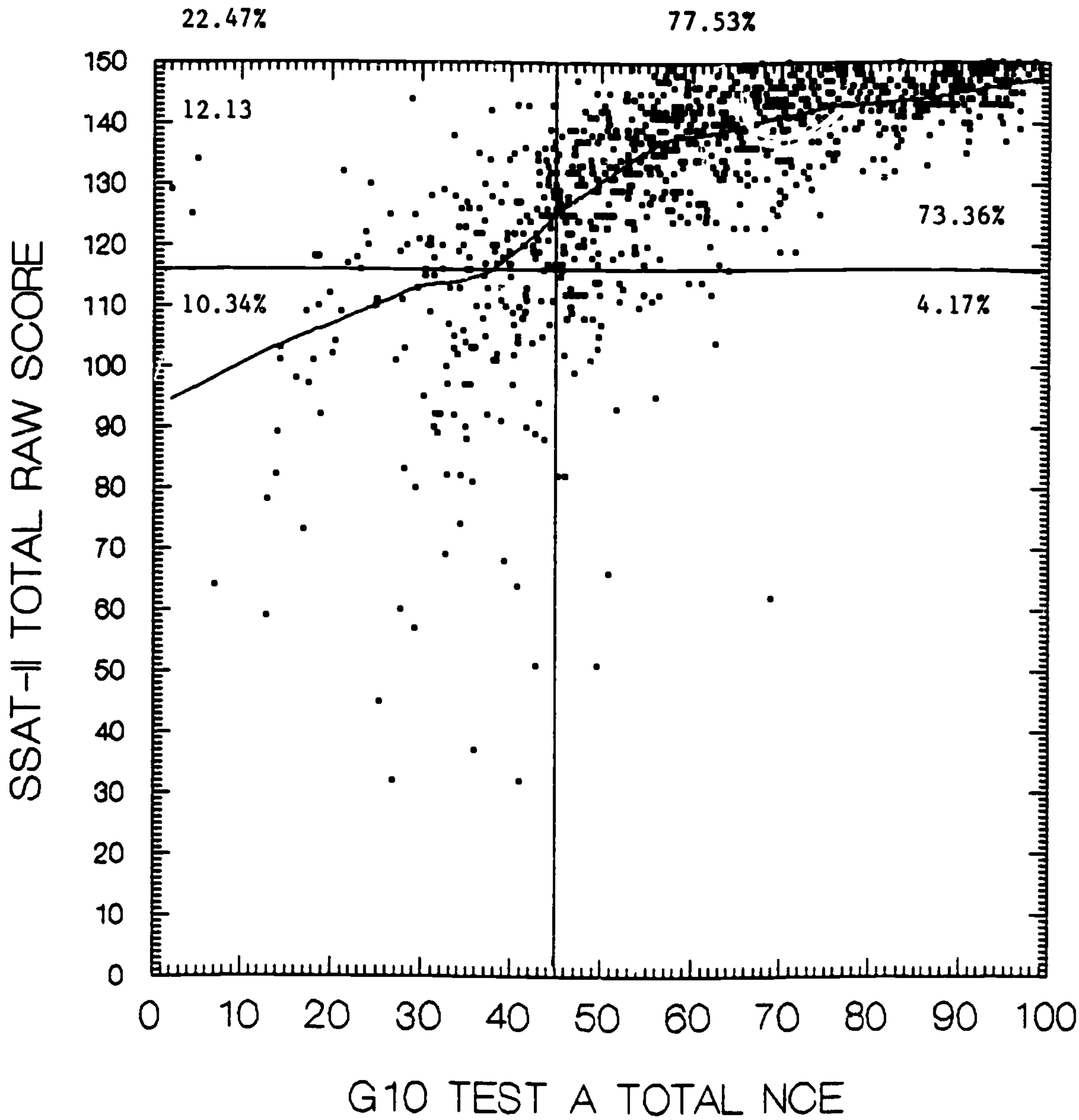
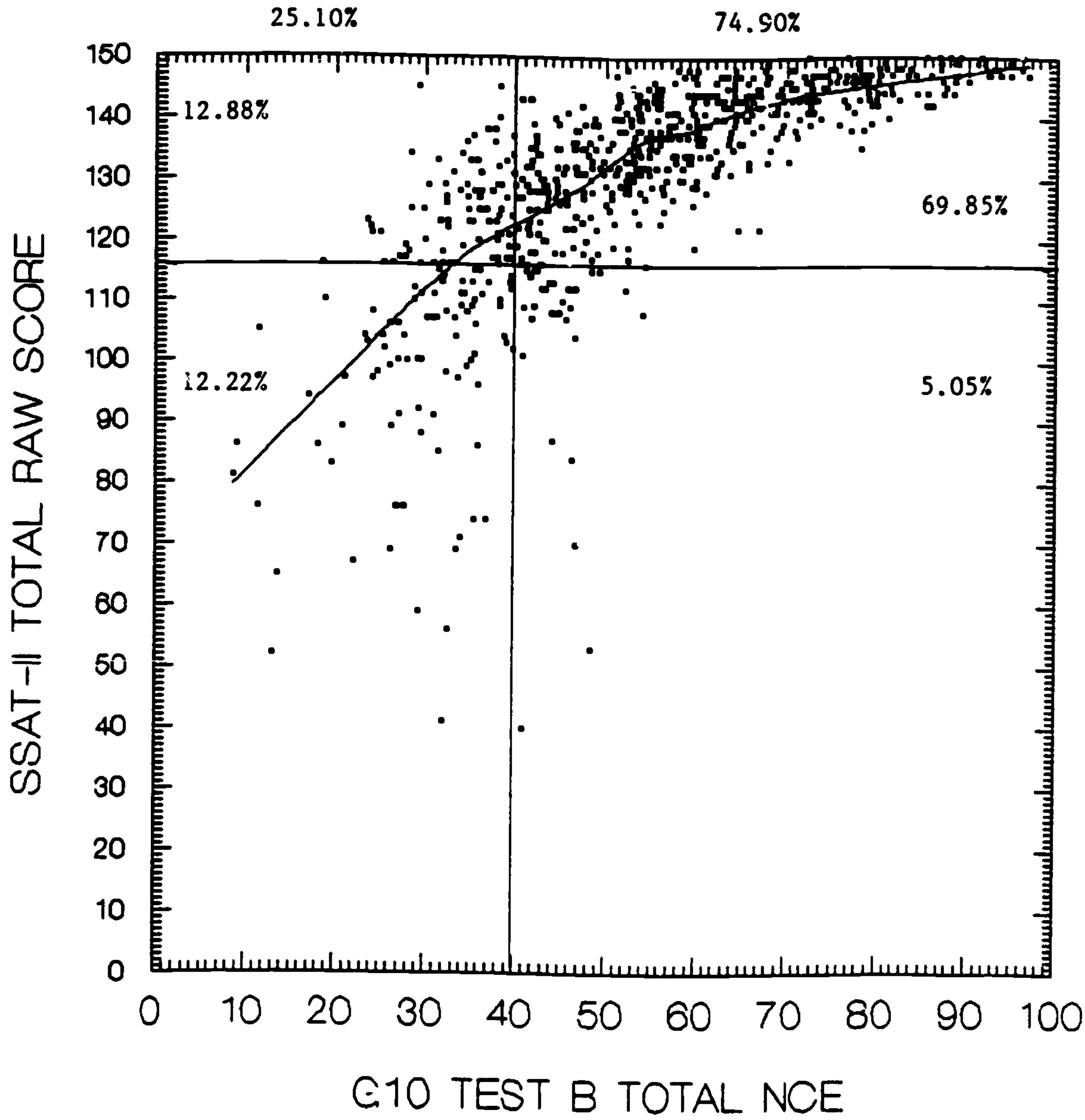


Figure A-13. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.



<u>Figure A-14</u>. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.

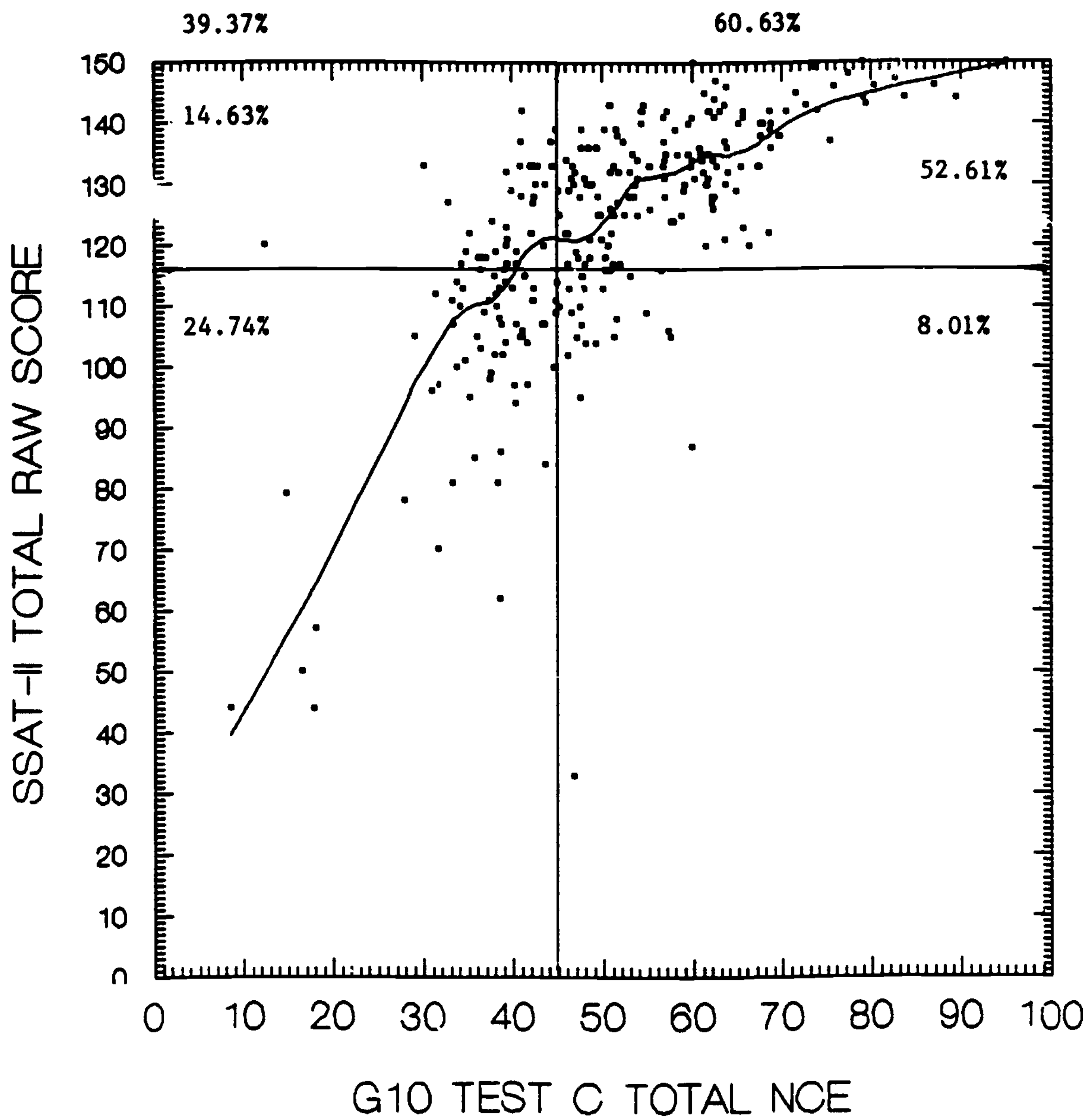


Figure A-15. Scatterplot showing the relationship between the Districts' norm-referenced and Florida's State Student Assessment Test scores.