DOCUMENT RESUME

ED 059 533 EA 003 970

TITLE	New York State Pupil Evaluation Program. School Administrator's Manual.
INSTITUTION	New York State Education Dept., Albany. Bureau of Pupil Testing and Advisory Services.
PUB DATE	Jun 70
NOTE	61p.
AVAILABLE FROM	New York State ERIC Service, Room 468 EBA, State Education Dept., Albany, New York 12224 (Microfiche copies available free to educators)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Academic Achievement; *Administrator Guides; Arithmetic; *Educationally Disadvantaged; Grade 3; Grade 6; Grade 9; Reading Tests; Student Testing; *Testing Programs; *Test Interpretation
IDENTIFIERS	New York

ABSTRACT

The Pupil Evaluation Program is a fall testing program required of all pupils in grades 3, 6, and 9 in the public and nonpublic schools of New York State. The program was originally established to provide information vital to the efficient management of ESEA Title I funds. It now provides an annual Statewide school-by-school inventory of pupil achievement in reading and mathematics. The test results identify, objectively and reliably, educational needs and indicate progress in meeting these needs irrespective of whether the funds originate from a local, state, or national source. This manual describes the tests and the general procedures for their administration, and discusses the interpretation of scores using the test results. (Author)

NEW YORK STATE

PUPIL EVALUATION PROGRAM

SCHOOL ADMINISTRATOR'S MANUAL

The University of the State of New York
The State Education Department
Bureau of Pupil Testing & Advisory Services
Albany, New York 12224

June 1970

THE UNIVERSITY OF THE STATE OF NEW YORK

Regents of the University (*with years when terms expire*)

1984	Joseph W. McGovern, A.B., LL.B., L.H.D., LL.D., D.C.L., *Chancellor*	New York
1985	Everett J. Penny, B.C.S., D.C.S. *Vice Chancellor*	White Plains
1978	Alexander J. Allan, Jr., LL.D., Litt.D.	Troy
1973	Charles W. Millard, Jr., A.B., LL.D., L.H.D.	Buffalo
1972	Carl H. Pforzheimer, Jr., A.B., M.B.A., D.C.S., H.H.D.	Purchase
1975	Edward M. M. Warburg, B.S., L.H.D.	New York
1977	Joseph T. King, LL.B.	Queens
1974	Joseph C. Indelicato, M.D.	Brooklyn
1976	Mrs. Helen B. Power, A.B., Litt.D., L.H.D., LL.D.	Rochester
1979	Francis W. McGinley, B.S., LL.B., LL.D.	Glens Falls
1980	Max J. Rubin, LL.B., L.H.D.	New York
1986	Kenneth B. Clark, A.B., M.S., Ph.D., Litt.D.	Hastings on Hudson
1982	Stephen K. Bailey, A.B., B.A., M.A., Ph.D., LL.D.	Syracuse
1983	Harold E. Newcomb, B.A.	Owego
1981	Theodore M. Black, A.B., Litt.D.	Sands Point

President of the University and Commissioner of Education
Ewald B. Nyquist

Executive Deputy Commissioner of Education
Gordon M. Ambach

Assistant Commissioner for Examinations and Scholarships
Sherman N. Tinkelman

Director, Division of Educational Testing
Victor A. Taber

Chief, Bureau of Pupil Testing and Advisory Services
Jack A. Maybee

TABLE OF CONTENTS

DESCRIPTION OF THE PROGRAM

Purpose

The Pupil Evaluation Program is a fall testing program required of all pupils in grades 3, 6, and 9 in the public and nonpublic schools of New York State. It consists of reading and mathematics achievement tests in grades 3 and 6 and reading and arithmetic minimum competence tests in grade 9. Over one million pupils enrolled in classes in more than 5,000 different school buildings take these tests each year.

The Program was established in September 1965 to help develop and maintain effective allocation, control, and evaluation procedures in the administration and use of ESEA Title I funds. It was designed to provide schools and the Department with a single uniform set of statewide test data that would identify educationally disadvantaged pupils regardless of where they attend school and give an objective picture of the severity of the problems of educational disadvantage in each and every public and nonpublic school building and school system in the State.

Although the Program was originally established to provide information vital to the efficient management of ESEA Title I funds, it now has a much broader function at both the State and local levels. It provides an annual statewide school-by-school inventory of pupil achievement in reading and mathematics. The test results objectively and reliably identify educational needs and indicate progress in meeting these needs irrespective of whether the funds originate from a local, State, or national source.

The results are used in the Governor's office, the legislature, and other State agencies and commissions as a basis for statewide education evaluation, planning, and financing. Personnel in school districts and in the Education Department also use the test results to formulate plans and make policy decisions ranging over the full spectrum of educational activities, including budget-making, supervision, allocation of personnel, improvement of instruction, evaluation of special projects and programs, and assessment of educational quality.

The use of test results in educational planning and administration is not new to those school districts in the State that have well developed districtwide testing programs. Schools, as they should, use a variety of different tests administered at different times and at different grade levels to meet their own special needs. However, it is technically impossible to equate the results of these different tests and forms into one meaningful, uniform statewide scale. Every school in the State must therefore administer the Pupil Evaluation Program tests to every eligible pupil regardless of the adequacy of the district's testing program. This is the only way the Program can maintain its value as a statewide product evaluation and continuous inventory system.

Description of the Tests

The tests used in the Pupil Evaluation Program are standardized tests developed and published by the Education Department. The tests used in the third and sixth grades are part of the New York State Elementary School Test Battery. The ninth-grade tests were developed specifically for use by New York State secondary schools to help identify those pupils in need of special attention if their reading and mathematical skills are to reach an established minimum level of competence before graduation.

The tests and their essential characteristics are presented in the table below:

Grade	Name of Test	No. of Questions	Testing Time in Minutes	Type of Scoring
3	Reading Test for New York State Elementary Schools – Beginning Grade 3 - Form A	53	35	Hand Only
	Pt. 1. Word Recognition	(25)	(15)	
	Pt. 2. Reading Comprehension	(28)	(20)	
3	Mathematics Test for New York State Elementary Schools- Beginning Grade 3-Form C	60	50	Hand Only
	Pt. 1. Concepts	(26)	(20)	
	Pt. 2. Computation	(16)	(12)	
	Pt. 3. Problem Solving	(18)	(18)	
6	Reading Test for New York State Elementary Schools- Beginning Grade 6-Form A	66	40	IBM 805
	Pt. 1. Word Recognition	(30)	(15)	IBM 1230
	Pt. 2. Reading Comprehension	(36)	(25)	Digitek
6	Mathematics Test for New York State Elementary Schools- Beginning Grade 6-Form C	67	60	IBM 805
	Pt. 1. Concepts	(27)	(20)	IBM 1230
	Pt. 2. Computation	(20)	(20)	Digitek
	Pt. 3. Problem Solving	(20)	(20)	
9	Minimum Competence Test in Reading for New York State Secondary Schools-Form A	40	40	IBM 805 IBM 1230 Digitek
9	Minimum Competence Test in Arithmetic Fundamentals for New York State Secondary Schools-Form A	25	40	IBM 805 IBM 1230 Digitek

The reading and mathematics tests for the third and sixth grades measure instructional goals described in syllabuses and curriculum guides developed and published by the Education Department. The tests were developed during 1959 and 1960 and made available to schools in 1961. During 1967 and 1968, the mathematics tests were completely revised to provide a better measure of the objectives of the current elementary school modern mathematics program, and a new form, Form C, was used for the first time in the fall 1969 program. The reading tests are now undergoing revision, and new forms are expected to be available for use in the fall of 1972.

The reading and mathematics tests are designed primarily for supervisory and administrative uses. They are brief survey tests which provide supervisors and administrators with the type of information needed to evaluate instructional programs in their schools. The test results, including those for the part scores as well as for the total scores, reveal weaknesses that have existed in previous learning situations, identify program elements in need of greater emphasis, monitor program accomplishments, and measure the effects of program innovations and changes.

When the test results are interpreted within their limitations, teachers can also use the test results effectively. The test results identify individual pupils whose scores indicate a need for further diagnostic testing or for other special consideration. Through an analysis of the test results, teachers can determine individual pupil and class strengths and weaknesses and can plan instruction accordingly.

The minimum competence tests used in the ninth grade serve an entirely different function. They are much easier tests and are designed specifically to discriminate among pupils with marginal skills in reading and arithmetic. They are not intended for use in diagnosing program strength and weakness but serve solely to identify pupils in need of special compensatory attention. Administering these tests to average and above average ability pupils or those with the abilities and potentialities for taking Regents examinations serves no useful purpose, because the results will indicate only that which is already well known - that the skills of these pupils are above the minimum level of competence.

According to present plans, the ninth-grade minimum competence tests will be replaced in the Pupil Evaluation Program in the fall of 1972 with broad range tests similar to those now in use in the third and sixth grades. The results of these new tests will provide administrators and supervisors with the type of information needed to evaluate the effectiveness of reading and mathematics instruction in grades six through eight and to improve skills in grades nine through twelve.

Program Administration

The administration of the Pupil Evaluation Program tests and the reporting of scores according to established procedures is the responsibility of the chief school administrator. Even though many phases of program administration can be delegated to building principals, some type of centralized control is required for coordinating the flow of test materials and reports to and from the Department and maintaining uniform accuracy and efficiency.

General Procedures In general, the Department sends order forms for test materials, one for each school building, to the chief school administrator. The administrator distributes these order forms to the schools where they are filled out and returned to the central office for forwarding to the Department. The Department ships the bulky test materials and the score report forms directly to schools according to directions on the order forms. Schools administer and score the tests, fill out the score report forms, and forward them to the central office. When the score reports for all schools in the district are on hand and complete, the central office forwards them to the Department.

Specific information and directions for ordering materials, for administering and scoring the tests, and for reporting the test scores to the Department are given in Appendix II.

Dates of Administration All Pupil Evaluation Program tests are to be administered during the 2-week period starting the fourth Monday after Labor Day. The actual administration dates through 1975 are as follows:

October 5 - 16, 1970	October 1 - 12, 1973
October 4 - 15, 1971	September 30 - October 11, 1974
October 2 - 13, 1972	September 29 - October 10, 1975

The test administration dates are set to begin 3 weeks after the opening of schools in most upstate and Long Island commu[illegible]es. Because test performance is significantly related to school attendance, schools should generally and consistently choose a date during this 2-week period that is between 15 and 25 days after the start of formal classroom instruction. Any schools in which the pupils will not have had 15 days of instruction by the end of the 2-week test administration period will be granted permission to administer the tests during the following week.

Adherence to the scheduled testing dates is of crucial importance. If the test results for the pupils in a school are to be interpreted on the basis of the test norms, and if comparisons of test performance are to be made from 1 year to the next, all pupils must be tested within the same 2-week period. Otherwise, actual changes in reading and mathematical competency will be obscured by changes in pupil test performance caused by differences in times of administration.

Pupils to be Tested Other than the exceptions noted below, all pupils enrolled in grades 3, 6, and 9 in all public and nonpublic schools must be tested. In ungraded classes, all pupils of equivalent ages must be included in the testing.

a. Exempted Pupils (all tests):

Pupils with special mental or physical handicaps, including those in CRMD classes and those with severe hearing or sight handicaps, impairing brain damage or birth defects, or serious emotional maladjustments, may be exempted from taking any Pupil Evaluation Program tests.

b. Pupils Excused from Ninth-Grade Tests:

Pupils who will be taking a Regents examination in business arithmetic or mathematics may be excused from taking the ninth-grade arithmetic test; those who will be taking the Regents Comprehensive Examination in English in either their junior or their senior year of high school may be excused from taking the ninth-grade reading test.

c. Non-English Speaking Pupils:

Pupils who do not have sufficient mastery of the English language to be able to comprehend the directions or read any of the test questions and would therefore earn only a chance score on the test should not be tested because of the adverse emotional and psychological factors involved in such stress situations. However, scores of zero (0) for each of these pupils should be included in all reports of test results.

All other pupils with language handicaps who will not be adversely affected by the testing situation should be given an opportunity to take the tests so that whatever test performance they are able to exhibit will become part of the grouped test results.

By including all non-English speaking pupils and all those with language handicaps in the program results in this way, these pupils can be identified and located by school so they can be given the same special compensatory help and attention as other educationally disadvantaged pupils.

THE SCHOOL TESTING REPORTS

Schools prepare frequency distributions of the scores of their pupils and report these distributions to the Department on optical scanning report forms. The distributions of scores are processed by computer, and reports which summarize the results in conveniently interpretable form are returned to each school and central office.

Report Tables

The school testing reports consist of (A) one School Summary Table for third-grade and sixth-grade tests, (B) one Total Score Distribution Table for each third-grade and sixth-grade test, and (C) one Ninth-grade Testing Report for both ninth-grade tests.

A. School Summary Table - Grades 3 & 6: This table shows the number of pupils tested in the school, the median raw scores for each test and subtest, and the corresponding statewide percentile ranks of these medians. For public schools only, the table gives the raw score medians and the statewide percentile ranks for the school system. In addition, for each test and subtest, the table contains the statewide percentile ranks for the median scores of pupils in selected reference groups.

B. Total Score Distribution Tables - Grades 3 & 6: These tables report the cumulative percent of pupils in the school (and for public schools in the system) at each of nine successive levels of achievement. Scores at achievement levels 1 through 3 indicate below minimum competence, 4 through 6 average competence, and 7 through 9 above-average competence. There is a separate table for third-grade reading, third-grade mathematics, sixth-grade reading, and sixth-grade mathematics.

Each table also provides raw score medians, means and standard deviations, and comparable information for selected reference groups.

C. Ninth-Grade Testing Report: This table reports the number and cumulative percent of pupils in the school (and for public schools in the system) with scores below minimum competence - those at levels 1 through 3. Below minimum competence raw scores on the reading test are 27 and below; on the arithmetic test, 19 and below. The table also gives the number tested, excused and exempted, and comparable information for selected reference groups. Ninth-grade medians and the cumulative percent of pupils at each of achievement levels 4 through 9 are not included in the report. As explained in the section describing the tests, the present ninth-grade tests are designed solely to identify pupils below minimum competence and the results, therefore, do not discriminate effectively or reliably among pupils of average and above-average competence.

Reference Groups

The reference group results provided in the report tables enable a school to interpret the scores of its pupils by comparing them with the scores of other pupils with certain similar characteristics and educational experiences. All reference group results are based on the scores obtained the year previous to the report year. Also, not all reference group results are reported to every school. Each school report contains only the results for those reference groups which will be of the most direct help in interpreting the test results for that school.

Test results for the following reference groups appear on every testing report:

1. Public and Nonpublic--State: All pupils in public and nonpublic schools statewide.

2. Public and Nonpublic--Community Type: All pupils in public and nonpublic schools located throughout the State in communities having similar size and type characteristics. The seven community type reference groups are:

Type 1... New York City
(population over 1,000,000)
Type 2... Large Size Cities
(population over 100,000)
Type 3... Medium Size Cities
(population 50,000-100,000)
Type 4... Small Size Cities
(population under 50,000)
Type 5... Village and Large Central Districts
(enrollment over 2,500 pupils)
Type 6... Large Rural Districts
(enrollment 1,100-2,500 pupils)
Type 7... Small Rural Districts
(enrollment under 1,100 pupils)

Test results for the following reference groups appear on public school testing reports only:

1. Public--County or Borough: All pupils in public schools in the county in which the school is located. In New York City, the borough is identical with the county.

2. Public--State: All pupils in public schools statewide.

Test results for the following reference groups appear on Roman Catholic school testing reports only:

1. Roman Catholic--Diocese: All pupils in Roman Catholic schools in the diocese in which the school is located.

2. Roman Catholic--State: All pupils in Roman Catholic schools statewide.

Test results for the following reference groups appear on private nonpublic school testing reports only:

1. Private--State: All pupils in private schools statewide. (Excludes pupils in Roman Catholic schools.)

2. Nonpublic--State: All pupils in nonpublic schools statewide. (Includes pupils in Roman Catholic schools.)

Norm Population

The percentile ranks and the achievement levels reported in the school testing reports and the test manuals are based on the scores of all New York State public and nonpublic school pupils tested in the fall 1966 administration of the Pupil Evaluation Program. Approximately 900,000 pupils in grades 3, 6, and 9 in over 5,100 different school buildings were tested. This constituted over 95 percent of all pupils enrolled in grades 3 and 6 and about 90 percent of all pupils enrolled in grade 9. Pupils with impairing mental, emotional, and physical handicaps were exempted from taking the tests and are not included in the norm population. All non-English speaking pupils are included in the norm population, however; zero scores were entered in the results for any pupil who was not tested for this reason.

The norms for the new Form C mathematics tests which replaced the Form A arithmetic tests in the fall of 1969 were obtained by equating raw scores on the new tests to those on the old tests, using the equi-percentile method. Thus, the norms for the new forms are still based on pupil competency as it existed in 1966 and are directly comparable to those for the reading tests.

The data for equating the scores of the new mathematics test to those of the former arithmetic tests were obtained in October 1969 by administering both Form A and Form C tests to about 2,000 third-grade pupils and 2,200 sixth-grade pupils.

These pupils were specially selected to be representative of all pupils statewide in respect to community type and public and nonpublic characteristics. The samples included all the pupils in 35 schools in 24 different public school systems and 19 nonpublic schools.

Minimum Competence

Minimum competence on the third- and sixth-grade tests is defined as the raw score that exceeds the scores of 23 percent of the pupils in the 1966 norm population. Pupils who obtain scores above this 23d percentile (achievement level 4 and above) are considered to have competency above this arbitrarily defined statewide minimum, and pupils who score at or below this point are considered to have competency below the statewide minimum standard.

The minimum competence point on each test was established after careful study by the staff members in the reading and mathematics education bureaus. In the combined judgment of these specialists, pupils who score below the minimum competence points on the Pupil Evaluation Program tests do not have the skills needed to make reasonable progress in an average New York State classroom, and therefore are in need of some type of compensatory attention and help. Depending upon the needs of the individual pupil, this attention and help can range from merely additional time given to the pupil by the classroom teacher to a concentrated effort by a number of specialists in a variety of special programs and services.

The minimum competence points are intended to provide only a convenient frame of reference for use in interpreting and understanding the test results. They are practical reference criteria and not absolute standards. As such, each minimum competence point is an operational or working definition of a rank on a scale of test scores that should be used only to identify and locate within the school and the State those pupils in most need of help and to establish a base line against which progress in meeting the needs of these pupils can be evaluated.

Interpretation

The school testing reports contain many columns of statistical data. Some data summarize the scores of the pupils in the school, while other data serve as a basis for interpreting these scores. Also, the results are presented in several different statistical forms which provide a variety of ways of looking at them and gaining insights into their meaning.

As with all statistical data, test results come alive and are useful only when given meaning by the user. Test interpretation starts with an examination of the test content so the skills measured and the techniques used to measure them are fully understood.

Next, the raw test scores must be converted into a meaningful statistic, such as a percentile, in order to determine its standing among the norms of pupils in a well-defined norm group. Finally, the relative standing of the score must be evaluated by relating it to pupil, school, and community factors that affect achievement.

Only after this interpretative process has been fully completed do test results become evaluative criteria ready for use in determining school program effectiveness.

The first step in interpreting the school testing reports is to analyze the data systematically and to formulate descriptive generalizations. For example, "third-grade skills are better than sixth-grade skills." These generalizations can then be restated in question form - "Why are third-grade pupils more competent in reading than in mathematics?" This procedure will quite naturally lead to a qualitative assessment of school programs, instructional services, and curricular materials.

Test results usually provide answers to questions. On the other hand, they often raise more questions than they answer. Differences between the reading and mathematics scores of third-grade pupils, for example, bring to mind several more questions. "Are pupils actually more competent in reading than in mathematics?" "What other evidence is there to confirm or refute the conclusion that there is a difference?" "What steps can be taken to improve pupil competence in mathematics?" Each statistic in a report should be viewed as if it had a question mark after it. Not to raise doubts about its authenticity, because it is accurate unless mechanical errors have intervened, but to emphasize its greater potential value when used as a question rather than a descriptive statement.

A Summary Report Table Interpretation: This table describes competence in each skill and nonskill area with a single representative statistic. The percentile ranks of the median raw scores give an immediate indication of whether pupil competence in a school is higher than, lower than, or about the same as the competence of pupils in the statewide norm population.

For example, if the percentile rank of a school median is about 50 (45 to 55), pupil competence in that skill area is similar to that of pupils in the norm population. If the school median is more than five percentile ranks above or below 50, competence in the school is above or below that of the norm group. As a rule of thumb, only differences in percentile ranks of more than five points indicate significant differences in competence.

Because the percentile ranks for the reference groups are also based on the same norm population as those for the school (all pupils statewide in 1966), the school and reference group percentile ranks are directly comparable and indicate differences or similarities in pupil competence. For example, if the percentile rank of a school's third-grade reading median is 50 and the percentile rank of the county median is 40, a school can interpret this to mean that its third-grade pupils are more competent in reading than are all third-grade pupils in the county combined.

The summary report tables for public schools also contain the medians and percentile ranks for all pupils in the system of which they are a part. For a public school, therefore, the pupils in the system become an additional reference group, and the percentile ranks of the system medians are an important basis of comparison. For a public school system, on the other hand, the system results are primary data and can be interpreted by making the same comparisons with the norm population and reference groups as those outlined for a school.

The results reported for the public school system are for the current year, the same year as those reported for the school. The reference group results, however, are for pupils tested the previous year. For most interpretations, this distinction is of more technical than practical significance. The reference group results, based as they are on large numbers of pupils, do not change notably from one year to the next.

An abbreviated sample of a school summary table on which a few selected percentile ranks have been entered is provided below to show typical illustrations that can be made from the data.

SCHOOL SUMMARY TABLE - GRADES 3 AND 6

GRADE	NAME OF TEST	SCHOOL RESULTS MEDIAN RAW SCORE -2-	SCHOOL RESULTS PERCENTILE RANK -3-	PERCENTILE RANKS OF REFERENCE GROUP MEDIAN RAW SCORES PUBLIC-NONPUBLIC C-TYPE -8-
3	WORD RECOGNITION	50	Competence in Word Recognition skills similar to that of norm population	58
	READ. COMPREHENSION	47		61
	TOTAL READING	50	Competence in reading skills similar to statewide but lower than for schools in similar type communities	60
3	CONCEPTS	56	Concepts less well developed than Problem Solving skills	63
	COMPUTATION	61		62
	PROBLEM SOLVING	63		63
	TOTAL MATHEMATICS	63	Competence in mathematical skills similarly higher than statewide	63

There are many different types of comparisons that can be made in analyzing and interpreting the data in the summary report table. Each type of comparison is suitable for achieving only certain specific purposes, however, and the types should not be used indiscriminately merely because the data are available. The following examples are illustrations of the different ways in which percentile ranks for school medians can be given additional meaning by comparing them with other report data.

1. Comparing the results for one skill area with those for another.

"The third-grade median for mathematics is 10 percentile points below that for reading," or
"The third-grade pupils are more competent in reading than in mathematics."

2. Comparing the results for one grade with those for another.

"The percentile rank of both the third-grade and the sixth-grade median scores in mathematics is 60," or
"The third- and sixth-grade pupils appear to be doing equally well in mathematics."

3. Comparing the results within a skill area with each other and total.

"Third-grade pupils rank at the 50th percentile in mathematical skills. However, their problem solving scores rank at the 40th percentile compared to those for conceptual and computational skills which rank at the 50th and 65th, respectively," or
"In general, the mathematical skills of the third-grade pupils are about average. Their computational skills are above average, and their problem solving skills are below average."

4. Comparing the results for a skill or subskill area with those for an appropriate reference group.

"The school median for third-grade reading has a percentile rank of 40. The county median ranks at the 35th; the community type median at the 42d," or

"Although the reading skills of the third-grade pupils are below the statewide average, their skills are better than those of the pupils in the county and only slightly below those of pupils in similar type communities throughout the State."

5. Comparing the results for the current year with those for previous years.

"During the past three years, sixth-grade median reading scores have dropped from percentile ranks of 60 and 55 to 50 this year," or

"The reading competence of sixth-grade pupils has steadily declined during the past 3 years. Their scores are now 10 percentile points lower than 3 years ago."

6. Comparing the differences between the results in two skill areas with differences in previous years.

"There is a difference of 10 percentile points between reading and mathematics achievement in the third grade. This difference has steadily decreased from the 20 points that existed 4 years ago," or

"Over the past 4 years, mathematics skills have improved to a point where the pupils' mathematics scores are only 10 percentile points below their reading scores.

7. Comparing the differences between the results in two skill areas with differences in a reference group.

"The third-grade pupils' reading scores rank at the 55th percentile; their mathematics scores at the 45th percentile. This 10 percentile point difference is similar to the difference between the reading and mathematics scores of other third-grade pupils in the county," or

"Third-grade pupils are more competent in reading than in mathematics. This is a common tendency among pupils in this county."

8. Comparing the differences between the results for the third and sixth grades with differences in previous years.

"The median reading score is at the 50th percentile in the third grade, the 40th in the sixth grade. Last year there was a 15 percentile rank difference. The third-grade percentile rank was 50, the sixth-grade was 35," or

"Last year, the reading scores of the third-grade pupils were 15 percentile points higher than those for the sixth-grade pupils. Because of a 5 percentile point improvement in the sixth-grade scores this year, the difference in reading competence between third and sixth grade pupils is now only 10 percentile points.

B Total Score Distribution Table Interpretation: Whereas the Summary Report Table describes competence in a skill or sub-skill area with a single representative statistic so that pupil competence in one area or one grade may easily be compared with that in another, the score distribution tables describe the full range of pupil competence within a skill area. The tables show whether the range is atypically broad or narrow, or whether it is similar to that found in most schools statewide. They also show the achievement level or levels at which many or most of the pupils in a school are functioning.

The achievement levels on the tables are fixed standards, defined by raw score values that remain constant from one year to the next. They were derived from the same 1966 norm population scores as those used to obtain the percentile ranks. There are nine units on the achievement level scale with each unit one-half standard deviation higher than the one below it. Thus, each unit is theoretically equal to the other.

An oversimplified but perhaps adequate view of achievement levels would be one in which the full range of achievement, from that represented by the pupil with the least competency in the State to that represented by the pupil with the most competency in the State, has been divided into nine equal parts and the parts labelled from the bottom up by the numbers 1 through 9. In such a scale, for example, the cumulative percents for the 1966 norm population from the first to the ninth level are 4, 11, 23, 40, 60, 77, 89, 96, and 100. Thus, the scores of the lower 4 percent are at level one, the next 7 percent at level two, the next 12 percent at level three, etc.

The percent of pupils within each achievement level can be obtained for a school or reference group by subtracting the cumulative percent for the next lowest level from the cumulative percent at that level. For example, in the 1966 norm population, 12 percent of the pupils obtained scores at level 3, (23 minus 11). The actual number of pupils within each level can be obtained for a school (not cumulative percent) by multiplying the percent of pupils at the level by the number of pupils tested, which is given at the bottom of the school results column.

If an analysis of pupil competence in terms of nine levels of achievement is too detailed to give a meaningful overall picture of pupil competence in a skill area, an analysis can be made in terms of only three descriptive classifications. Achievement levels 1-3 represent competence below the minimum statewide established level, levels 4-6 represent average competence, and levels 7-9 represent above-average competence. The percentages and numbers can easily be consolidated to describe competency in terms of these classifications.

As in the school summary table, the reference group data provide a basis for comparing the competence of the pupils in a school with that of other groups of pupils to determine whether or not the distribution of competence is similar to or different from that of pupils in each of the reference groups. For example, a lower cumulative percent for a school than for a reference group indicates greater competence among pupils in the school than in the reference group.

That is, a smaller proportion of the pupils in the school than in the reference group obtained scores at and below that level.

In general, the principles and procedures for analyzing, interpreting, and comparing the data in the total score distribution report tables are the same as those outlined in the preceding section on Summary Report Table Interpretation. The abbreviated table below offers some specific examples of initial interpretations and is followed by some typical generalizations derived from the data in the table.

TOTAL SCORE DISTRIBUTION TABLE - GRADE 6 READING

ACHIEVEMENT LEVEL	CUMULATIVE PERCENT OF PUPILS IN THE SCHOOL -1-		CUMULATIVE PERCENT OF PUPILS IN OTHER REFERENCE GROUPS: PUBLIC AND NONPUBLIC C-TYPE -5-	
9	100	38% have above average competence (100 - 62 = 38)	100	
8	92		95	30% above average
7	78		85	
6	62	55% have average competence (62 - 7 = 55)	70	
5	41		50	57% average
4	10		29	
3	07	7% have below minimum competence (7 - 0 = 7)	13	
2	05		4	13% below minimum
1	04		1	
N Tested	200		111,671	

Typical generalizations obtained from the above data:

1. Most sixth-grade pupils have average or above reading competence (93 percent, or 186 out of 200).

2. Sixteen pupils are reading at the very highest level of competence (100 percent - 92 percent = 8 percent times 200).

3. There are 14 pupils with below minimum competence skills in reading (7 percent times 200), and eight of these pupils have skills that are far below the minimum (4 percent at level 1).

4. The percent of pupils with above-average competence in reading is higher than that in schools in similar type communities statewide (38 percent compared to 30 percent).

5. Schools in similar types of communities have about the same proportion of pupils reading with average competence (55 percent to 57 percent).

6. Although fewer pupils are below minimum competence in reading skills (7 percent to 13 percent), schools in similar type communities do not have as large a proportion at the lowest level (4 percent to 1 percent).

A descriptive report of the above results could be written in many ways depending upon the reading audience and the purposes of the writer. One form might be as follows:

"The test results show that 38 percent of the third-grade pupils are above average in reading competence, 55 percent have average competence, and 7 percent are below the statewide established standard of minimum competence. This compares to percentages of 30, 57, and 13, respectively, in schools in similar type communities statewide. Even though there are 8 percent more pupils above average in competence and 10 percent fewer pupils below the minimum level of reading competence, eight pupils (4 percent) are reading far below the minimum acceptable level of competence and need considerable compensatory help. There are also six other pupils whose reading competence is below the minimum needed by them to make normal progress in those classroom activities involving reading."

C Ninth-grade Report Table Interpretation: The ninth-grade table provides only the percent of ninth-grade pupils below minimum competence in reading and in arithmetic. The lower the percent, the fewer the number of pupils below minimum competence.

The percents for both the school and the reference groups are based on the number of pupils tested plus the number excused. It is assumed that all pupils excused from taking the ninth-grade minimum competence tests would have scored higher than achievement level 3, or, in other words, above the minimum competence level.

USING THE TEST RESULTS

The results from the same tests can be used for several purposes. They can be analyzed in different ways to provide teachers, supervisors, coordinators, and administrators with the special kind of information they need about pupils. The same test results help develop hindsight - to evaluate what has happened, and to help improve foresight - to plan for what should happen. They also contain implications for the present by showing what pupils are learning and what needs require immediate attention.

Classroom Use

Third-grade and sixth-grade classroom teachers will find the test results particularly helpful in planning instruction. Third-grade results are available at the beginning of the year, and, therefore, are an excellent basis for determining what skills need further development before going on to new work and what skills have already been sufficiently mastered.

The test manuals contain detailed suggestions on how to interpret pupil scores for instructional uses. The reading manuals have brief, separate sections on analyzing and treating word recognition and reading comprehension problems. The mathematics manuals have similar sections on concepts, computations, and problem solving.

Administrative Uses and Techniques

The Pupil Evaluation Program is an information system designed primarily to help the school administrator know his pupils and to alert him to changes in pupil competence as they occur. The program is also a device with which the administrator can monitor school programs and determine their effectiveness. The results provide the type of objective basic information that the administrator needs to report and discuss pupil accomplishments and needs at meetings with the school staff, the board of education, and the public. The results give clues as to what new resources are needed, how present resources can best be allocated, and where different priorities must be established.

The critical point in using test results is the selection of relevant data and the presentation of comparisons that clearly contribute to the purpose for which the results are being used. Rather than attempt to use all the data in one general analysis intended to serve all purposes, the administrator who can extract from the reports only that data which has a specific bearing on a particular problem or situation and can interpret the data in terms of a single purpose will find the results easier to understand and easier to communicate to others.

A list of purposes for which administrators find test results helpful is given below. The list is not intended to be all inclusive, or any one item exclusive of the other.

In every case, the results must be interpreted and then related to all the other relevant information the school has available concerning the pupils, the community, and the school instructional, curricular, and special services. Once this has been done, the administrator is in a position to evaluate, to discriminate between the effective and ineffective, and to plan accordingly. Each analysis can be made for the school system as a whole or by school buildings, grades, or classrooms.

1. Evaluating pupil strengths and weaknesses.
   How do reading skills compare with those in mathematics? Why are pupils stronger or weaker in some skills than in others? Is this acceptable? If not, what can be done about it?

2. Evaluating appropriateness of present curriculum materials and determining what areas need what types of new materials.
   Is this course of study and set of materials more effective than another? Are the levels represented and assumed in these materials appropriate to the competence of pupils? Are the quantities of materials at each ability level proportional to the number of pupils at each level of competence? What type and level of new materials are needed?

3. Measuring trends in pupil achievement and applying appropriate treatments to offset any negative trends.
   Are pupils' skills improving? What is causing the change? How can a decline in a skill area be overcome? How do these trends relate to trends in pupil, school, and community factors?

4. Determining reasonable levels of achievement and establishing realistic goals.
   Considering the abilities of the pupils, community conditions, and school resources, what is a reasonable goal for the school and the school system? Is it realistic to expect every pupil to develop competence above the statewide minimum level? Can more of the able pupils develop above-average competence?

5. Diagnosing instructional strengths and weaknesses.
   Is instructional time properly allocated? Is it based on actual pupil needs or on assumed needs? Considering pupil learning rates and initial pupil competencies and readiness, are teachers uniformly meeting appropriate and realistic instructional goals? Are specialists assigned to the school buildings and grade levels where they are most needed? Are Title I and other projects serving the appropriate group of pupils?

6. Identifying pupils with special problems who are in need of special services.
   How many pupils are below minimum competence? Where are they and what is being done for them? What expectations are assumed from present special services and programs? Are they succeeding? Are they economical in terms of the results obtained?

7. Evaluating program changes and innovations.
   What is the effect of changes in schedules? of special classes? of ungraded class situations? Have special projects indirectly related to reading and mathematics, such as those designed to improve motivation, interest, nutrition, self-concept, etc., contributed to an improvement in pupil skills?

8. Improving public relations.
   Do teachers, board members, and the public have a factual basis for forming opinions about the school and school system? Are the accomplishments of pupils and the school as well understood as the failures? Are the costs and the results of regular and special programs related?

9. Allocating professional staff time.
   What skill areas need additional staff time and attention? Considering the results, are present staff assignments realistic? From what areas can staff time be shifted?

10. Budgeting materials and other resources.
    What should be given priority? What equipment and materials are most needed on the basis of conditions revealed by the test results? What items should be requested in the annual budget? Do the results show a need for additional staff? Which areas?

11. Coordinating, supervising, and following-up.
    Do the results show the effects of planned changes? Why not? Why were expected goals not achieved? Is program supervision effective? Are staff training needs apparent, and are opportunities available?

Even though the tests in many local school testing programs measure the same skills as those measured by the tests in the Pupil Evaluation Program, the school testing reports for this program, with their State norms, reference group results, achievement levels, and minimum competence points, give a slightly different view of pupil performance which augments rather than duplicates the information provided by other standardized tests. For this reason, and because the test results for one district are easily compared with those for other districts, each chief school administrator is urged to analyze and to interpret the test results for his district and to prepare, or have a member of his staff prepare, a report of these results for distribution to the community. He should not leave to the news media and the public the task of interpreting the results and drawing conclusions solely on the basis of their own information.

The following questions provide guidelines that may be helpful in preparing such a report.

1. In general, how do reading and mathematics skills in the school or system compare with those of pupils statewide?

2. How do these skills compare with those of pupils in other reference groups?

3. Is there a difference between the reading and mathematics tests results? Is the percentile rank for the reading median higher or lower than that for the mathematics median?

4. Are pupils relatively more competent at one grade level than at another? Is the reading or mathematics program more (or less) effective below grade 3 than below grade 6?

5. Is there a difference among the component skills in reading (or mathematics)? Is the percentile rank of the word recognition subtest median higher or lower than that for reading comprehension? Which mathematics component has the highest and which has the lowest rank?

6. Has competence increased or decreased during the past few years? Have past weaknesses been remedied?

7. What proportion of the enrollment in each grade level scores in the low, middle, and high score ranges? Are there proportionally more high, low, or average achievers than elsewhere?

8. Is there a disproportionately high concentration in a particular part of the score range? Does the school enroll a heavy concentration of pupils who are below minimum competence? Is the need for remedial instruction greater than for enrichment programs?

9. Are pupils equally competent in all buildings and classrooms? Which ones are atypically high or low? Is the concentration of low (or high) achievers greater in one school than in another?

10. Has there been an increase (or decrease) in low achievers at a given grade in reading or mathematics over the years? What changes have occurred in the proportion of high achievers from year to year?

One of the clearest ways to present test results is through the use of graphs. The illustrations below are examples of types of graphs that can be used.

Fig. 1. Percentile Rank of Medians

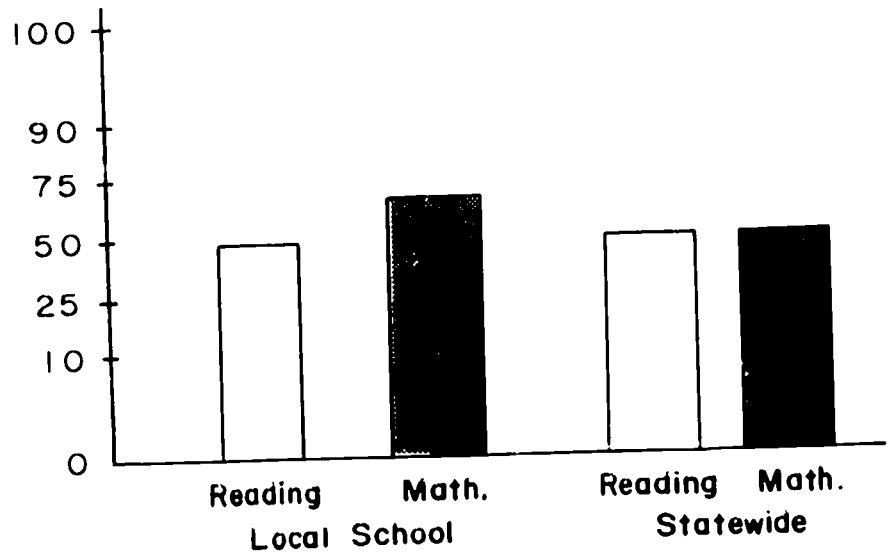


Fig. 2. Percent of Pupils Within Three Score Ranges

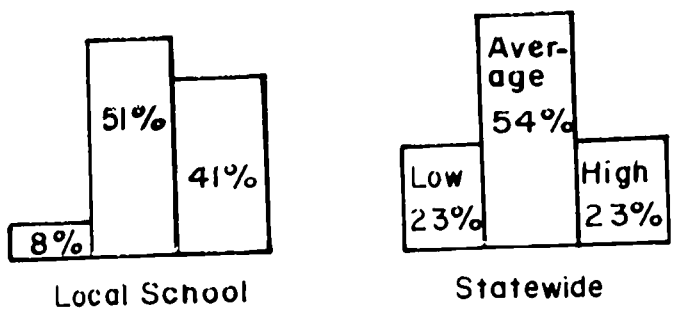


Fig. 3. Trends in Achievement

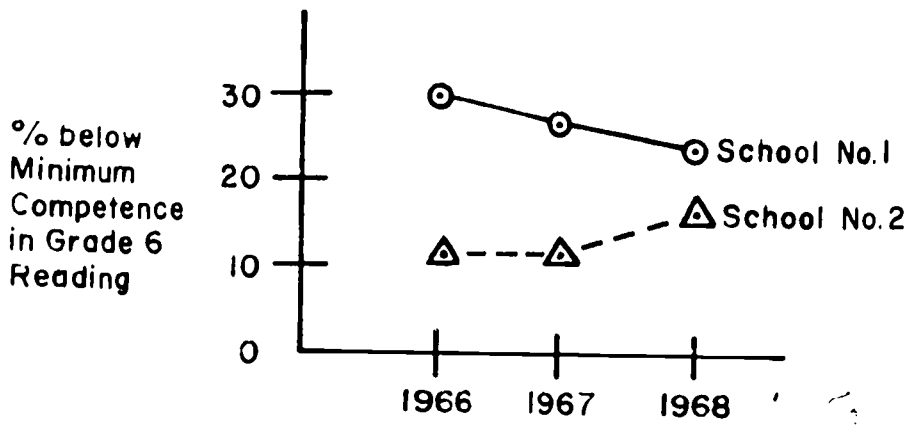


Fig. 4. Subskill Analysis

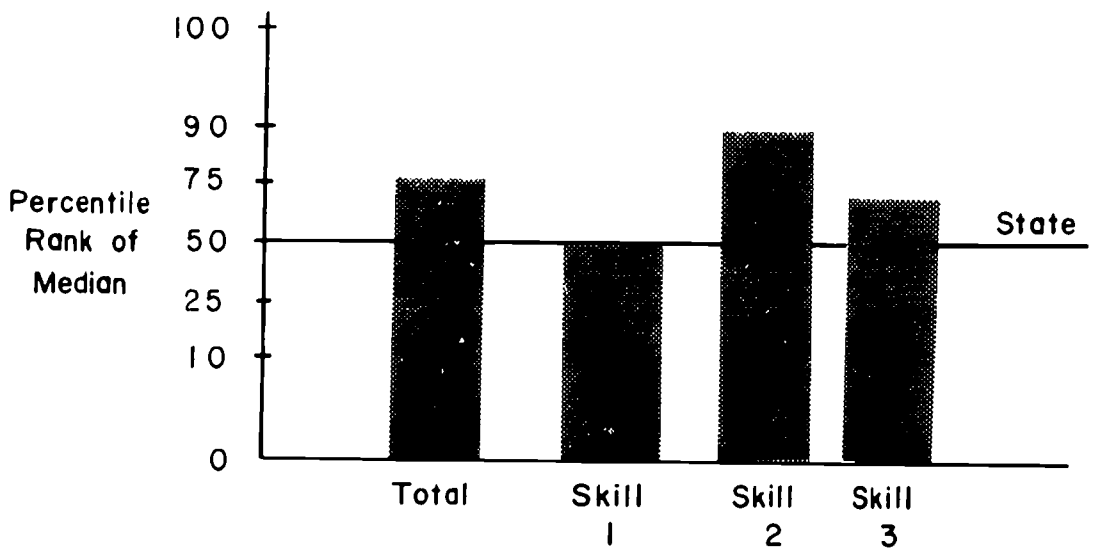


Fig. 5. Schools with Problems

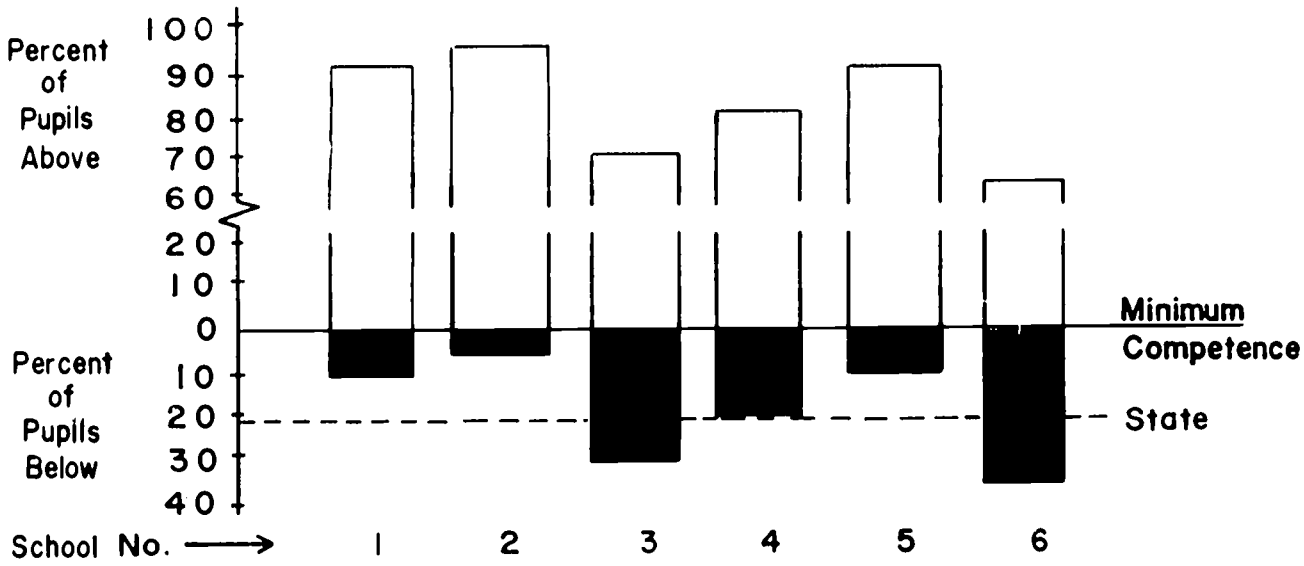
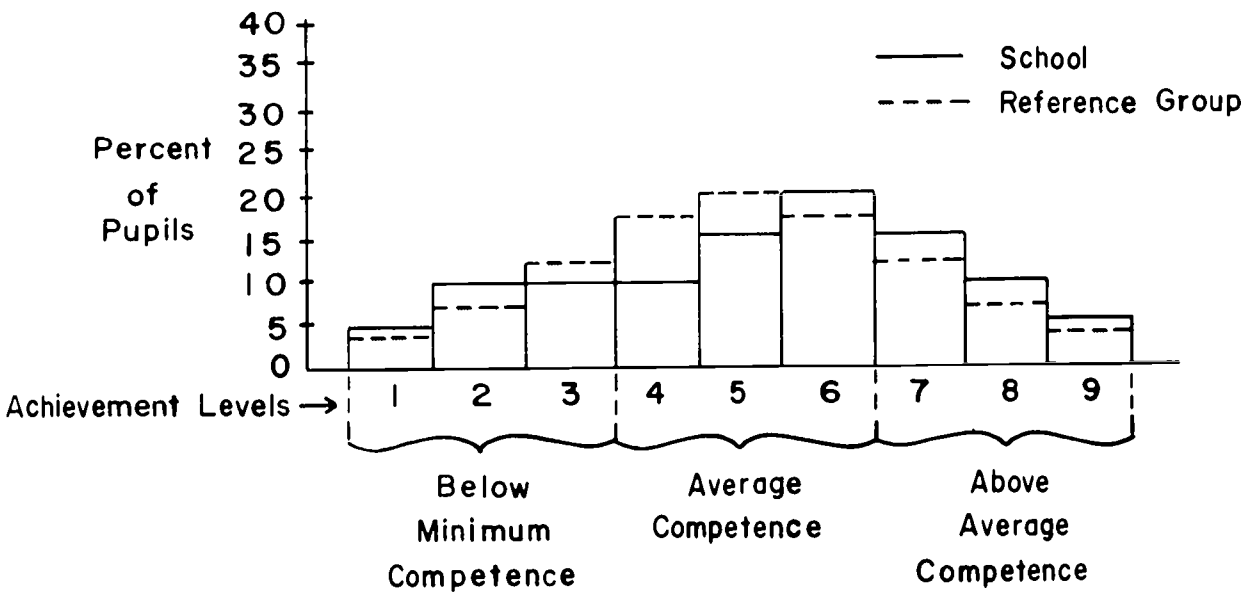


Fig. 6. Range of Achievement



Release of School Test Data

The State Education Department, as a matter of policy, does not make a general release of individual school or school district data. Only statewide and area summaries are published. All requests from public information media or from parents for individual school test data are referred to the local principal or superintendent, on the basis that only persons thoroughly acquainted with the local school situation are in a position to provide the interpretation necessary to make the data meaningful.

Nevertheless, it must be borne in mind that all test data on file in the State Education Department relating to schools, whether public or private, are public records, by law. If a person requesting such information can demonstrate that such data are not otherwise reasonably available to him, the Department has no option but to make its files available for inspection. In such event, the Department will, of course, make every effort to explain the limitations of the data and the cautions that should be observed in interpreting them, and to notify the school district that the information has been made available. Further, individual school test results will in no case be made available prior to a general release date, which will be at least 8 weeks after the school reports have been sent to all schools in the State.

Under these circumstances, school officials will recognize the advantage of planning for the proper distribution and utilization of test results within their own school districts. By summarizing the significant outcomes and pinpointing the unmet pupil needs revealed by the test results, greater direction and meaning can be given to the efforts of teachers, supervisors, and school board members; and parents can be provided with a sound, objective basis for understanding school program needs.

Cautions in Interpreting Test Results

Test results, by themselves, do not indicate the quality or effectiveness of instruction, nor do they indicate the extent to which pupils are attaining many of the other important and generally accepted goals of education. The achievement of a single pupil or all the pupils in a school, a community, or the State is a result of the interaction of at least three types of factors:

> Educational Resources - the total environment in which the school or school system is located, including community aspirations, financial support, and other socioeconomic conditions.
>
> Teaching and Learning Setting - the appropriateness and quality of instruction, curriculum, supervision, organization, and other educational services provided by the school or school system.
>
> Pupil Potential - the physical, emotional, social, and mental characteristics of the pupil, including motivation, interests, readiness, attitudes, and abilities.

Test results, therefore, should never be the sole basis for evaluating teaching or the effectiveness of a school or a school program. Just as low test results should not be rationalized as being caused only by inadequacies in pupil potential, high test results should not be casually accepted as indication of an effective teaching and learning setting. In each particular school situation, all the factors affecting pupil achievement, including educational resources, must be carefully examined and related to the test results in order to develop a realistic basis for constructive action.

In addition, undue importance should not be attached to small differences between the results for a school or school system and those for a reference group. Schools with 100 or more pupils in a grade will find that differences in percentile ranks or cumulative percents based on median raw score differences as small as 1.5 raw score points are statistically significant. However, even though a difference may be statistically significant, it may not be large enough to be of any practical significance.

The decision as to the size of a difference that is of practical significance is a subjective matter. A well-endowed school may feel that smaller differences are of more concern than would a school in more disadvantaged circumstances. As a convenient rule of thumb, schools may wish to consider only a difference greater than 5 or 10 points in percentile rank or in cumulative percent to be worthy of practical concern.

Maintaining Records of Pupil Scores

Records of the scores on the tests of the New York State Pupil Evaluation Program should be kept in the pupil's cumulative record file. The raw scores, percentile ranks, and testing date should be recorded.

Example

Test	Raw Score	Statewide Percentile	Date
Grade 6 NYS Reading Total	22	10	10/70
Word Recognition	10	15	
Comprehension	12	10	
Grade 6 NYS Mathematics Total	23	25	10/70
Concepts	7	10	
Computation	8	35	
Problem Solving	8	35	

Help is Available

The content and the objectives measured by all tests in the program are determined by subject specialists in the Division of General Education, with the help of special committees of classroom teachers and other recognized authorities in the State. The staffs of the Bureau of Reading Education and the Bureau of Mathematics Education therefore can provide answers to any specific questions concerning test content, test validity, and measurement objectives that are not included in the test manuals. Call or write:

Jane B. Algozzine, Chief
Bureau of Reading Education
New York State Education Department
Albany, New York 12224
Area Code 518: 474-2885

Frank S. Hawthorne, Chief
Bureau of Mathematics Education
New York State Education Department
Albany, New York 12224
Area Code 518: 474-3900

General questions concerning the program and its administration, as well as requests for help in interpreting school results or in reviewing or improving summary reports, should be addressed to:

Jack A. Maybee, Chief
Bureau of Pupil Testing and Advisory Services
New York State Education Department
Albany, New York 12224
Area Code 518: 474-5099

The Bureau of Pupil Testing and Advisory Services would be glad to receive copies of reports summarizing the test results for the school district so they may share these with other school districts needing help. All comments and suggestions on how to make the program fit more appropriately into school building and school district needs would also be most helpful.

A P P E N D I C E S

I School District Test Policy Guidelines

II Detailed Administrative Procedures

III Tables of Norms

Appendix I: School District Test Policy Guidelines

Below are 12 brief statements that can be used by a school district as a start in expressing its philosophy concerning testing and the use of test results in policy form. These statements pertain specifically to the Pupil Evaluation Program, but can easily be expanded to include the total school testing program. By adopting, formally or informally, some or all of these statements (or modified or expanded versions of them), a school board can provide the district testing program with the kind of direction and leadership that will encourage proper interpretation and use of test results.

Role of Tests and Their Limitations

1. Tests measure the cumulative effect of prior educational experiences. These experiences begin at birth and are influenced by the home and the community environment as well as by formal inschool programs.

2. Test results reveal pupil and school program needs in respect to only those skills or content areas measured by the tests. Although highly valid for measuring carefully delimited achievement objectives, paper and pencil tests do not measure all or even most of the important and generally accepted goals of education.

3. Test results alone do not indicate program effectiveness or school quality; nor do comparisons of test results for individual schools or school systems reveal true differences in program effectiveness or school quality. Many conditions other than school programs determine pupil achievement, and these conditions vary considerably from one school or school district to another.

Organization and Administration of the Testing Program

4. Centralized coordination of the school district testing program is necessary if the district is to obtain maximum value from test results.

5. Tests must be carefully administered under controlled conditions. Failure to follow the verbatim instructions and to observe the specified time limits invalidates results.

6. Teachers must be thoroughly familiar with the testing procedures prior to the administration of the tests. Teachers should understand the purposes for which tests are administered and should foster the development of healthy attitudes toward testing among pupils.

7. Teachers must not prepare pupils specifically for test content by giving them special instruction (cramming). Test results are interpreted as if the test questions were a random sample of all the questions that could appropriately be asked of pupils and not as if they were questions on which pupils had been given specific drill.

Use and Dissemination of Results

8. Test results provide teachers with a basis for planning instruction relevant to pupil needs. After the tests have been administered, teachers may use the test content and the pupils' answers to strengthen any weaknesses in skills and concepts revealed by the test results.

9. Teachers and other school program personnel need an opportunity to discuss districtwide test results and can help interpret and relate them to other factors. Every member of the school district program staff should be aware of the achievement status of the pupils in the district and of the progress or lack of it in achieving districtwide instructional goals.

10. Meaningful interpretations of district test results are a very practical means of explaining to school district residents some of the basic educational accomplishments and needs of their children. As taxpayers and parents assume a greater role in the governance of education, the educator's job is to provide them with as much objective basic information about pupil achievement as possible and to make the extra effort needed to help them understand this information and use it intelligently.

11. Grouped test results are public records. As such, they must be made available, upon demand, to reporters or other persons having a reasonable interest or concern in such records. The public has a right to know how well or how poorly its children are achieving.

12. Principals and superintendents are the primary source for release of test data. Any release of test results, to be meaningful, must be accompanied by an interpretation which relates the results to all other factors in the school and community situation that affect pupil achievement, and only persons thoroughly acquainted with these factors are in a position to interpret the data properly.

Appendix II: Detailed Administrative Procedures

Ordering Test Materials

Order forms for test materials are distributed during the spring to chief school officers. On these order forms, each school specifies the exact quantities of test booklets, scoring keys, teachers' manuals, answer sheets, and class record sheets needed to administer the testing program the following fall. Schools also indicate the type of answer sheet (IBM 805, IBM 1230, or Digitek) needed and check one of three dates on which the test materials should be shipped (August 1, August 15, or September 1). The completed order forms are collected by the central office and returned to the Department in one complete package by June 1 so that processing and packaging can begin well in advance of the shipping dates.

In ordering test materials, schools should plan for a fall inventory of about 5 percent over the number of materials needed to test the estimated enrollment. This will reduce the need for supplementary orders caused by unanticipated increases in enrollment and small variations in the quantities packaged for shipment. If additional materials are needed, however, supplementary orders can be placed by writing or telephoning the Bureau of Pupil Testing and Advisory Services (Area Code 518: 474-8220).

Unless schools are notified to the contrary, all materials left over from the previous year's administration are usable, and schools should order only the quantities needed to restore inventories to the amounts required for the next fall's administration. The third-grade test booklets are consumable (pupils write their answers in the test booklets); therefore, only the unused third-grade booklets are suitable for use in the next administration. The test booklets for the sixth and ninth grades, however, are reusable from one year to the next and should be replaced only as copies become worn or marked up.

Each teacher administering the tests needs an up-to-date manual and set of scoring keys. A school should order a sufficient supply of manuals and keys to provide a set to each new teacher and to replace any copies that have been lost or mislaid during the year. Even if tests are scored by outside agencies, each teacher needs a scoring key to check on the accuracy with which the answer papers have been scored. The teacher should also be provided with enough class record sheets to keep a duplicate set for classroom use. This usually amounts to at least four sheets per teacher.

All materials for the Pupil Evaluation Program are provided to schools free of charge.

Directions for Administering the Tests

The test manuals contain complete directions for administering the tests. As for all standardized tests, there should be no variations from the uniform procedures described in detail in these manuals.

To avoid the risk of invalidating the test results, teachers administering the tests - especially those doing so for the first time - should look over the test booklets and study the directions for administering the tests well in advance of the testing date. Teachers should also have an opportunity to ask questions and clarify their understanding of the directions before attempting to administer the tests. Any changes in prescribed time limits or any distracting interruptions in the test administration process will seriously impair the comparability needed for proper interpretation of the test scores. Except for the ninth grade tests, which may be administered to large groups of pupils, all tests are to be administered to pupils by their teachers in regular classes or in groups no larger than class size.

Scoring the Tests

The directions for scoring the tests are provided in the test manuals. The third-grade pupils mark their answers in the test booklets. The sixth-grade and ninth-grade tests have separate answer sheets which may be scored either by hand or by IBM or Digitek equipment. In addition to keys and stencils, tables of correct answers are provided for use by scoring machine operators. Test scoring is a school or school system responsibility.

Reporting the Test Scores to the Department

In general, schools combine the scores of individual pupils into frequency distributions, enter these distributions on score distribution report forms, and forward the forms to the Department where the information is translated into computer-readable form by optical scanning machines.

Because these machines are sensitive, or sometimes insensitive, to variations in the way the forms are marked, the directions below must be followed exactly. One small item of information omitted or one seemingly insignificant variation in gridding can cause untold man-hours to be spent in locating and correcting the source of trouble, or, worse yet, can go unnoticed and result in negating many hours of pupil, school, and Department time by producing erroneous data.

School Information Sheet Procedures

Complete one copy per school of the School Information Sheet (printed in blue ink) for return to the Department, according to the following procedures:

1. Print the information requested in the heading spaces provided in the upper left-hand corner of the form.

2. Write the appropriate numbers in each number box (blank space at top of each grid) for (1) number of Pupils Exempted, (2) School Enrollment, (3) School Code Number, and (4) number of completed Score Distribution Reports (SDR's) that are forwarded with the School Information Sheet.

3. Be sure to fill in each number box. For example, for zero, write 00; for 1, 2, 3, etc., write 01, 02, 03, etc.

4. For School Code Number, public schools should use the BEDS program code number for the first 12 digits and the Pupil Evaluation Program community type classification 71 through 77 for the last two digits.

5. When all the number boxes are filled in, use a No. 2 soft-lead pencil to make a heavy horizontal mark in each column on each grid in the space corresponding to the number in the number box.

6. Doublecheck the numbers in every number box and the accuracy with which they are gridded. Every number column on each and every grid should have one and only one grid mark in it.

Example of a Completed Grid for School Code Number

School Code Number													
CO		CITY		DIST		DIST KIND		SCHOOL				CMTY TYPE	
0	4	0	1	0	7	1	6	0	0	0	3	7	2
	0		0		0	0	0				0	0	0
1	1	1		1	1		1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	
3	3	3	3	3	3	3	3	3	3	3		3	3
4		4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6		6	6	6	6	6	6
7	7	7	7	7		7	7	7	7	7	7		7
8	8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9	9

Score Distribution Report Procedures

Complete one form per school for each grade tested. Detailed procedures for collecting and compiling the test data and for completing the Score Distribution Report forms are given below:

1. Have each teacher in the grades tested fill out the class record sheet included with the test materials. One side provides spaces for the name of each pupil and his scores. Enter the raw scores only. Extra sheets will be needed if there are more than 30 pupils in the class and if the teacher is to make a duplicate copy for her file and reference. The other side (entitled "Distribution of Raw Scores") provides places for tallying the part and total scores and for converting the tallies to frequency counts (in the F column).

Example of a Completed Distribution of Raw Score Table

SCORE	TALLY	F
8-9	//	2
6-7	~~HH~~ //	7
4-5	~~HH~~	5
2-3	///	3
0-1	/	1

2. Collect the Class Record Sheets and combine the distributions separately for each part and total into single schoolwide distributions for each test in each grade by adding the numbers (frequencies) at each score interval on one class record sheet with those in the same intervals on the other class record sheets. Enter the sums of the frequencies on blank copies of the Distribution of Raw Score Tables so as to produce a single set of schoolwide tables. Doublecheck for errors in addition and in copying.

3. On the Score Distribution Report forms, print the information requested in the heading box on the front, upper right-hand side of the forms and fill in the number boxes for the school code number.

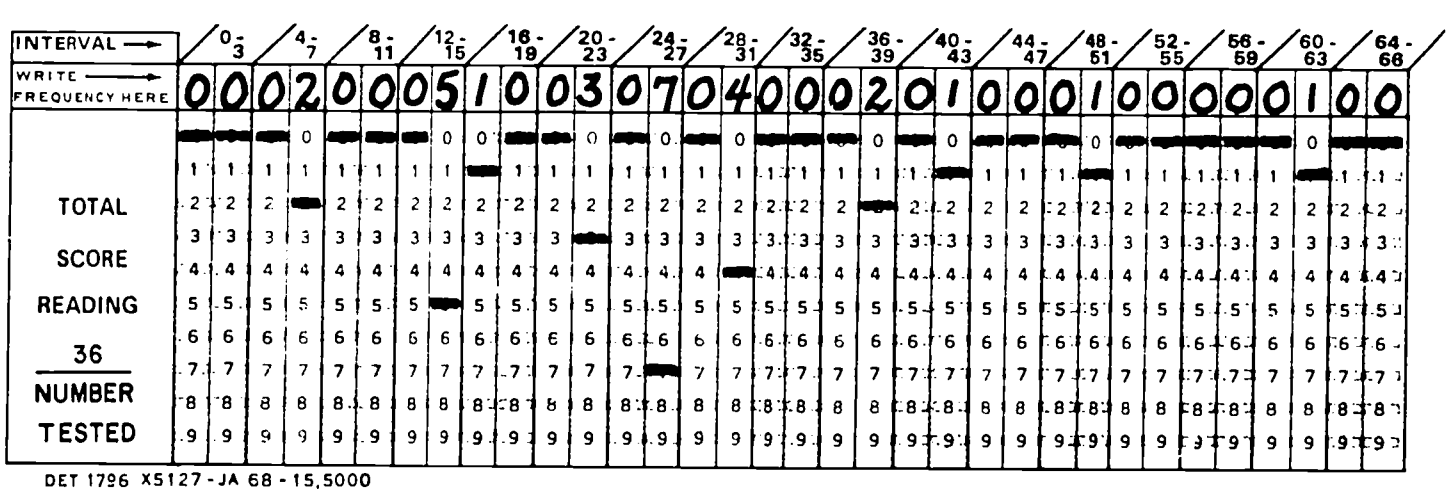
4. Enter the number of pupils tested for each part and the total test in the appropriate spaces.

5. Copy the numbers from the F columns of the schoolwide Distribution of Raw Score tables into the corresponding number boxes (blank spaces at top of each grid). Enter the numbers (frequencies) for each subtest and total just as they appear on the schoolwide tables. Do not add the numbers in the subtest boxes to obtain those for total score.

6. Be sure to use the proper form for each grade (red for grade 3, brown for grade 6, green for grade 9) and to fill in each and every number box. Write each of the frequencies as a two digit number (00 for 0, 01 for 1, 02 for 2, etc.)

7. Using a No. 2 soft-lead pencil, make a heavy horizontal mark in each column on each grid in the space corresponding to the number in the number box. Erase carefully and completely if it is necessary to make corrections. Do not use ink or colored pencil.

Example of a Completed Grid for Sixth-Grade Total Reading Score



8. Grid in the School Code Number on each report, if you have not already done so.

9. Doublecheck all copying and gridding work. There should be one and only one mark in each column on each grid.

Additional Points of Information

1. Do not fold or crease the forms. If a form is damaged or filled out wrong, complete a new one.

2. Fill in all numbers in the number boxes before attempting to grid them. Fewer mistakes will occur when the gridding is done as a separate operation.

3. For pupils who take part of a test, but not the whole test, report their scores for the part of the test taken. Report total test scores only for pupils who take all of the subtests.

4. Use a second form to report any score interval frequencies over 99. For example, a large school with 120 pupils at one score interval should report 60 in that interval on each of two forms or 99 on one form and 21 on the other, or use any other convenient set of two digit numbers totaling 120. If there are 240 pupils, the principle is the same, using three forms instead of two. Do not split any two digit frequencies, however; report all two digit frequencies on one form and report 00 for these intervals on the other forms.

5. If all ninth-grade pupils in a school are excused from taking the tests, complete the School Information Sheet and a Score Distribution Report for the ninth grade. In this case, the number of pupils reported as enrolled will be the same as the number excused.

6. Report as zero the scores of any non-English speaking pupils who are unable to take the tests because of their language handicaps.

7. In returning the completed forms to the central office or to the Department, do not fold the forms. Insert one or more pieces of cardboard in the envelope so that the forms do not become damaged in the mail.

8. Send all forms to the Department by first class mail. Mail all the report forms for the schools in a public school system together in one envelope. Do not send school results singly or separately, but hold all forms until those for all schools in the system are ready for mailing. Send the forms to:

Pupil Evaluation Program
Bureau of Pupil Testing and Advisory Services
New York State Education Department
Albany, New York 12224

Unused test Materials

Except when new forms of a test are developed and introduced into the program, materials left over from one year's administration can be used the following year. Also, the same sixth-grade and ninth-grade test booklets can be reused until they are worn out or marked up. Immediately after completion of the test administration, therefore, all unused and reusable materials should be collected, checked for marked answers, inventoried, and stored. Principals will find an accurate inventory of materials extremely helpful the following spring when they make out orders for next year's program. No test materials, used or unused, should be returned to the Department, unless a specific request is made for them.

APPENDIX III: TABLES OF NORMS

TABLE 1: READING TEST - BEGINNING GRADE 3 - FORM A

A. PERCENTILE RANKS

PERCENTILE RANK	WORD RECOGNITION	READING COMPREHENSION	TOTAL SCORE	PERCENTILE RANK
99	...	27–28	52–53	99
95	25	26	49–51	95
90	24	25	47–48	90
85	23	24	46	85
80	...	22–23	44–45	80
75	22	21	42–43	75
70	21	20	40–41	70
65	20	19	38–39	65
60	19	18	36–37	60
55	...	16–17	34–35	55
50	18	15	32–33	50
45	17	14	30–31	45
40	15–16	13	28–29	40
35	14	12	26–27	35
30	13	11	24–25	30
25	12	10	21–23	25
20	10–11	9	19–20	20
15	8– 9	8	16–18	15
10	6– 7	6– 7	13–15	10
5	4– 5	4– 5	9–12	5
1	0– 3	0– 3	0– 8	1

B. ACHIEVEMENT LEVELS

ACHIEVEMENT LEVEL	WORD RECOGNITION	READING COMPREHENSION	TOTAL SCORE	CUMULATIVE PERCENT
9	25	27–28	51–53	100
8	24	25–26	48–50	96
7	23	22–24	44–47	89
6	20–22	18–21	37–43	77
5	16–19	13–17	29–36	60
4	12–15	10–12	21–28	40
3	8–11	7– 9	15–20	23
2	5– 7	5– 6	11–14	11
1	0– 4	0– 4	0–10	4

TABLE 2: READING TEST - BEGINNING GRADE 6 - FORM A

A. PERCENTILE RANKS

PERCENTILE RANK	WORD RECOGNITION	READING COMPREHENSION	TOTAL SCORE	PERCENTILE RANK
99	30	35–36	63–66	99
95	29	34	60–62	95
90	28	33	58–59	90
85	27	32	57	85
80	26	31	55–56	80
75	25	30	53–54	75
70	24	29	51–52	70
65	23	28	50	65
60	22	27	48–49	60
55	21	26	46–47	55
50	20	24–25	44–45	50
45	19	23	41–43	45
40	17–18	22	39–40	40
35	16	20–21	36–38	35
30	15	18–19	34–35	30
25	13–14	17	31–33	25
20	12	15–16	27–30	20
15	10–11	13–14	24–26	15
10	8– 9	11–12	20–23	10
5	5– 7	7–10	14–19	5
1	0– 4	0– 6	0–13	1

B. ACHIEVEMENT LEVELS

ACHIEVEMENT LEVEL	WORD RECOGNITION	READING COMPREHENSION	TOTAL SCORE	CUMULATIVE PERCENT
9	30	35–36	62–66	100
8	28–29	33–34	59–61	96
7	26–27	31–32	55–58	89
6	22–25	27–30	49–54	77
5	18–21	22–26	40–48	60
4	13–17	17–21	31–39	40
3	9–12	12–16	22–30	23
2	6– 8	9–11	16–21	11
1	0– 5	0– 8	0–15	4

TABLE 3: MATHEMATICS TESTS - BEGINNING GRADE 3 - FORM C

A. PERCENTILE RANKS

Percentile Rank	Concepts	Computation	Problem Solving	Total Score	Percentile Rank
99	24-26	15-16	18	54-60	99
95	22-23	13-14	-	50-53	95
90	20-21	12	17	47-49	90
85	19	11	16	45-46	85
80	18	10	-	42-44	80
75	17	9	15	40-41	75
70	16	-	14	38-39	70
65	15	8	-	36-37	65
60	-	7	13	33-35	60
55	14	-	12	31-32	55
50	13	6	11	28-30	50
45	12	5	10	26-27	45
40	11	-	9	24-25	40
35	10	4	8	22-23	35
30	-	-	7	20-21	30
25	9	3	6	18-19	25
20	8	2	5	14-17	20
15	6-7	-	3-4	12-13	15
10	5	1	2	9-11	10
5	2-4	0	1	4-8	5
1	0-1	-	0	0-3	1

B. ACHIEVEMENT LEVELS

Achievement Level	Concepts	Computation	Problem Solving	Total Score	Cumulative Percent
9	25-26	14-16	18	53-60	100
8	21-22	12-13	17	48-52	96
7	18-20	10-11	16	42-47	89
6	15-17	7-9	13-15	35-41	77
5	12-14	5-6	9-12	25-34	60
4	9-11	3-4	6-8	18-24	40
3	6-8	1-2	3-5	11-17	23
2	3-5	0	2	6-10	11
1	0-2	-	0-1	0-5	4

TABLE 4: MATHEMATICS TEST - BEGINNING GRADE 6 - FORM C

A. PERCENTILE RANKS

Percentile Rank	Concepts	Computation	Problem Solving	Total Score	Percentile Rank
99	24-27	19-20	19-20	60-67	99
95	22-23	17-18	18	53-59	95
90	20-21	16	17	50-52	90
85	19	15	16	47-49	85
80	18	14	15	45-46	80
75	17	13	14	42-44	75
70	16	12	13	41	70
65	15	-	-	38-40	65
60	-	11	12	36-37	60
55	14	-	11	34-35	55
50	13	10	10	33	50
45	-	9	-	31-32	45
40	12	-	9	29-30	40
35	11	8	8	27-28	35
30	10	-	7	25-26	30
25	-	7	-	23-24	25
20	9	6	6	21-22	20
15	8	5	5	19-20	15
10	7	-	4	17-18	10
5	5-6	3-4	3	13-16	5
1	0-4	0-2	0-2	0-12	1

B. ACHIEVEMENT LEVELS

Achievement Level	Concepts	Computation	Problem Solving	Total Score	Cumulative Percent
9	23-27	18-20	19-20	58-67	100
8	21-22	16-17	17-18	51-57	96
7	18-20	14-15	15-16	44-50	89
6	15-17	11-13	12-14	37-43	77
5	12-14	9-10	9-11	30-36	60
4	10-11	7-8	7-8	23-29	40
3	8-9	5-6	5-6	18-22	22
2	6-7	4	3-4	15-17	11
1	0-5	0-3	0-2	0-14	4