DOCUMENT RESUME

ED 108 971 SE 019 476

AUTHOR Bernstein, Margery R.
TITLE Evaluation of Elementary Math Program.
INSTITUTION Mamaroneck Public Schools, N.Y.
PUB DATE Jun 73
NOTE 12p.

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS *Criterion Referenced Tests; *Curriculum Evaluation; Elementary Education; *Elementary School Mathematics; Evaluation; Geometry; *Number Concepts; Problem Solving; Test Construction; *Testing Programs

ABSTRACT

A criterion-referenced test battery was devised for use with the mathematics program of the Mamaroneck Public Schools (New York) in 1971. After two annual administrations, the tests were revised in four ways. The format was changed, and the number of items dealing with geometry was increased; some items were revised to conform better with the curriculum and others were modified in the interests of clarity. Results on the revised tests are compared with those on the original forms. (SD)

EVALUATION

OF

ELEMENTARY MATH PROGRAM

Margery R. Bernstein

Program Evaluation & Research
Mamaroneck Public Schools

June 1973

PROGRAM EVALUATION AND RESEARCH DEPARTMENT
Mamaroneck Public Schools

Dr. Margery R. Bernstein

ELEMENTARY MATH PROGRAM

Brief History

The elementary math program was developed in the Mamaroneck schools over a period of two years and adopted on a systemwide basis for all elementary grades in September 1970. A criterion-referenced test was developed by the math steering committee. The six areas covered were as follows:

Number and numeration
Sets, language and symbols
Operations whole numbers
Operations fractions
Problem solving
Geometry and measurement

The test was administered in May 1971 and each year thereafter as follows:

May 1971 - A random sample of about 1/6 of the pupils at each grade level was tested.

May 1972 - All pupils were tested with each pupil taking only one of five forms of the test, each of which contained one-fifth of the original items.

May 1973 - All pupils were tested on all items.

The May 1973 test was revised in four ways. 1) The format was improved. 2) Items were added to improve the coverage of geometry. In order to make a comparison with previous years possible, these additional items were tabulated separately. 3) Some items were changed or reworded to make them clearer. 4) Some items were changed to conform better with the curriculum.

The changes which affected the difficulty of the items were in the grades and areas indicated below. The percents of the scores affected are noted in parentheses.

Grade 1: Sets, language and Symbols (22%)
Operations on fractions (50%)
Problem Solving (40%)
Geometry (10%)

Grade 2: Operations on Whole Numbers (2%)

Grade 5: Operations on fractions (4%)
Problem Solving (29%)

Grade 6: Numbers and Numeration (6%)
Operations on Whole Numbers (3%)
Operations on Fractions (4%)
Problem Solving (37%)
Geometry and Measurement (12%)

Thus comparisons in these areas with previous years' results must be made with caution. The changes are further detailed in Table 3 in the Appendix.

Results

The Criterion-Referenced Tests were administered during the period of May 21-23, 1973. Table 1 indicates the percentages of pupils tested. (See Table 4 in the Appendix for numbers of students enrolled and numbers tested.)

TABLE 1

Percentage of Enrolled Students Tested

Grade	Central	Chats	Mam'k	Murray	Total
1	98	92	96	96	95
2	98	93	91	98	95
3	92	98	91	96	95
4	93	94	91	97	94
5	93	93	95	94	94
6	94	99	93	98	96

Table 2 indicates the percentage of correct responses in each of the math curriculum areas by grade for each year the test was given systemwide. Each principal and math committee member has received a similar table for his school. Circled percentages indicate the areas where items were revised. The data for 1973 only is shown in Figure 1.

TABLE 2

Math Criterion-Referenced Test
Percents of Correct Responses for 1971, 72 & 73

Gr.	Numbers, Numeration			Sets, Lang. Symbols			Op. Whole Numbers			Op. Fractions			Problem Solving			Geom. & Measure. Old			Geom. & Measure. Comb. Old & New
	71	72	73	71	72	73	71	72	73	71	72	73	71	72	73	71	72	73	73
1	94	88	90	87	84	(80)	89	76	81	79	71	(73)	68	75	(83)	81	74	(83)	86
2	90	90	94	91	83	95	94	81	(83)	83	82	88	77	86	83	86	87	91	
3	91	85	92	84	87	90	91	76	77	73	60	68	54	63	64	86	79	83	
4	84	72	84	74	69	76	86	78	80	64	63	73	42	44	52	35	45	69	
5	64	70	75	70	73	79	83	71	75	69	62	(65)	50	61	(73)	20	51	70	72
6	68	51	(68)	81	70	76	86	72	(79)	78	55	(70)	48	47	(66)	33	51	(59)	65

If a comparison is made between the 1972 and 1973 results, there are 36 possible comparisons. Of these, substantial improvement defined as an increase of 10% was made in the following areas:

Grade 2: Sets, Language and Symbols

Grade 4: Numbers and Numeration
Operations on Fractions
Geometry and Measurement

Grade 5: Geometry and Measurement

There was a 10% improvement in four additional areas (5th grade: Problem Solving; 6th grade: Numbers and Numerations, Operations on Fractions, and Problem Solving) but these are parts of the tests where adjustments in items were judged to have made the tests somewhat easier.

It must be noted, however, that teachers considered the 1972 test format with item sampling to have been confusing and difficult to administer. In

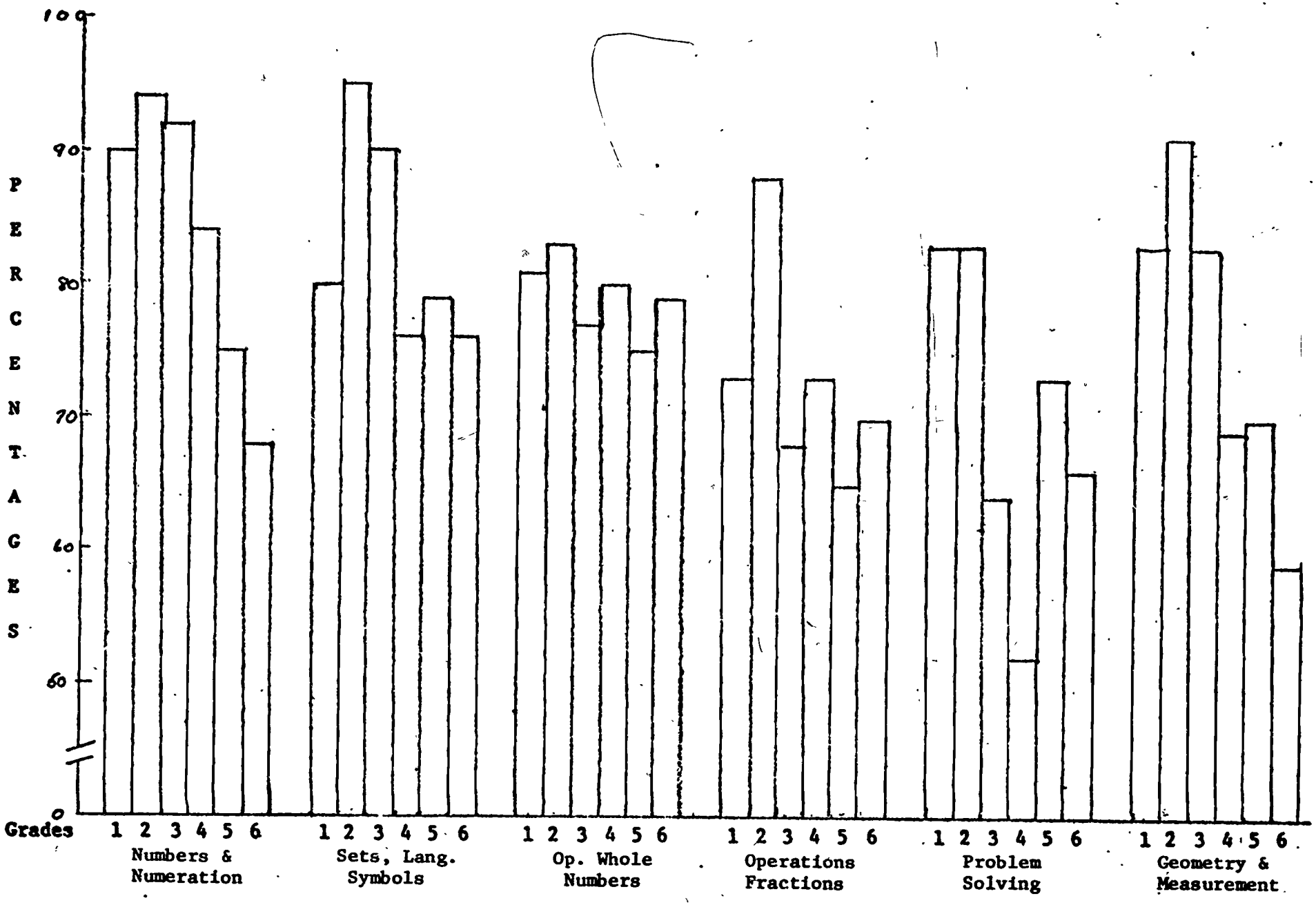

Figure 1. Percentages of correct responses, Math Criterion-Referenced Tests, Grades 1 - 6, May 1973

the judgment of the math committee this factor served to lower test scores for 1972. A comparison was therefore made between 1971 and 1973 results.

Substantial changes, defined as above, were made in the following areas:

Decreases:

Grade 2: Operations on Whole Numbers (Item changes made the 1973 easier and so did not make this change less meaningful).
Grade 3: Operations on Whole Numbers

Increases:

Grade 3: Problem Solving
Grade 4: Problem Solving
Geometry and Measurement
Grade 5: Numbers and Numeration
Geometry and Measurement

In addition there were increases in scores in five areas where items were made easier. An illustration might be helpful here. In the 1st Grade Problem Solving test, the 1971-72 test had two problems to be solved where the addition facts involved required the use of numbers higher than those included in the Operations on Whole Numbers part of the curriculum (7 + 15 = 12). The corrections made these questions better tests of children's problem-solving ability.

It is noteworthy that there was a decrease in the percentage of correct responses in the category of Operations on Whole Numbers in every grade even though the lowest percentage correct was 75. On the other hand, Problem Solving was seen to go up in every grade, even though results still show weaknesses from 3rd Grade on.

The math committee also studied 1972 and 1973 results for the same groups of pupils (e.g. Grade 1 in 1972 compared with Grade 2 in 1973). Although it was realized that the changes in the student body caused by pupils' moving in and out precluded the formulation of definite conclusions,

the results are indicated for the sake of the hypotheses they suggest. Thirty comparisons are possible. Ten percent changes occurred as follows:

Increases

Content Area	Grade in 1973
Sets, Language and Symbols	2, 5
Operations on Fractions	2, 4
Geometry and Measurement	2, 5

Decreases

Sets, Language and Symbols	4
Operations on Fractions	3
Problem Solving	3, 4
Geometry and Measurement	4

One other increase (an improvement in problem solving ability from Grade 4 in 1972 to Grade 5 in 1973) is of questionable validity because of changes in the test.

A similar comparison of roughly the same groups of pupils from 1971 to 1973 showed 11 substantial changes out of a possible 24.

Decreases

Numbers and Numeration	5, 6*
Sets, Language and Symbols	4
Operations on Whole Numbers	3, 4, 5
Operations on Fractions	3, 4
Problem Solving	4
Geometry and Measurement	4, 5

The three improvements noted (Problem Solving to Grades 5 and 6, Geometry and Measurement to Grade 6) may be attributable to changes in the test.

It should be noted that Grades 1 and 2 could not be included in this three-year study. Five of the 11 decreases occurred at the 4th Grade level.

* Change in the test does not detract from the meaningfulness of this change.

From the above, two hypotheses are suggested:

1) Computational skills are getting weaker;

2) There is an area of weakness in middle elementary grades.

Discussion

There are several problems involved with the continued use of the CRT's which need to be faced:

1) The CRT's were designed to determine how well the math curriculum is being learned. No criteria, however, have ever been set. The various math committees have worked hard to devise and revise a test which accurately reflects the Mamaroneck curriculum. They should now take the next step and find a way to set criterion levels. In doing so one need not set a uniform criterion level for all areas and grades. For example, Operations on Whole Numbers may be considered so crucial that a 90% criterion level is desired, while Geometry and Measurement may be considered less important and may therefore have a criterion level of 60% to 70%.

2) Attempts have been made this year to use the CRT's for purposes for which they were not intended.

a) Individual diagnosis. The tests cannot legitimately be used for this purpose, first, because of their administration at the end of the year, second, because there are too few items on each topic, and third, because their reliability has not been determined. Any conclusions about an individual child's performance based on the CRT results are indefensible since the standard error of measurement is unknown. It is therefore recommended that individual results not be retained in the cumulative records.

b) Program evaluation. When plans for evaluating some elementary programs were made, it seemed sensible to use the existing tests for this additional purpose. Such use is necessary but not sufficient for this purpose. An innovative

program might well set its own objectives in math and find suitable evaluative approaches. Thus, for example, a program whose aim was to encourage selected students to go ahead in math learnings as far as possible might also use appropriately selected norm-referenced tests at higher grade levels.

3. If the tests are used for their intended purpose, some sampling procedure should be used to cut down on the time spent by pupils in taking tests, by the staff in scoring them, and by clerks in tabulating the results. Sampling of pupils or test items or both would be statistically satisfactory.

It is recommended that each grade's CRT be divided into parallel short forms, four or five per grade. Each class teacher would then distribute the forms randomly to her students. The problems involved in giving the tests this way would be no greater than those involved when groups of pupils work on different tasks and could be readily resolved.

The Mamaroneck schools have made great strides in revising the math curriculum and devising a way of determining the extent to which it is being learned. It should not now lose the momentum gained by diluting the value of the test through inefficient use.

6/73
MRB

TABLE 3

SUMMARY OF CHANGES IN 1973 REVISIONS

Grade	Changes in form only	Items made harder	Items made easier	Items Added
1	6	1 (S2/18)*	5 (G1/10) (F2/4) (P2/5) (S2/18)	1 (G3/10)
2	1	-	1 (W1/55)	-
3	9	-	-	-
4	6	-	-	-
5	10	-	2 (P2/7) (F1/23)	2 (G3/6)
6	12	-	7 (N1/17) (W1/35) (F1/28) (P3/8) (G1/8)	2 (G3/8)

* The parentheses contain the code letter for the section of the test and the number of value points changed/total value points in the section. Example: on the first grade test, one question was made harder; it involved two of the 18 points in the "Sets, language and symbol" section.

Code letters

Number and Numeration (N)
Sets, Language and Symbols (S)
Op. on Whole Numbers (W)
Op. on Fractions (F)
Problem Solving (P)
Geometry and Measurement (G)

TABLE 4

Pupils Tested

Grade	CE		CH		MK		MU		TOTAL	
	N Enrolled	N Tested	N Enrolled	N Tested	N Enrolled	N Tested	N Enrolled	N Tested	N Enrolled	N Tested
1	87	85	98	90	68	65	98	94	351	334
2	98	96	107	99	75	68	104	102	384	365
3	102	94	124	121	64	58	104	100	394	373
4	123	115	129	121	92	84	125	121	469	441
5	129	120	108	100	86	82	127	119	450	421
6	126	118	144	143	82	76	128	125	480	462
	665	628	710	674	467	433	686	661	2528	2396