DOCUMENT RESUME

ED 127 131 SE 020 685

AUTHOR	Zoet, Charles J.
TITLE	Michigan Fourth and Seventh Grade Student Performance in Mathematics, 1970-1973, Monograph No. 6.
INSTITUTION	Michigan Council of Teachers of Mathematics.
SPONS AGENCY	Michigan Education Association, East Lansing.
PUB DATE	Oct 74
NOTE	92p.; For monographs 5-9 of this series, see SE020684-688
AVAILABLE FROM	MCTM, Box 16124, Lansing, Michigan 48901 ($1.50 each, prepaid)
EDRS PRICE	MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS	*Achievement; Curriculum; *Educational Assessment; Elementary School Mathematics; Elementary Secondary Education; *Evaluation; *Instruction; *Mathematics Education; Secondary School Mathematics; State Programs; Testing Programs; Test Interpretation
IDENTIFIERS	*Michigan

ABSTRACT

This monograph, addressed primarily to Michigan teachers, presents an analysis of the item data from the Michigan Educational Assessment Program (MEAP) study of fourth and seventh graders. Summary data related to construction and administration of the tests are presented, and reliability, validity and equating of tests are discussed. Items were classified as belonging to one of the categories: number and operations; geometry and measurement; relations, functions, and graphs; logical thinking; mathematical sentences; and applications. Items within each category are discussed with regard to a number of indices of difficulty level and correlation with overall test scores. Performance on items within categories is compared and possible sources of difficulty are suggested. Summarizing his findings, the author opines that the findings identify a number of critical points which need attention from school districts and teachers, but that a criterion referenced instrument might provide more useful information. (SD)

THE MICHIGAN COUNCIL OF TEACHERS OF MATHEMATICS

M
C
T
M

Guidelines for Quality Mathematics Teaching

A Monograph Series

MICHIGAN FOURTH AND SEVENTH GRADE STUDENT PERFORMANCE IN MATHEMATICS 1970-1973

Charles J. Zoet

Monograph No. 6
October 1974

Additional copies of this monograph can be obtained at $1.50 each (prepaid) by writing to:

MCTM
BOX 16124
LANSING MICH 48901

Make checks payable to Michigan Council of Teachers of Mathematics (M.C.T.M.).

Other Monographs available from same address:

1. Classroom Proven Motivational Mathematics Games ($.60)
2. An Introduction to the Minimum Performance Objectives for Mathematics Education in Michigan ($1.00)
3. Order in Number ($1.00)
4. Metric Measurement Activity Cards ($2.00)
5. An Activity Approach to Fractional Concepts ($1.00).

MICHIGAN FOURTH AND SEVENTH GRADE STUDENT PERFORMANCE IN MATHEMATICS 1970-1973

Charles J. Zoet

7

TABLE OF CONTENTS

5

FOREWORD

Since 1969, the Michigan Council of Teachers of Mathematics has steadily maintained its interest and involvement in the Michigan Educational Assessment Program. Knowing that the state legislature and the department of education were bent toward an extensive assessment program, the MCTM proposed to make every effort to assist in assuring that the program would make a positive contribution to mathematics instruction.

The MCTM recognizes the need for a more objective basis for decision-making in education. It has continuously assisted in efforts to identify minimal objectives and develop test items. This monograph is a part of this effort to provide for better local and state-wide determination of mathematics curricula.

Dr. Charles Zoet has done an excellent job in sifting the 1970-73 mathematics assessment data for implications. His considerable experience and talents have enabled him to make many insightful observations.

Terrence G. Coburn

TO THE READER

I would like to take this opportunity to alert the reader to the nature of the analysis presented in this document. The analysis has assumed that the primary group of readers will be teachers of mathematics and others interested in improving mathematics programs. Therefore, the conjectures and observations presented are those which were thought to be of interest to that group.

While a great amount of data was presented for this analysis, there was little opportunity to assemble comparative statistics which fit a classical research model. Since the results of classical research in the educational field frequently conflict and since it was believed that such research would not be of primary interest to the readers, I have worked freely with some very simple data. Substantial statistical data could be assembled to support some of the findings; with others, I have simply drawn on my background and experience to present conjectures in which I may be completely off base. The context is such in those cases that it should be clear that I am drawing heavily on my own perceptions.

With so much data available on so many concepts, it did not seem to me that there was any great virtue in timidity. At the very least it is my hope that the document does stimulate interest in the development of mathematical concepts in young students.

While the MCTM has permitted me a great amount of freedom in developing the nature of this monograph a number of people did contribute substantially to the final document. First of all, Mr. Jules Shrage of Oakland Schools did some preliminary organization and analysis of the data which was most useful in narrowing the focus of this analysis. Dr. Tom Fisher and Dr. Robert Huyser of the Michigan Department of Education, Doctors William Swart and Jim Bidwell of Central Michigan University and Dr. Terry Coburn of Oakland Schools all submitted extensive and helpful editorial comments on the first draft of the document and assisted in bringing it to its final form. I am grateful to all of them for their help.

Finally whatever ability I ~~may~~ have to see instructional implications in data of this nature has been cultivated through my work in Livonia Public Schools--a district which believes that critical analysis of the results of instruction is absolutely necessary to the continued improvement of programs. They supported their effort in many ways and I am particularly grateful that I have had the opportunity to spend my professional years in such a setting.

Charles J. Zoet
May 1974

INTRODUCTION

The purpose of this monograph is to analyze the results of student performance on the mathematics portion of the 4th and 7th grade tests administered by the Michigan Department of Education early in 1970, 1971, 1972, and 1973 and to identify information which may be useful to mathematics teachers in Michigan.

Achievements in mathematics by Michigan students are now being assessed through objective oriented tests. Therefore no recommendations are made here for changes in past testing programs. A later monograph will deal with information coming from the new objective oriented program. The major focus of this analysis will be on the performance of Michigan students on individual test items.

Under the Michigan Educational Assessment Program (MEAP) tests covering a variety of basic skills in language arts and mathematics were administered to nearly all Michigan 4th and 7th graders during the school years of 1969-70 to 1972-73. These achievement tests were prepared by the Educational Testing Service under contract to and in cooperation with the MDE. Instruments designed to measure a variety of socio-economic characteristics of school districts (and other organizational units) accompanied some of these tests.

PURPOSE AND DIRECTION OF MEAP

The decision-making and implementing process at the state level is an extremely complex mix of professional judgment, political convenience, economic fluctuations, personalities, practical necessity and historical deadlines. While MEAP has become part of a six-step, long-range plan to manage Michigan schools rationally, the commitments necessary to such long-range management have not yet been made.

In order to avoid being ensnared by the intricacies of the politics involved let us simply recognize that the same financial problems and absence of long-range commitment which enmesh local school units extend to the state Department of Education and complicate attempts to develop more effective schools. Readers interested in critical analysis of MEAP testing programs are referred to critiques such as these by the Michigan Association of Professors of Educational Administration (23) and Oakland County (24) as well as the MDE responses (11,12).

In August, 1969, the State Superintendent of Public Instruction introduced the initial thrust of the Michigan Educational Assessment Program emphasizing that:

> "The full implementation of a meaningful assessment program will not be achieved in the period of one year. Nor will it be achieved without the cooperation and involvement of professional educators and lay citizens. The task at hand is a complex one and will necessitate systematic planning and development over

a period of many months. The activities which will be undertaken during the 1969-70 school year represent only a beginning step in a long-range program designed to provide better and more comprehensive information concerning the level, distribution, and the progress of education in the schools of our state." (1, Foreword)

The evolution of MEAP, projected in that paragraph, was apparent in a year-to-year shift of emphasis from better allocation of monies, to accountability to parents, students, school boards and the state; and to assessment of education needs to implement the MDE six-point program.

However, the major objectives for MEAP remained consistent through the four years of testing examined in this analysis. Briefly stated they were

(1) to provide the state level decision-makers with information for use in allocating state resources.

(2) to provide local decision-makers with information regarding their schools for use in the design of future educational programs.

(3) to provide basic information about students to help assess their progress and to identify students (and districts?) who have special need of assistance.

(4) to provide all interested people with information regarding the progress of education statewide and in individual schools or school districts.

As it applied to basic skills in mathematics, the assessment program strove to answer three questions:

--Is the level and distribution of basic skills improving among the state's school districts?

--Is the level and distribution of basic skills improving within a particular school district?

--What is the achievement level in mathematics in Michigan and its school districts?

In order to find a starting point which would give some breadth to their look at educational inputs, basic skills, etc., the MDE elected to begin its testing at the 4th and 7th grade levels. While there have been some indications that a further expansion would include grade ten, a firm commitment to extend the program was placed in reserve.

The limited amount of useful instructional information and inconclusiveness of the test results during these years, represent the pitfalls of attempting to derive data for the improvement of instruction from a testing program not designed for that purpose. The tests involved literally millions of hours of student and professional time but frequently by-passed mathematical concepts which are critical at the level tested. The limitations are more sobering when viewed against what might have been.

Those remarks are intended to be critical of the educational decision-making process and not of attempts to evaluate programs. The latter is probably essential to statewide improvement of education. However, it is very difficult to conduct a meaningful evaluation in a sociological arena even when there is a clear, long-range commitment to getting accurate and useful curriculor data.

The absence of a specific long-range design for that purpose simply increases the difficulty.

THE MATHEMATICS TESTS

The mathematics tests with which this monograph is concerned were administered to 4th and 7th graders during the month of January in 1970, 71, 72, and 73. The following table summarizes the number of students and school districts tested during those ~~years~~s.

	1969-70		1970-71	
	4th	7th	4th	7th
Number of Districts	585	566	612	584
Schools	2,492	909	2,529	902
Pupils	158,713	159,407	161,599	162,585

	1971-72		1972-73	
	4th	7th	4th	7th
Number of Districts	608	581	604*	577*
Schools	2,485	917	2,470*	911*
Pupils	162,280	164,601	157,854	160,200

*Approximated from the data available.

Educational Testing Service (ETS), Princeton, New Jersey was assigned prime responsibility for developing the tests to be used in assessing basic skills in mathematics. The MDE supervised the contract for this development and frequently consulted with ETS as the work developed. This four year testing period began with a crash program authorized by the legislature in August 1969 for implimentation in January 1970. Since the initial reaction to that program the MDE has continually looked for ways to involve the educational community in decisions regarding the testing program. By 1972-73 a MEAP advisory council of eight members representing the educators of the state had been formed to advise the MDE.

Specifications

The construction of a test begins with a set of specifications describing the objectives to be tested. In this case:

> "So that the achievement test used in the Michigan Assessment Program might accurately reflect the objectives of education in the state, the professional staff of each of the several departments of the Test Development Division of Educational Testing Service wrote a set of preliminary test specifications based on texts in use in Michigan Schools." (15,24)

The specifications were modified through interaction with an ad hoc committee appointed by the MDE. The final mathematics test specifications for the 1969-70 school year covered these broad areas:

Number and Operation

Geometry and Measurement

Relations, Function and Graphs

Logical Thinking

Mathematical Sentences

Applications

A more detailed description of those specifications and the published distribution of questions among them during the four-years appears in APPENDIX A.

Since these were taken from "texts in use," the tests might more reasonably be expected to assess the effectiveness of those texts than any other single ingredient. However instructional materials are used by teachers and students in a variety of school settings. It would be misleading to suppose that whatever strengths appear can be attributed to the texts or that weaknesses can be remedied by merely changing texts. Mathematics delivery systems are a blend of all four factors; materials, teachers, students, and programs.

It is important to note that this process for identifying objectives cannot be expected to identify a mathematics program which is distinct to "Michigan". Since the textbooks used in Michigan probably represent a cross-section of those used nationally, there is nothing uniquely "Michigan" in these specifications for mathematics. It is also worth noting that the assumption that teachers do teach what is in the textbook is a very broad generalization. In fact they teach some subset of that material and it is likely that a small ad hoc committee could accurately identify its elements without some systematic sampling.

This is illustrated by the fact that during these four years of testing, there was one whole number addition item in one of the 4th grade tests and the other tests contained none. It is very likely that whole number addition is a major objective in 3rd grade mathematics in Michigan and it should have been well covered in the 4th grade testing program.

The tests for 1969-70 contained 30 items and were administered in twenty-five minutes. In the next three years, the number of items grew to 40 and the time to 30 minutes.

The relationship of the test items used from year to year does not appear to have followed any particular pattern. A total of 83 items were used in the 4th grade testing and 79 items in the 7th grade. This table shows the number of items which were used on one, two, three, or four of the tests.

Frequency of Use of Test Items	Number of Items	
	Grade 4	Grade 7
Used on 1 test	44	37
Used on 2 tests	18 (9,4,5,)*	21 (6,6,9)*
Used on 3 tests	14 (4,10)*	13 (6,7)*
Used on all 4 tests	7	8
Total Items	83	79

*Parenthetical numbers show how those items were distributed among the first two years, second 2 years, and last two years or first three and last three years. For example, at grade 4, 9 items were common to the first two years ('70,'71) 4 to the second two years ('71,'72) and 5 to the last two years ('72,'73).

There was a commitment to continue a sufficient number of items from year to year to facilitate a statistical comparison of performances. Item changes permitted a gradual improvement in the selection of items as well as some shift in the objectives tested.

Technical Analyses of the Tests

ETS published technical reports describing the tests used in each of the first three years that the tests were administered. These reports discussed mean scores, standard deviation, reliability, validity, difficulty level, item difficulty, timing effects and discrimination power of each of the tests. In March 1973, the MDE also published an "Equating Report" (18) which discussed the technique for comparing scores achieved on the tests administered during the first three years. These reports were based on a "spaced sample" of approximately 1000 students each year with checks being made against the total group to establish the representativeness of the sample.

Statistics describing the mathematics tests in terms of the characteristics listed above are contained in APPENDIX B.

Based on generally acceptable standards (25,27) the tests for mathematics were reported to be reliable, time limitations were not a factor, they discriminated fairly well between strong and weak students and they were somewhat difficult for tests of this type. The equated scores indicated that the average performance did not shift significantly from year to year. In the case of mathematics, "the 1970, 1971, and 1972 test averages are identical." (18,p.18)

Later (13) the MDE also published equated scores for the 1972-73 testing program. That report appeared to show a slight improvement in 4th grade mathematics scores for the fourth year of testing.

The "content validity" of a test "is an indication of the extent to which it does the job it is intended to." (15,p.17)

> "Content validity is ensured by entrusting development of the tests to specialists in test construction and in the subject to be covered." (15,p.17)

> "The content validity of the achievement tests,...,can be judged by the procedure used in their development and by inspection of the classification of the test questions into various segments of the content domain." (16,p.21)

> "Committees of Michigan educators assisted in the development and review of items. In doing so and because of the representativeness of the content domain, the test can be judged as having a high content validity." (17,p.24)

These statements represent gradual shifts in the view of content validity in mathematics tests. In September 1973 the use of objective referenced tests related to objectives determined completely by Michigan mathematics educators represented the MDE recognition that assessment must be in terms of what mathematics educators in Michigan schools are trying to accomplish. The new tests consisted of sets of items developed to assess the achievement of specific objectives contained in <u>Minimal Performance Objectives for Mathematics Education in Michigan.</u> (14)

You are referred to the technical reports themselves for technical questions about the earlier tests which require greater detail.

Item Statistics

Finally, for the purpose of this analysis, ETS submitted the statistical data which it uses to examine performance on test items. The following statistics were among the nineteen pieces of data available on each item.

- --Number of students who selected each response to the item.
- --A mean criterion score for each response; this was a score from 6 - 20, average 13, describing the over-all test achievement level of the students who selected that response. The higher the criterion score the more capable the group of students.
- --A mean criterion score for the students still answering questions and the percentage of students still answering at that item.
- --Percentage of correct response to the item based on those still responding.
- --A mean criterion score for the item. This was also based on a 6 - 20 scale as described above. In this case the higher the criterion score the more difficult the item. In this analysis we will abbreviate this score as DL - difficulty level.
- --A corresponding criterion score for the item assuming it has been given to a base group, in this case students writing in 1969 - 70.
- --A correlation coefficient comparing performance on this item to overall performance on the entire test.

A more detailed description of these data taken from an ETS manual (19) are included as APPENDIX C.

EVALUATION OF THE TEST RESULTS

OVERALL TEST PERFORMANCE

Analysis of the overall test performance offers little information of value to Michigan mathematics educators. It has already been reported that the Equating Report identified no differences on the mathematics tests during the first three of the four years. Those statistics for the fourth year computed by MDE showed no great change in that respect.

The overall level of the performance was predetermined by ETS through a calculated distribution of easy and difficult items. It is therefore not a direct measure of the quality of performance by Michigan students, but rather a measure of the level of difficulty of test.

Since no other norming group is available, there is no immediate basis for comparing Michigan students to other students. The overall test performance data for Michigan districts or areas having common socio-economic characteristics were not assembled for this analysis. Future analysis should explore the relationship between the learning of mathematics and such forms of educational input.

ITEM PERFORMANCE

Inasmuch as the descriptive statistics of overall test performance are not available in a form which provides information of general interest to the Michigan mathematics educators, the major thrust of this analysis will relate to the performance of fourth and seventh grade students on test items. Particularly, that performance will be examined for information which may provide some answers for the following types of questions.

What kind of items were difficult for 4th graders (7th graders)?

What kind of items were easy for 4th graders (7th graders)?

How did 4th graders (7th graders) perform on collections of items such as:

- Place value
- Number concepts
- Fractions
- Word problems, and
- Problems using algebraic equations?

Do the tests provide evidence of specific learning problems in 4th and 7th grade mathematics?

The analysis of performance by many students on a large number of test items such as this poses many problems. The absence of objective data can be a serious handicap to decision-making but here we have so much data that organizing it into meaningful components is a challenge. While there is a large amount of data about the base groups there is only limited opportunity for comparisons between different grade levels, between similar or related items, or between sub-groups of students on the same item.

Besides the absence of comparative data there are other factors which will hamper our efforts:

1. Frequently two or three concepts are interwined in the same item. That poses no particular problem to the creator of the normative test, but it means that a student may have missed an item in spite of the fact that he had mastered all but one of the concepts. It also is very difficult to group such items into related categories. The specifications for these tests were clearly indentified earlier but different persons assigning items to those categories have not agreed on those assignments, and doubtless others would seem them still differently.
2. Multiple response items are made more or less difficult by the nature of the possible responses. Indeed, manipulating the choices is a device used by the test maker, to achieve a desired level of difficulty for the item. Once more, that handicaps a person attempting to detect program implications from performance on the items.
3. The normative test maker must use items appropriate for the students being tested so that the resulting statistics are within an acceptable range. Particular emphasis on specific areas is not important to him. Hence, some of the more critical areas for mathematics educators and their students are either not

represented or are masked by item construction techniques. On these tests the meaning of decimal fractions, operations with common fractions, and ability to use the whole number algorithms are not well covered.

The above remarks are not intended to be an exhaustive listing of the difficulties inherent in using normative tests to discern specific strengths, weaknesses or trends in the learning process. They are rather intended to put the readers on guard as they attempt their own analysis and to place this analysis in a proper context.

While we have attempted to exercise caution in interpreting results, it must be recognized that frequently the observations represent attempts to interpolate or extrapolate from the statistics. Since all statements relate to statewide results, it has also been impossible to anticipate how individual schools or districts may reflect different results.

Finally, the determination of what is worth calling to the readers' attention is subject to judgmental limitations. The raw data are available through the MDE and it is hoped that they will be scrutinized for other clues as to how we can improve instruction in mathematics.

It is also hoped that as other professionals arrive at alternative interpretations of the data they will utilize the journal <u>"Mathematics in Michigan"</u> or other suitable media to air their viewpoints.

With those precautionary remarks, let us examine the test items for information of interest. Remember the basic statistics discussed for the items will be difficulty level (DL) and % of correct response.

DL will be a number between 6 and 20 with an average DL for test items being about 12.7. DL's for more difficult items will be higher than that and for easier ones they will be lower. APPENDIX C contains a table relating DL's to percentile levels.

From time to time, assumptions will be made concerning the number of students who guessed the correct responses. Since there were 4 possible choices for each item, we will assume that those who "knew" the answer will equal correct responses decreased by 1/3 of the wrong responses. Special note will be made when that is done, otherwise the percentages will simply represent correct responses.

The tests were identified as form S, T, U, V for 1970, '71, '72, and '73 respectively; hence, we will use the letters and item numbers to identify an item in a particular test(s). UV6 would be the 6th item on both the 1972 and '73 tests.

Whole Number Computation and Place Value

While there was an average of nine whole number items dealing with computation and place value on each 4th grade test, there was a relatively small number concerned with any one concept. Probably the outstanding feature of student performance on these items was the failure of 4th graders as a group to excel on any of them.

The two easiest items were:

(V9) $10 \times \square = 100$

What number goes in the $\square$ above?

(A) 0
(B) 1
(C) 10
(D) 100

Difficulty Level 8.4, 87% correct.

(S5, TUV6) $12 \,\Delta\, 4 = 8$

What sign goes in the Δ above?

(A) +
(B) −
(C) ×
(D) ÷

DL 9.1, 82%

The more difficult were:

(T9) $9 - 7 = \square - 9$

What number goes in the box above?

(A) 2
(B) 7
(C) 11
(D) 16

DL 17.5, 13%

But careless 4th graders selected 2 or 7 and the item provided little insight into their understanding.

(TUV23) A: $2\overline{)252}$ B: $3\overline{)252}$ C: $4\overline{)252}$ D: $6\overline{)252}$

If A, B, C, and D are the answers for the division problems above, which is the greatest?

(A) A
(B) B
(C) C
(D) D

DL 15.7, 22%

(T40) Which of the following numbers would be 40 more if the digit 2 were changed to 6 ?

(A) 1,092
(B) 1,129
(C) 1,243
(D) 2,040

DL 15.7, 23%

It will not be surprising to middle elementary grade teachers that when guesses are eliminated the best that the 4th graders could do on division problems was an approximately 1/3 knowledgeable performance on (V29): $3\overline{)693}$

(A) 132
(B) 212
(C) 231
(D) 312

This is (S30)

```
23) 5175 | 200
   -4600 |
     575 |  20
   - [ ] |
     115 | △
         Total quotient
```

In the division example above, the correct way to fill in the spaces is

(A) $\boxed{460}$, △0
(B) $\boxed{460}$, △5
(C) $\boxed{560}$, △6
(D) $\boxed{560}$, △5

DL 15.0, 32%

Does the higher performance in S30 when compared to TUV23 above illustrate that the early emphasis on division is on the algorithm and not on its relationship to multiplication?

4th grade performance on place-value items must be disappointing to teachers and indicative of a need for either improvement in methodology or just greater emphasis. Place-value concepts have become very basic with the strong reliance on them in more recent development for whole number algorithms. Besides T40, which was listed earlier among the more difficult items, students performed on place-value items as follows:

(V18) $4{,}271 = 4{,}000 + 200 + \square + 1$

Which goes in the $\square$ above?

(A) 7
(B) 70
(C) 700
(D) 7,000

DL 13.3, 47%

(STU13) $634 = 600 + \square + 4$

What number goes in the box above?

(A) 3
(B) 30
(C) 34
(D) 63

DL 12.2, 50%

(V30) Which is NOT a way of writing 387 ?

(A) 300 + 80 + 7
(B) 200 + 180 + 7
(C) 200 + 170 + 17
(D) 100 + 280 + 17

DL 13.9, 43%

These demonstrate rather clearly that most of our students were not ready for more standard approaches to any of the whole number algorithms. However, V30 does seem to indicate that some students were ready. Under these circumstances perhaps the performances on the following items, which measure a major part of two years of mathematics instruction, can hardly be surprising.

(V12)

$$\begin{array}{r} 146 \\ 32 \\ +133 \\ \hline \end{array}$$

(A) 211
(B) 301
(C) 310
(D) 311

DL 10.3, 74%

(V13)

$$\begin{array}{r} 489,263 \\ -265,051 \\ \hline \end{array}$$

(A) 224,102
(B) 224,212
(C) 244,212
(D) 254,314

DL 11.1, 67%

(U15)

$$\begin{array}{r} 506 \\ -223 \\ \hline \end{array}$$

(A) 273
(B) 283
(C) 323
(D) 383

DL 12.9, 51%

Perhaps the most surprising of all performances by fourth graders in these items was on:

(V33) $4 + 4 + 4 + 4 + 4 = \square \times 4$

Which number goes in the $\square$ above?

(A) 4
(B) 5
(C) 20
(D) 24

DL 14.5, 37%

Imagine! Less than one in four 4th graders demonstrated an understanding of the association between multiplication and repeated addition. Assume that the 20% who selected "(C) 20" were careless and knew the correct response. Even then, after discarding guesses, only about 1/3 do understand the relation. What does it mean, then, when 55% correctly respond to (U24) $3 \times 604 =$

(A) 192
(B) 1,807
(C) 1,812
(D) 1,912

There were 40% who "knew" the correct response. Could they make use of knowledge so lacking in background?

Perhaps better than any other group of items these illustrate the enigma of the middle elementary grades in mathematics education. So many concepts included in the program are appropriate for some students and not for others. The need to teach different concepts to different students in the same classroom, i.e. differentiated instruction, is very strong. It makes little sense to try to teach the algorithms to youngsters who do not have a basic background. Yet the dominant instructional mode is very likely that of taking the whole class page by page. There is little wonder that many starry-eyed second graders who are intrigued by numbers turn into sixth graders who shy away from them.

Perhaps because the 7th-grade texts and programs have tended to move on to other things the group of whole number test items is less extensive for 7th graders. Questions which many 7th-grade mathematics teachers would have liked to have included, such as two and three digit multiplication and division, were not asked. Those which are included indicate that the basic weaknesses cited for 4th graders are still present in the 7th grade.

(ST5 UV12) 70,060 =

(A) $(7 \times 10{,}000) + (6 \times 10)$

(B) $(7 \times 10{,}000) + (6 \times 100)$

(C) $(7 \times 10{,}000) + (6 \times 1)$

(D) $(7 + 10{,}000) \times (6 + 10)$

DL 9.8, 79%

Because of the nature of the choices, that item established that 1/3 of the students could not correctly identify the place value of the "6".

(ST17, U28) If $\square + \Delta = 9$, where $\square$ and Δ are whole numbers, which of the following is true?

(A) $9 + \square = \Delta$

(B) $9 - \Delta = \square$

(C) $9 + \Delta = \square$

(D) $\square - 9 = \Delta$

DL 12.5, 55%

(T12) $25 \times 8 =$

(A) $(25 \times 6) \times 2$

(B) $(25 \times 5) \times 3$

(C) $(25 \times 4) \times 2$

(D) $(25 \times 4) \times 4$

DL 11.4, 66%

(S13) $21 \times 32 =$

(A) $(20 \times 32) + (1 \times 32)$

(B) $(2 \times 32) + (1 \times 32)$

(C) $(20 \times 30) + (1 \times 2)$

(D) $(21 \times 30) + (2 \times 30)$

DL 13.3, 47%

(UV6)

$$\begin{array}{r} 6 \\ 62\overline{)376} \\ \boxed{} \\ \hline 4 \text{ remainder} \end{array}$$

In the division problem above, the number that goes in $\boxed{}$ is

(A) 4×62
(B) 6×62
(C) $(4 \times 62) + 6$
(D) $(6 \times 62) + 4$

DL 11.1, 68%

(STU8) If $36 - (n \times 3) = 0$, then $n =$

(A) 4
(B) 8
(C) 12
(D) 33

DL 10.9, 70%

Deducting for guesses, the best success on these items was just over 2/3; the average is about 1/2. They are all related to a meaningful use of numbers. These performances clearly indicate that a large group of Michigan 7th graders are not very facile with whole number concepts. Again, implication for differentiated mathematics instruction at 7th grade and some modification of elementary programs are inescapable.

Questions related to multiples of whole numbers constituted the balance of these whole number items for 7th graders. Around 2/3 of the 7th graders demonstrated familiarity with the concept of "multiple" as in:

(V8) If x is a multiple of 21, it must also be a multiple of

(A) 11
(B) 7
(C) 5
(D) 2

DL 10.7, 73%

However, two other questions concerning common multiples were found to be so difficult and complex that they were not informative; i.e.,

(ST29) The least common multiple of 6, 8, 18, and 30 is

(A) 1
(B) 2
(C) 180
(D) 360

DL 16.5, 20%

Properties of Whole Numbers and Their Operations

Items treating the special properties of zero and one, and whole number operations are examined together in this section.

While nearly all (93%) of the 4th graders used zero correctly in (UV1) $13 + 3 + 0 =$

(A) 0
(B) 10
(C) 16
(D) 17

only 2/3 could generalize the role of zero in addition to (TUV3) $7 + \square = 7$

What number goes in the box above?

(A) 0
(B) $\frac{1}{7}$
(C) 1

2/3 of the 4th graders could also generalize multiplying by "1" in (S2, T3):

$3 \times \square = 3$

What number goes in the box above?

(A) 0
(B) 1
(C) 2
(D) 3

Since it seems very likely that nearly all 4th graders know that "3 x 1 = 3" and "7 + 0 = 7" those last two items seem to indicate that 1/3 of the 4th graders are confused by "box".

A related question (U9)

$84 \times \square = 84$

What number goes in the $\square$?

(A) 84
(B) 83
(C) 1
(D) 0

was responded to correctly by 50% of the 4th graders, suggesting that there was still another group who understood the use of the "box", and that "3 x 1 = 3", but had not formed the generalization "N x 1 = N".

Consider (V26) for 7th graders: Which is true for all numbers N ?

(A) $N - 2 = 2 - N$
(B) $N + 0 = 0$
(C) $N \times N = 2 \times N$
(D) $N \times 1 = N$

Since 7th graders selected "N x 1 = N at a 41% rate, it is likely that not more than 1/3 of the 7th graders had formed the generalization for multiplying by 1.

About 1/2 of the 7th graders able to respond correctly to (S15) If $x \neq 0$ and $y = 1$; which of the following expressions equals 1 ?

(A) $x + y$ (C) $\frac{x}{y}$

(B) xy (D) $\frac{y}{x}$

90% of them responded correctly to (V4) If $7 \times n = 0$, then $n =$

(A) 7

(B) 1

(C) $\frac{1}{7}$

(D) 0

and to (S1, TU2) If $4 + 3 + N = 7$ and N is a number, then $N =$

(A) 0
(B) 3
(C) 4
(D) 7

These three items point to the pesky nature of zero and one for both 4th and 7th graders. The generalizations may be mastered quickly by some, others are at least periodically unsure and some just do not know.

Two items covered the commutative law for addition,

(ST16, UV22) $23 + 4 = 4 + \square$

What number goes in the $\square$ above?

(A) 27
(B) 25
(C) 24
(D) 23

(UV5) If $24 + 75 = 75 + N$, then $N =$

(A) 0
(B) 12
(C) 24
(D) 48

They were answered correctly by about 63% of the 4th graders and 80% of the 7th graders respectively.

Two more items related to the distributive law:

(STUV27) $\Delta \times 9 = (5 \times 9) + (3 \times 9)$

What number goes in the Δ above?

(A) 2
(B) 8
(C) 9
(D) 15

(S13) $21 \times 32 =$

(A) $(20 \times 32) + (1 \times 32)$
(B) $(2 \times 32) + (1 \times 32)$
(C) $(20 \times 30) + (1 \times 2)$
(D) $(21 \times 30) + (2 \times 30)$

They were managed by 38% of the 4th graders and 47% of the 7th graders respectively.

(T29) Which is twice $(8 + 8)$?

(A) $2 + 8 + 8$
(B) $16 + 8$
(C) $4 + 12$
(D) $16 + 16$

received a 45% correct response from 4th graders.

Thus, in spite of textbooks which stress these basic number relations, a large proportion of our students have not mastered them by the 7th grade. They appear to provide a very shaky basis for the teaching of computational algorithms.

A question involving an application of the concept "average" was one of the most difficult on the entire test--

(STUV30) If the average of 5 integers is 12 and the sum of 4 of the integers is 52, then the 5th integer is

(A) 8
(B) 12
(C) 40
(D) 60

DL 16.0, 25%

There is only the slightest statistical assurance that any 7th grade students responded to that question knowledgeably.

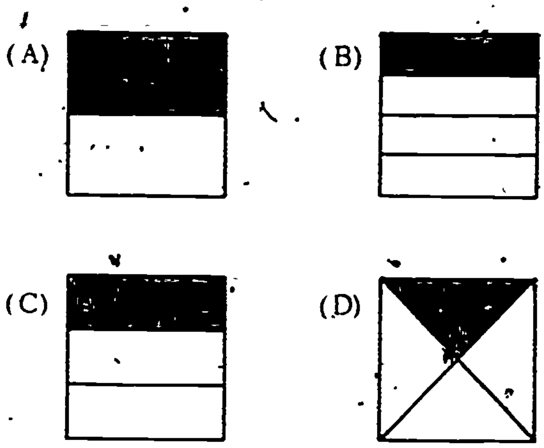
Fractions

By January of their 4th grade year, many students have had only limited experience with fractions. There is some rather strong sentiment that the major emphasis at that time should be on improving proficiency with the whole number algorithms. In many classes, fractions tend to be given a cursory treatment late in the year.

One might hold that 4th graders should study the mathematics for which they are ready. Is two-digit multiplication easier or more difficult for 4th graders than the introductory concepts related to fractions? We have already seen that an uncertain knowledge of place value and other number properties makes a meaningful introduction to the multiplication algorithm doubtful for at least half of the 4th graders. Does performance on these introductory concepts related to fractions suggest that greater attention to fractions at the 4th grade level would be more timely? Have we devoted too much time in the 3rd and 4th grade to larger whole numbers and too little to the simple fraction concepts?

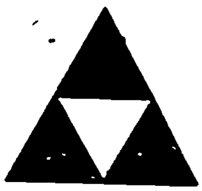
Three test items asked students to associate fractions with geometric regions.

(TUV34) Which square has the greatest part shaded?



DL 11.3, 70%

(S21)



What fraction of the figure above is shaded?

(A) $\frac{3}{4}$

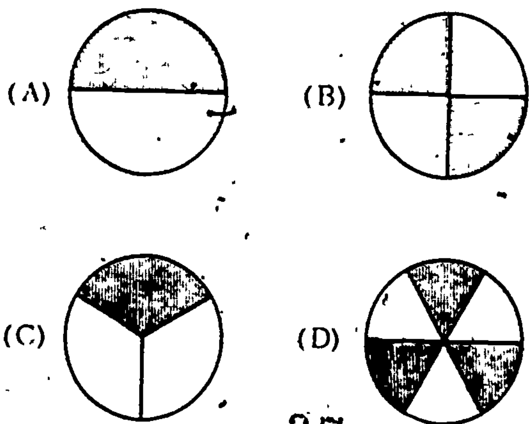
(B) $\frac{3}{5}$

(C) $\frac{1}{4}$

(D) $\frac{1}{5}$

DL 13.1, 50%

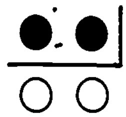
(TU28) Which figure is NOT $\frac{1}{2}$ shaded?

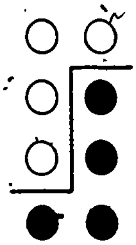


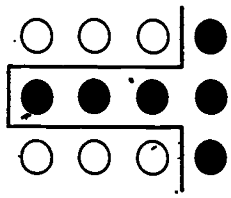
DL 13.0, 48%

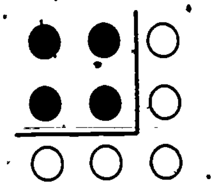
And five others with some other form of meaning of fractions:

(V32) Which set does NOT have $\frac{1}{2}$ the dots shaded?

(A) 

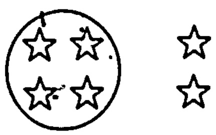
(B) 

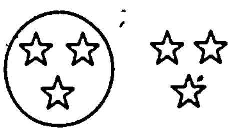
(C) 

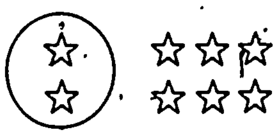
(D) 

DL 13.3, 49%

(U38) In which set are $\frac{1}{3}$ of the stars circled?

(A) 

(B) 

(C) 

(D) 

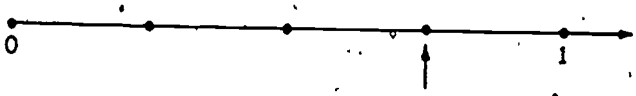
DL 18.0, 12%

(S24) 6 is $\frac{1}{3}$ of

(A) 18 times 6
(B) 9 times 6
(C) 6 times 6
(D) 3 times 6

DL 12.5, 56%

(S29)



The arrow above points to a number named by all the fractions in which of the following sets?

(A) $\left\{\frac{1}{4}, \frac{2}{8}, \frac{3}{12}, \frac{4}{16}\right\}$

(B) $\left\{\frac{1}{3}, \frac{2}{6}, \frac{3}{9}, \frac{4}{12}\right\}$

(C) $\left\{\frac{1}{2}, \frac{2}{4}, \frac{3}{6}, \frac{4}{8}\right\}$

(D) $\left\{\frac{3}{4}, \frac{6}{8}, \frac{9}{12}, \frac{12}{16}\right\}$

DL 16.5, 28%

(S9) If 4 sticks are each cut into fourths, how many pieces are there?

(A) One
(B) Four
(C) Eight
(D) Sixteen

DL 12.4, 56%

Assuming that 4th graders have not had an extensive exposure to fractions, the following observations appear to emerge from that collection of items:

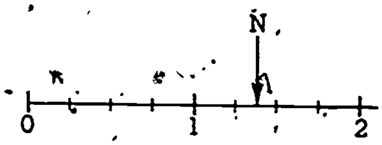
Fractions related to regions are easier for 4th graders to understand than when they are related to sets. (Compare S21 to U38).

Fourth graders have a strong intuitive grasp of 1/2 as indicated by the strong performance on V32.

The set interpretation of fractions appears to represent a stumbling block (compare S24 and U38). This is perhaps because the basic unit to which the fraction applies is more obscure in a set than it is in a region.

It is unfortunate that S29 was complicated by the concept of equivalent fractions. One would like to believe that having identified the point as "3/4" students would have seen only one correct choice, but a low correlation to performance on the rest of the test (.16) and other items statistics suggest confusion. The only other item which asks 4th graders to associate fractions with the number line (T38)

(T38)



On the number line above, N is halfway between $1\frac{1}{4}$ and $1\frac{1}{2}$. What do you add to $1\frac{1}{4}$ to get N ?

(A) 1

(B) $\frac{1}{2}$

(C) $\frac{1}{4}$

(D) $\frac{1}{8}$

was impossibly complex and offered no additional insight.

It is interesting to note that one of the few items on the test on which there is solid evidence of change over the four years of testing is (TUV34) (p.29) on which performance in successive years improved from 66% to 76%.

(T36, UV35) Students are asked to give meaning to mixed numbers,

$7\frac{3}{4} =$

(A) $7 + \frac{3}{4}$

(B) $7 - \frac{3}{4}$

(C) $7 \times \frac{3}{4}$

(D) $7 \div \frac{3}{4}$ DL 13.6, 45%

There are a number of complexities in that item and it is surprising to find that (DL 13.6, 45%) one-third of the 4th graders are that knowledgeable about mixed numbers.

The weakness of using a norm-referenced test for facilitating instruction stands out clearly in the items on fractions used with 7th graders. There is so much specific information which would have been useful but the item selection did not provide a basis for getting that information. An item (STUV3) Which statement is true?

(A) $\frac{1}{5} > \frac{1}{4}$

(B) $\frac{1}{2} > \frac{1}{3}$

(C) $\frac{1}{9} > \frac{1}{8}$

(D) $\frac{1}{6} > \frac{1}{5}$ DL 10.5, 72%

deals with size comparison for unit fractions but there is no follow-up for nonunit fractions. A single question (V7)

(V7) Each of the following is equal to $\frac{1}{2}$ EXCEPT

(A) $\frac{9}{18}$

(B) $\frac{8}{16}$

(C) $\frac{6}{3}$

(D) $\frac{5}{10}$

DL 9.9, 80%

deals with equivalent fractions but the fraction involved is 1/2, one with which students show such strong intuitive indentification that the 80% performance tells nothing about an understanding of the generalized concept.

In spite of interfering factors in two of them, these three 7th grade items do provide some meaningful information about the effectiveness of instruction:

(V36) $\frac{2}{6} + \frac{1}{3} =$

(A) $\frac{2}{18}$

(B) $\frac{3}{9}$

(C) $\frac{1}{2}$

(D) $\frac{2}{3}$

DL 14.8, 32%

(U26) $\frac{1}{2} + \frac{1}{3} + \frac{1}{4} =$

(A) $\frac{9}{12}$

(B) $\frac{11}{12}$

(C) $1\frac{1}{12}$

(D) $1\frac{1}{6}$

DL 13.4, 46%

(V40) $\frac{7}{8} - \frac{1}{2} =$

(A) $\frac{(2 \times 7) - (8 \times 1)}{8}$

(B) $\frac{7-4}{8}$

(C) $\frac{7-1}{8}$

(D) $\frac{7-1}{8-2}$

DL 14.8, 33%

Each of these items was used on just one test. That provides a very unsatisfactory look at 7th graders' ability to compute with fractions! Performance on these items speaks for itself. Michigan 7th graders are very weak in the addition of fractions with unlike denominators and are probably also weak in subtracting them.

Two items related simple fractions directly to decimals. One (TU20) $2 + 2\frac{1}{2} =$

(A) 0.45
(B) 2.25
(C) 4.05
(D) 4.5

required the students to represent "2 + 2 1/2" as a decimal and another (S7) $1 + \frac{1}{2} =$

(A) .15
(B) .5
(C) 1.05
(D) 1.5

required the same for "1 + 1/2". Average DL for these items was about 12.0 with 60% correct responses. In other words, about half of the 7th graders tested "knew" how to represent 1/2 as a decimal fraction. About the same number were able to manage (UV19) 200.1 - 199.0 =

(A) 1.0
(B) 1.1
(C) 10.1
(D) 11.1

Large numbers of Michigan 7th graders do not have even a minimal ability to work with decimals; but the nature of the items and the percentage of correct responses is such that there are likely to be large differences among districts in the performance of students in decimal fractions.

The test makers put unusual stress on the concept that a number multiplied by a fraction could be less than the number itself. They administered 3 items dealing with that concept a total of 7 times and discovered that perhaps 1/3 to 1/2 of the 7th graders were aware of this or could deal with multiplication in a way which permitted them to make accurate deductions about the relations.

(STUV22) Karen says, "When I multiply any number by 2, the answer I get is always 2 or more." The answer to which of these problems shows Karen is wrong?

(A) $2 \times 1\frac{1}{2}$

(B) $2 \times \frac{11}{10}$

(C) 2×1

(D) 2×0

DL 11.8, 62%.

(UV38) If 300 times a number N is greater than 300, then N could be

(A) $\frac{1}{8}$

(B) 0.888

(C) 1

(D) 8

DL 13.6, 46%

(T38) For which of the following does n represent a number greater than 300 ?

(A) $300 \times \frac{3}{4} = n$

(B) $300 \div 2.3 = n$

(C) $300 \times 1.2 = n$

(D) $300 \div \frac{3}{2} = n$

DL 15.0, 33%

Collectively these items seem to indicate that most 7th graders have not achieved the ability to make these kind of comparisons. They suggest that 7th graders' "feel" for numbers of this type is not very strong.

The most difficult item in this section for 7th graders was:

(T40) If $82 \times \frac{1}{N} = 0.082$, then N =

(A) 0.001
(B) 0.01
(C) 100
(D) 1,000

DL 17.1, 17%

Perhaps its placement as the last item on the test had some influence on the outcome, but on face value, and particularly when considered with TU20, S7, (p.35) and UV19 (p.36) it appears to point to a severe breakdown in the development of decimal fractions concepts. With the present emphasis on metric measurement, it most certainly demands more attention.

The comparison of:

(U16) 3 is $\frac{1}{2}$ of

(A) 2×3
(B) 3×3
(C) 4×3
(D) 6×3

DL 10.4, 74%

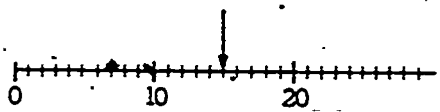
to S24 (p.31) for 4th graders is perhaps worth noting. The 7th graders had the advantage of an intuitively more familiar fraction, but their performance advantage (18%) is small enough to conclude that this concept was not stressed in the intervening years.

Measurement

A large number of items applied to the general area of measurement. 8 items collectively appeared 16 times on tests for 4th graders, and 17 items appeared 30 times on tests for 7th graders. That is an average of 4 items per test to 4th graders and 7 items per test for 7th graders. But the number of important concepts in this area is also great, so only linear measurement received fairly detailed attention; other measurement areas were barely touched.

Linear Measure

(T33)

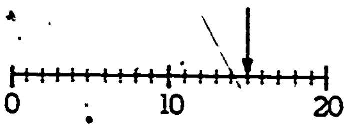


On the number line above, the arrow points to

(A) 12
(B) 15
(C) 17
(D) 25

77% of 4th graders correctly selected (B) for that question. When the scale was terminated at 20 (UV21),

(UV21)



On the scale above, the arrow points to

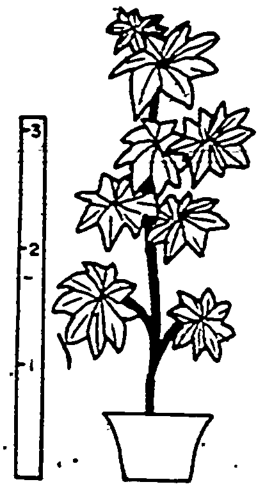
(A) 12
(B) 15
(C) 17
(D) 25

the percentage for the same question increased significantly to 83%. Not a great change, but why should the open-ended scale keep 6 students in every 100 from succeeding?

S29 (p.31) was discussed earlier with fractions. It indicated that the use of fractions on the number line with 4th graders should be approached very cautiously. It also suggested that expecting the average 4th grader to measure distance to quarter units is not very realistic for many Michigan 4th graders.

Another perceptual problem in linear measurement was encountered by 4th graders when asked to respond to:

(ST15)



This plant is next to a yardstick. About how tall is it?

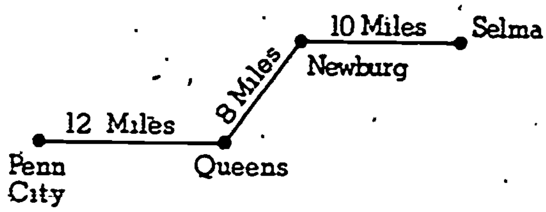
(A) Almost 4 yards tall
(B) Almost 4 feet tall
(C) Almost 40 inches tall
(D) Almost 4 inches tall

DL 14.5, 36%

About 36% responded correctly. Discounting guesses, about 1/6 of them were able to perceive the fact that the 2 1/2 inch scale was intended by the adults to be a yardstick. Very few appeared to be distracted by distinguishing 40 inches from 4 feet. The message seems clear: 4th graders will have a tough enough time measuring in the real world; don't ask them to visualize a yardstick shrunk to 2 1/2 inches.

Interestingly, on this map, a few more students were able to identify the correct response for the distance from Penn City to Newburg.

(TU32)



On the map above, what is the distance from Penn City to Newburg?

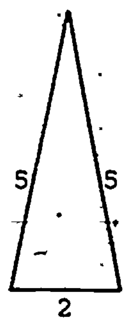
(A) 12 miles
(B) 18 miles
(C) 20 miles
(D) 30 miles

DL 13.7, 45%

And the correlation to performance on the rest of the test was significantly better. Perhaps this was because they had never seen a mile?

Nearly 3/4 of the 4th graders could correctly identify the perimeter of this triangle, but fewer than half could tell how much wire it would take to make the square.

(ST12)

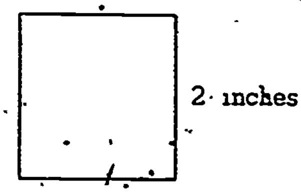


The perimeter of the figure above is

(A) 10
(B) 12
(C) 25
(D) 50

DL 9.8, 79%

(UV11)



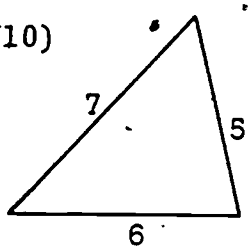
How many inches of wire would be needed to make the square above?

(A) 4 inches
(B) 8 inches
(C) 12 inches
(D) 16 inches

DL 12.3, 57%

This could indicate that while they recognized the square, for many of them it did not have equal sides. Or maybe they just couldn't imagine that the side, which appeared to be 1 inch, was actually 2 inches in length.

Seventh graders displayed a similar problem. Nearly 7/8 of them correctly identified the perimeter of a given triangle (V10)



The perimeter of the figure above is

(A) 11
(B) 13
(C) 18
(D) 30

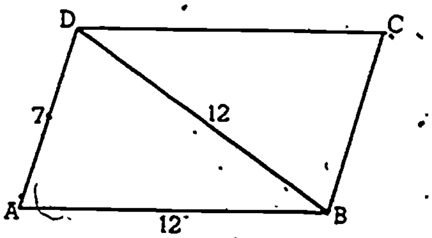
but less than 1/4 responded correctly to:

(TU18) If the perimeter of a square is 36, then the length of one side is

(A) 6
(B) 9
(C) 12
(D) 18

Even though the alternatives were not very attractive, only 1/4 of the 7th graders correctly gave the perimeter of the parallelogram:

(S21)

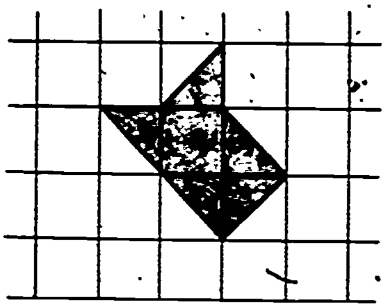


In the figure above, what is the perimeter of parallelogram ABCD ?

(A) 19
(B) 31
(C) 38
(D) 84

This seems to indicate that the characteristics of common geometric figures do not come through very strongly in the K - 7 mathematics program.

While 70% of the 7th graders responded correctly to (STU10)



If each small square above has an area of 1 square inch, then the area of the shaded region in square inches is

(A) $3\frac{1}{2}$

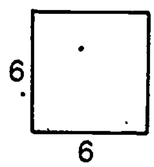
(B) 4

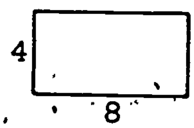
(C) $5\frac{1}{2}$

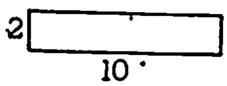
(D) 6

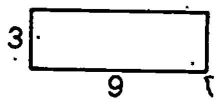
a question calling for them to determine the area of a polygon displayed on a grid, only 60% could tell: (V20)

(V20) Of the following rectangles, which has the greatest area?

(A) 

(B) 

(C) 

(D) 

And finally on (T37) fewer than 1/4 managed.

(T37) What is the greatest number of squares of cake 2 inches on a side that Jane can cut from a cake 15 inches by 8 inches if the thickness of each square is the same as the thickness of the cake?

(A) 21
(B) 28
(C) 32
(D) 34

In other words, only the best Michigan 7th graders are functional with the area of simple rectangles. Fortunately, there were no area questions covering triangle, parallelograms, or trapezoids.

The balance of the measurement questions for 7th graders dealt with units of measurement. There were six unit conversion problems:

(ST9) minutes to hours--76%

How many hours is 150 minutes?

(A) 2

(B) $2\frac{1}{2}$

(C) $3\frac{1}{2}$

(D) 5

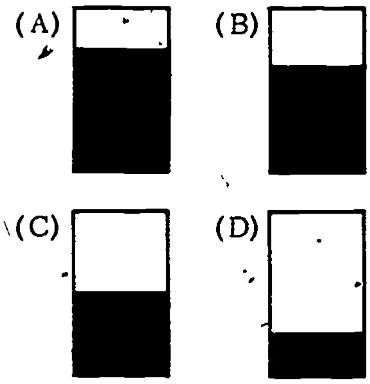
(TU15) dimes to nickels--64%

24 dimes are equal in value to

(A) 480 pennies
(B) 48 nickels
(C) 12 quarters
(D) 5 half dollars

(TUV34) gallons to quarts--54%

Which of the following 1-gallon cans contains about 1 quart?



(ST27) feet to inches application-35%

How many quarter-inches are in two feet?

(A) 96
(B) 48
(C) 24
(D) 4

(V13) decimeter to meter (given 1 meter = 10 decimeters)--63%

One meter is equal to 10 decimeters. How many meters are equal to 80 decimeters?

(A) 800
(B) 8
(C) 0.80
(D) 0.08

(T21) poundto ounces application--36%

How many 8-ounce packages of hard candy can be made from $2\frac{1}{2}$ pounds of hard candy?

(A) 3
(B) 4
(C) 5
(D) 6

Together they indicate that about, one-half of Michigan 7th graders are not comfortable with the more common units of measurement. Even the dimes to nickels conversion of 24 dimes to 48 nickels confused more than 1/3 of them.

This item was answered correctly by 53% of the 7th graders.

(UV29)

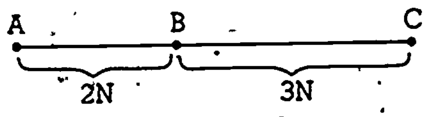
```
  3 hours 20 minutes
 -1 hour  35 minutes
```

(A) 2 hours 15 minutes
(B) 1 hour 45 minutes
(C) 1 hour 25 minutes
(D) 1 hour 15 minutes

When compared to the conceptually easier problems in the last paragraph, one can only conclude that not much time is spent in Michigan schools in teaching students to use units of measurement and that they do not pay much attention to them outside of school either. One would suspect that without schooling a 7th grader would be able to tell which of four partially filled 1 gallon containers contained 1 quart, but fewer than half were able to on these tests, (TUV34) (p.45).

One of the most difficult measurement questions for 7th graders was a proportion problem.

(U40)



In the figure above, if the length of segment AB is 12 inches, then the length of segment AC is

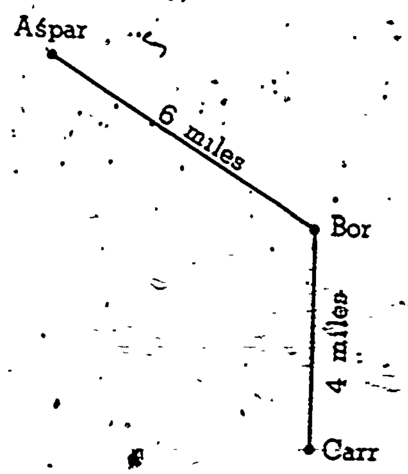
(A) 18 inches
(B) 24 inches
(C) 30 inches
(D) 36 inches

So few students knew how to deal with that problem that it would not be possible to statistically establish that anyone "knew". Yet students did respond at the 84% level to: (STU4) If 5 inches on a map represents 120 miles, how many miles does 1 inch represent?

(A) 600
(B) 125
(C) 115
(D) 24

There was also a rather strong performance (59% on V21) in another problem requiring direct use of a scale.

(V21)



Scale: 1 inch = 3 miles

On the map above, the length of the segment connecting Aspar and Bor is how many inches?

(A) 2
(B) 3
(C) 4
(D) 5

This group of problems concerning measurement tend to make it quite clear that Michigan mathematics programs either do not stress or are not particularly effective in their stress on measurement and other related topics which apply numbers to geometric figures (metric geometry). This is one of the most obvious areas for giving practical meaning to numbers. Since that has seemed to be a general weakness in modern mathematics programs, and with the more recent emphasis on career education, it appears that this area of instruction does need considerably more attention.

Geometric Properties of Figures

Geometric properties of various figures were covered by 3 items which collectively appeared 5 times on 4th grade tests and 6 items which appeared 9 times on 7th grade tests -- an annual average of slightly over one item per test for 4th graders and 2 items per test for the 7th graders. The results are correspondingly sparse. At the 4th grade level:

--90% of the students selected a triangle from a number of plane figures (S1, T2, U3)

--80% of the students identified points which were inside both of an overlapping square and circle (TUV5)

--About 50% correctly picked a shortest distance between two points on a grid (V26)

At the 7th grade level:

--just over 50% correctly selected a polygon (square) which had parallel sides (UV23)

--about 33% correctly identified a diameter as a line twice as long as a radius (ST16)

56

--about 50% were able to manage a rather wordy problem involving distances and direction on a gridded map (ST23, V27)

Perhaps the outcomes most worthy of noting here are the failure of the mathematics program (1-7) to deal effectively with parallel and the radius/diameter relationship. These are so obviously within the experience of these students that one wonders whether the more abstract geometric concepts investigated in many elementary mathematics books have hidden relationships which are easily within the students' grasp and perhaps far more useful to them. These books may even imply that a geometric relationship is not considered useful or significant unless it is couched in proper mathematical terms.

The facility of both 4th and 7th graders to work on a gridded plane is very interesting and could be useful. It is obvious that in many respects the regularities within our cities and buildings does result in students living in a gridded world. The geoboard attempts to capitalize on this. Perhaps other concepts like perpendicularity and parallelism should be taught through the use of the geoboard.

Formulas and Graphs

The use of algebraic expressions in a formula to generalize methods for arriving at certain values, such as $C = 2\pi r$ or $V = lwh$, or to describe relations such as

$\frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd}$ was not directly explored by these tests. However, three items did examine the ability to substitute numbers into algebraic expressions to determine values. Results on these items showed that, despite general student familiarity with the operations involved, well over half of the students had difficulty with such substitutions.

In the 4th grade this was true for (STUV19) $2 \times k$, and in the 7th grade for (TUV11) $\frac{a \times b}{c}$ and (S2) $\frac{2k}{r}$. This seems to imply that textbooks which have relied heavily on such expressions to describe relationships tend to reach only the better students.

In examining students ability to work with graphs, tests of this nature tend to be restricted to items which require an interpretation of existing graphs. The construction of a graph involves different skills and it seems probable that many students are likely to be able to interpret graphs which they could not construct.

Since these tests were not different in that respect the results are probably somewhat optimistic about students ability to work with graphs. Bar graphs, pictographs and line graphs were examined at the 4th grade (3 items) and circle graphs were added to give a total of 7 items at grade 7. Except for a 72% score on a very simple pictograph (V20) at grade 4, the results for the 10 items, which collectively appeared 15 times, were surprisingly consistent.

The minimum correct response percentage was 42% and the maximum 62% indicating that between 1/4 and 1/2 of the students were able to interpret the graphs. We must therefore believe that not over 1/3 of our 7th graders are able to construct such graphs. Ability to interpret bar graphs was only slightly stronger than it was for circle graphs.

Since the ability to interpret graphs appears on nearly every list of functional competencies and since it increases in importance for individuals as their use of mathematics grows, this should be an area for concern. Unless great stress is placed on this area in later instruction it seems likely that most consumers will not be even minimally competent in this area. The fact that it needs attention is not likely to surprise mathematics educators. This is perhaps an area in which mathematics instruction might profit by greater cooperation with other curricular areas, such as science and social studies.

Equations and Inequalities

At the 4th grade level, the findings on problems involving equations and inequalities are frequently inconclusive. As with word problems, it is sometimes difficult to determine whether the mathematical concepts or the equation format represents the difficulty.

On the test placeholders (□ , Δ , etc.) are used in equations to represent numbers or operation signs. Some evidence that these are not particularly confusing can be seen in these two problems.

(ST5 U8 V6) $12 \ \Delta \ 4 = 8$

What sign goes in the Δ above?

(A) +
(B) −
(C) ×
(D) ÷

(V17) Kim dropped a dozen eggs. If seven eggs were not broken, how many were broken?

Which number sentence can be used to solve the problem above?

(A) $12 - 7 = \square$

(B) $\square - 7 = 12$

(C) $7 + 12 = \square$

(D) $\square - 12 = 7$

81% and 70% respectively responded correctly to these items. Since the second problem also has some format and language difficulties not present in the first, it would appear that the equation form of the responses to the second did not represent difficulty.

It is interesting to note, however, that in item (STUV19) which uses a letter to represent a number, the situation was quite different.

(STUV19) If $k = 4$, then $2 \times k =$

(A) $\frac{1}{2}$

(B) 2

(C) 8

(D) 16

Actual correct responses to that item was about 50%. Indicating that roughly 1/3 of the students dealt with it knowledgeably: Since other evidence indicated that at least 90% of the 4th graders knew that 2 x 4 = 8, it seems quite clear that only the best students were comfortable with the use of "k" in that way. This seems to indicate that the expression of concepts in terms of letters should be used very cautiously with 4th graders.

In the area of inequalities there is an interesting comparison between two approaches to the ">" concept.

(UV2) Which is the greatest?

(A) 98
(B) 89
(C) 888
(D) 908

90% correctly selected (D). The 90% indicated that 1/8 of the 4th graders still need remedial help in this area. Clearly <u>they</u> are not ready for the multiplication and division algorithms. It is also interesting when compared to (ST8) to which less than 1/2 responded knowledgeably.

(ST8) Which statement is true?

(A) 145 > 155
(B) 125 > 125
(C) 135 > 145
(D) 155 > 135

About half of the students fell victim to the words "statement", "true", and ">". A 70% correct response to

(V15) Which statement is true?

(A) $9 - 6 \neq 3$
(B) $9 - 6 < 3$
(C) $9 - 6 > 3$
(D) $9 - 6 = 3$

61

indicates that a sizeable group of our 4th graders are confused by the symbol ">".

By the 7th grade a 72% gave a correct response to

(STUV3) Which statement is true?

(A) $\frac{1}{5} > \frac{1}{4}$

(B) $\frac{1}{2} > \frac{1}{3}$

(C) $\frac{1}{9} > \frac{1}{8}$

(D) $\frac{1}{6} > \frac{1}{5}$

That seems to indicate that by grade 7 ">" is primarily a remedial problem.

A 90% correct response by 7th graders to S1, TUV2,

If $4 + 3 + N = 7$ then $N = ?$

indicates that 7th grade students are not completely uncomfortable with the use of letters.

However a 48% correct response to (TUV11)

If $a = 3$, $b = 8$, and $c = 4$, then $\frac{a \times b}{c} = ?$

and a 44% correct response to (S2)

If $k = 5$ and $r = 2$ then $\frac{2k}{r} = ?$

indicates that letters must still be used with great caution. Using letters as a means of expressing a relationship or, making a definition, as is common at the 7th grade level, could be confusing to nearly 2 out of every 3 students. That is enough of a problem to ban the use of letters for such purposes at the 6th and 7th grade level unless it follows a long developmental process through which the relationship has been internalized. In 1935 Harold Fawcett commented on that problem in Mathematics Teacher:

> "Behind every symbol is an idea. It is the idea which is important and it is familiarity with that idea which puts life into the symbol. It is, therefore, of greatest importance that the idea be identified before it is symbolized."(20)

That appears to be a very appropriate caution for "modern mathematics."

Word Problems

The summary of mathematical specifications indicated that at least a third of these problems could also be examined within other specific categories. For the purpose of this analysis, 22 fourth grade items and 13 seventh grade items were placed in this category. They represented a total of 40 and 27 items respectively for the four years of testing. This was 26% and 18% of all items administrated at those levels. Medium difficulty level for the items was 13.0 and 14.0 compared to 12.6 for all math items. Hence, the word problems tended to be slightly more diffiuclt than other items.

The word problems which were easiest for 4th graders were:

(V8) Kathy has only 3 black kittens and 6 white kittens. How many kittens does she have?

(A) 18
(B) 9
(C) 6
(D) 2

DL 7.8, 90%

63

(S3, TUV4) Jill had 203 stamps and lost 4 of them. How many stamps does she now have?

(A) 199
(B) 207
(C) 209
(D) 243

DL 9.5, 81%

The most difficult word problems for 4th graders were:

(V40) Jack's spelling test has 100 words. He spelled $\frac{3}{4}$ of the words correctly. How many words did Jack misspell?

(A) 75
(B) 40
(C) 25
(D) 4

DL 15.9, 33%

(U40) If Bob is fifth in a line of children and Bill is twelfth in the line, how many children are between Bob and Bill?

(A) Six
(B) Seven
(C) Eight
(D) Nine

DL 16.4, 23%

(T39, UV37) John bought 7 pencils for 5 cents each and a box of chalk for 20 cents. Which could be the correct change he got from a dollar?

(A) 2 dimes and 1 quarter
(B) 2 quarters
(C) 1 nickel and 2 quarters
(D) 1 dime and 2 quarters

DL 15.5, 30%

The easiest items dealt with simple addition and subtraction which are well within the 4th graders repertoire, while the harder, except for U40, involved multiplication of whole numbers and the use of fractions, topics which are new to 4th graders. These results would certainly support the hypothesis that proficiency in the mathematical concepts involved is a necessary condition to solving a word problem. They may even suggest that proficiency in the mathematical concepts is the dominant factor. Pursuing that a step further, whole number addition or subtraction were required for six word problems in which average performance was 73% correct and the difficulty level 10.1. In 9 items involving multiplication or division the average performance was 46% with a difficulty level of 13.8. That performance was not greatly different from the 65% - 42% comparison obtained from comparing two pure addition, subtraction items to two multiplication, division items.

<u>For the 7th grade, the easiest of 13 items were:</u>

(T1, UV9). On a 2-hour trip, Jack drove at an average rate of 45 miles per hour. What was the total number of miles that he drove?

(A) $22\frac{1}{2}$

(B) 60

(C) 90

(D) 135

<u>DL 9.3, 82%</u>

(TUV35) On Monday Tim had $20. If he then earned $8 selling newspapers, spent $2 for a book, spent $1 for a movie, and earned $6 running errands, how much did he have?

(A) $11
(B) $26
(C) $31
(D) $37

<u>DL 11.0, 78%</u>

65

The most difficult items were:

(V32) On field day, 7 children ran the 50-yard dash and 8 children ran the 100-yard dash. If a total of 12 children entered the two races, how many ran in both races?

(A) 3
(B) 4
(C) 5
(D) 15

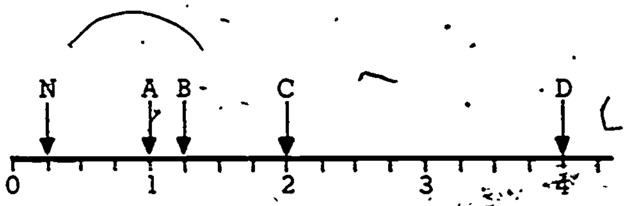
DL 14.5, 35%

(T37) What is the greatest number of squares of cake 2 inches on a side that Jane can cut from a cake 15 inches by 8 inches if the thickness of each square is the same as the thickness of the cake?

(A) 21
(B) 28
(C) 32
(D) 34

DL 14.7, 36%

(ST28)



Arrow N is pointing to $\frac{1}{4}$ on the line above. Some number times $\frac{1}{4}$ is equal to 1. Which arrow is pointing to this number?

(A) A
(B) B
(C) C
(D) D

DL 14.6, 36%

The ability to reason, rather than computational skills was prerequisite for more difficult items for 7th graders. However, the easier items were concerned with mathematics with which 7th graders are quite proficient. Pursuing the hypothesis that proficiency on mathematical concepts implies ability to do word problems, and assuming 7th graders are quite proficient in multiplying, adding, and subtracting whole numbers and weaker in dividing whole numbers and work with fractions, one finds a 67% - 40% comparison in performance on five pure computation items in each of those categories.

On items in which reasoning might be classed as a dominant factor, as in the most difficult above, the 4th graders performed at a 45% level and the 7th at 36%. After deductions for guesses, that becomes 27% and 15%. Assuming the average reading level is comparable, that would indicate that the ability to analyze and deduce is a severe handicap to all but the better students--a finding which is not very surprising. One wonders if those percentages may not indicate that very little time is spent in giving students experiences which require reasoning.

Another interesting comparison which emerged from the 4th grade testing is related to a change in wording in two otherwise identical items. The two versions are listed below:

(T14) Ed's house is 24 blocks from Al's. Joe lives halfway between. How many blocks is it from Joe's house to Ed's?

(A) 2
(B) 6
(C) 12
(D) 48

(S14) Ed lives 24 blocks away from Al. Joe lives halfway between Ed's house and Al's. How many blocks must Ed ride on his bike to go from his house to Joe's house?

(A) 2
(B) 6
(C) 12
(D) 48

The change is slight. Ten more words were used in the second version and perhaps a less direct sentence structure. The result was a change in difficulty level from 11.1 to 11.8 and percentage correct from 66 to 62. Not much difference, but it does represent around 4 students in a sample of 100 and it gives some insight into the subleties of helping students learn mathematics. It also challenges one earlier hypothesis concerning mathematical concepts and word problems.

SUMMARY

As part of an assessment of Michigan educational programs, the Department of Education tested nearly all Michigan 4th and 7th graders early in 1970, 71, 72, and 73. Tests in mathematics were prepared by Educational Testing Services and based on specifications drawn from mathematics tests in use in Michigan. The tests were intended to establish the performance level of the 4th and 7th graders in the basic skills and to determine the relationship of the performance and progress to certain socio-economic characteristics.

The mathematics tests were carefully constructed to provide normative information. While mathematics educators have expressed concern about the selection of topics, any such test can include only a sampling of items, and criticisms related to the choices are virtually unavoidable. However, since the performance level was broadly established by the test makers, and since the test was not administered to another norm group, it failed to yield comparative information concerning the level of performance of Michigan youngsters. In mathematics, the tests discerned no change in performance for the first four years of administrations. Public reaction to attempts to determine socio-economic characteristics complicated attempts to relate achievement to those characteristics.

The period covered by the tests was marked by significant fluctuations in local factors which affect school programs, but they have received minimal attention in the interpretation of results. The effects of such critical items as educators per 1000 students, investment per student in real dollars, reduction in staff, and public attitude towards education have not been determined. Perhaps the sheer stability of test results through so turbulant a period represents a substantial accomplishment.

Extensive data drawn from a statistically stable sampling of student performance on each of the test items has been used as a basis for this analysis. That data for subsets of the school districts has not been available for the analysis. It has not been possible, therefore, to differentiate the implications of performance on specific items for school districts which share common characteristics.

The analysis has been further limited by the normative nature of the test. Frequently, items did not isolate concepts. Observations related to the performance on those items reflects a qualitative interpretation of the results. If the improvement of instruction in specific concepts is a goal of the assessment process, then the new criterion referenced tests should be much more useful. However, neither test yields the comparison of performance by Michigan students to other students which is necessary to determine "the achievement level of mathematics in Michigan". (p.4)

Based on this author's examination of the results of

70

student performance on the items, the following remarks describe those features of performance in broad general areas which seemed worthy of note.

On Whole Numbers

Less than 50% of the 4th graders had a functional command of the place value concept of the number system. Performance improved somewhat by grade 7, but even there ~~It~~ must be seen as limiting students' ability to function with ~~the~~ whole numbers. There is also very strong evidence that a large portion of the students did not understand the relation of multiplication to addition.

Perhaps the most significant finding here, one not peculiar to this analysis, is that in most concepts related to whole numbers there is a large number of students who can perform and do understand, as well as another large group who cannot and do not. This is a strong indication of a need for differentiated instruction. Recent MDE interest in identifying more effective delivery systems could foster the development of improved techniques for differentiating instruction.

There are also rather strong indications that 4th graders would find some of the work with fractions easier than the more complex work with whole numbers. That raises the question as to whether some redistribution of instructional time between fractions and whole numbers in grades 3 - 6 might not be helpful.

Fractions

When considering the limited time spent on fractions prior to January of the 4th grade, performance on fraction items was quite good. Fourth graders seemed to be able to relate fractions to geometric regions more easily than to sets of objects. The number line offers very questionable support to their understanding. This is probably related to the fact that, historically, fractions emerged from the necessity to describe part-whole relationships, and while the unit tends to be clearly identified in regions, it is far more obscure in sets and on the number line.

Intuitive understanding of one-half and one-quarter was quite strong and could be useful in work with fractions.

There was very limited opportunity to look at 7th grade performance in this area. However, 7th graders were very weak in their ability to add and subtract common fractions and only the more able were functional with decimal fractions.

Performance by 7th graders on this section of items was very disappointing. It clearly establishes that instruction in fractions was not effective. Whether or not that applied to local districts will need to be determined through local evaluations. However, a more detailed evaluation than this one seems likely to simply add more particulars to the indictment.

Properties of Numbers and Operations

While generalizations concerning numbers have been given a major emphasis in modern mathematics texts, the test responses indicate that only a small number of 4th and 7th graders understood these generalizations. Whether this reflects instructional deficiencies or a lack of learner readiness, present knowledge of those concepts forms a very shaky base for the development of algorithms.

Measurement

Performance on measurement items seems to point to the fact that there are a number of perceptual subleties at work when younger students are requested to work with fractions of units.

Both 4th and 7th graders had difficulty using metric characteristics of squares and parallelograms and only the best students were able to find the area of rectangles with simple dimensions.

Seventh graders also had difficulty with items requiring conversion from one unit of measurement to another.

Performance on problems involving simple applications of numbers to geometric figures or other forms of measurement caused Michigan students a great amount of difficulty. Of all areas on the tests, this was one in which need for improvement was most obvious.

Geometric Properties of Figures

Michigan students seemed to show a strong ability to analyze problems involving distance on a gridded plane. This offers opportunities which we have begun to explore with the geoboard, but perhaps those experiences more rightfully belong in the regular program than among enrichment experiences as has frequently been the case.

Performance on more specific geometric relationships was as weak here as it was in the measurement items. Parallelism and the radius-diameter relationship, for example, troubled 1/2 and 2/3 of the 7th graders tested.

Formulas and Graphs

More than half of the students had difficulty in evaluating simple algebraic expressions involving operations with which they were otherwise familiar. This raises questions concerning the practice of expressing generalizations through the use of such expressions; i.e.,

(a x b) x c = a x (b x c).

About half of the 4th and 7th graders were able to make straight-forward interpretations from bar, line, and pictographs. Seventh grade performance on circle graphs was only slightly lower than that. Because of the frequent practice of presenting quantitative information in this form, work is needed to enable all students to understand such information.

Equations and Inequalities

Many of the comments which might be summarized here have already been discussed. We have already discussed difficulties with the use of letters to represent numbers. Students did, however, demonstrate some ability to work with statements expressed symbolically.

The test indicated a gradual growth in the ability to work with the symbol ">". Most students beyond the 7th grade will understand its meaning.

Word Problems

In spite of an unusually large number of items in this category, performance on them does not support any firm conclusions. However, they do suggest two observations. First of all, while reading is a factor in working word problems, the item statistics seem to suggest that understanding of the mathematical concept involved was probably a more significant factor for most students.

One reviewer remarked that the following item was trivial for fourth graders: "Kathy has only 3 black kittens and 6 white kittens. How many kittens does she have?" The same reviewer did not see these items as trivial. "Jill had 203 stamps and lost 4 of them. How many stamps does she now have?" or "Jack's spelling test has 100 words. He spelled 3/4 of them correctly. How many words did Jack misspell?" The reviewer was obviously focusing on the mathematical concepts and not the reading concepts involved.

A summary of items used on this test suggest that facility with the mathematical concepts is the major stumbling block to solving word problems.

Students had predictable difficulty with items which involved somewhat complex reasoning even though the concepts were quite familiar to them. Despite a "modern" emphasis on mathematical reasoning, this test does not indicate that the ability to reason is a strength of Michigan students.

CONCLUSIONS

First of all the Michigan Department and Board of Education should be congratulated on this massive effort to put Michigan education on a more rational basis. The obstacles to gathering information on this scale are many and the state-level educators must be commended for the hard work and persistance which made these tests a reality. The MDE publications consistently reflect a basic concern that the assessment program should become an instrument for use by local districts to improve education throughout Michigan. Those responsible are also to be commended for recognizing that the vigor and understanding necessary to general improvement must come from the districts and cannot emerge from a rigid regulatory system. As long as that posture is maintained, the assessment progrmm needs and deserves the contributions of Michigan educators and their organizations.

This analysis has raised many questions; it has answered few. It is especially apparent at this point that another reviewer might have detected other significant bits of information and added a more judicious interpretation than that which is reported here. However, the economics of time and money have dictated this narrow base, and the report cannot escape the associated limitation.

It must be concluded that in the area of mathematics these tests did not determine the level of achievement of

Michigan 4th and 7th graders in comparison to any reference group which exists outside of Michigan. That objective for the tests could have been met by using any appropriate nationally normed standardized mathematics test.

The tests did indicate that there was no noticeable change during the first four years. But that assessment did not clearly establish how mathematics performance is affected by such factors as: economic support, class size, student ability, student effort, or other forms of support or interference. Hence, it is not possible to determine whether absence of change represents stagnation, stability, or a tremendous effort against overwhelming odds. It would be very useful to have good information on how such factors do affect the learning of mathematics.

These tests did provide enough information concerning instruction that the time spent in taking them was worthwhile. This is particularly true if local districts examine the data for implications for instruction at the local level. However, if the primary purpose of the tests had been to obtain information to serve as a basis for improving mathematics instruction, a criterion-referenced instrument would probably have provided more useful information. Having made the decision to move in that direction, a sampling program involving many more items would also yield a far greater amount of information. This entire analysis has, in fact, been based on samples of approximately 1000 students each. If it had been designed on that basis, it could have provided much more information. However, the conduct of such a program would require a far greater amount of support for design, administration, and interpretation than has been available for this program.

This analysis deals with average performance by all 4th and 7th grade students in the state. It, therefore, has very restricted meaning for any particular district and even less for a particular classroom. At best, the findings identify a number of critical points which deserve the attention of school districts and classroom teachers. If less than half the 7th grade students in the state can work in a meaningful way with the area of rectangles, 7th grade teachers throughout the state might suspect that their students may have difficulty. More basically, if large numbers of 7th graders in Michigan do have difficulty with symbolic representation of concepts by the use of letters, it is very likely that almost every district (classroom?) has some of these students.

The results of the study do point to at least two characteristics of the mathematics program which require careful examination. First of all, the test performance points to the inappropriateness of teaching all students at a particular grade level as though they were ready to learn the same mathematics. Gagne's studies (21) have pointed out the futility of attempting to learn a mathematical concept without being suitably fortified by the background concepts. Variability in the understanding as well as the ability to use a particular concept is a normal characteristic of a group of students. While some variability can result from ineffective instructional practices, an instructional program which enables students to realize their full potential is likely to increase variability among a group of students even more.

Mathematics instruction responsive to the needs and readiness level of students must be based on differentiated instruction for students at a particular grade level. Management problems have frustrated attempts to achieve a suitable balance between more personalized instruction and opportunities to interact with other students and the teacher. But mathematics instruction simply must become more responsive to the learner.

Secondly, there seems to be continually increasing evidence that the mathematics programs examined by these tests did result in undesirable outcomes for many students. Those students are not learning mathematics in a context which is meaningful to them. They cannot use mathematics to answer questions which are of everyday importance. They do not become quantitatively literate. Is the focus on basic number properties appropriate for some students but not others? Is the focus appropriate but our teaching methods ineffective? Do students have sufficient number sense based on the real world to understand the abstractions or are the abstractions simply not within the perceptual capabilities of some students? We do not know the answer to these questions but we do know that the pursuit of the abstract has produced students who lack understanding of both the abstraction and the real world interpretation.

We must conclude with Dr. Suppes that:

> "as yet theories of learning have little to offer in providing insight into how one learns to think mathematically."

> "...we do not yet understand with any reasonable degree of scientific detail what goes on when a student learns a particular piece of mathematics whether the mathematics in question be first grade arithmetic, undergraduate calculus or graduate school topology."

These observations, coupled with this analysis and data from Piaget on studies of progressive changes in behavior and thought in the developing child, require that we proceed cautiously. We must continue the search for more effective ways to help students understand mathematics.

One last comment concerning the new and promising development in MEAP to apply objective-referenced tests to examine the achievement of minimal objectives. The identification of minimal objectives should sharpen our perception of the effectiveness of instructional practices for certain students. But it carries with it the danger that we will not pursue appropriate objectives for all students. We are inexperienced in the use of minimal objectives and must make certain that, in providing all students with the tools necessary for everyday living, we do not deny to more capable students the understanding necessary to confront the complex problems of an increasingly technical society. We need to provide for both of these extremes and the many variations between them.

APPENDIX A

Mathematics Test Content Specifications

	Grade 4			Grade 7		
	70-71	72	73	70-71	72	73
1. NUMBER AND OPERATIONS	15-20 Items			10-16 Items		
A. Operations with integers (whole numbers)	4	7	7	1	4	3
B. Place-value	2	4	2	1	1	1
C. Properties of integers, divisibility	2	1	1	1	2	1
D. Proper fractions	2	3	4	1	3	4
E. Decimals and percents	0	0	0	1	1	1
F. Properties of operations (Commutative, associative, distributive, closure)	2	2	2	2	2	2
G. Estimation	1	1	1	0	1	1
H. Special properties of zero and one	2	2	2	2	1	1
I. Average	0	0	0	1	1	1
2. GEOMETRY AND MEASUREMENT		ITEMS			ITEMS	
	3	6	6	7	9	10
A. Units of measure, length, weight, time, temperature, money	1	1	1	2	3	4
B. Perimeters and areas of simple polygons	1	1	1	1	3	3
C. Scale drawings and maps	0	1	1	1	1	1
D. Properties of polygons and the circle	1	1	1	2	1	1
E. Angles and intuitive ideas of geometry	0	1	1	1	1	1
F. Non-metric geometry						
3. RELATIONS, FUNCTIONS, GRAPHS	2	2	2	4	4	4
A. Use of mathematical formula	1	1	1	1	1	1
B. Reading and interpreting graphs	1	1	1	3	3	3
4. LOGICAL THINKING	1	1	1	1	1	1
A. Intuitive ideas; Counterexample reasoning	1	1	1	1	1	1

Number of ITEMS

		ITEMS			ITEMS	
5. MATHEMATICAL SENTENCES	2	3	3	4	4	4
A. Equations	1	2	2	2	2	2
B. Inequalities	1	1	1	2	2	2
6. APPLICATIONS						
Word problems (other than those already listed in one of the categories above)	7	8	8	4	6	6

NOTE: At least one-third of the problems could be classified as applications.

APPENDIX B.

CHARACTERISTICS OF THE MICHIGAN ASSESSMENT MATHEMATICS TESTS 1969-73

GRADES 4 & 7

	GRADE 4				GRADE 7			
	1970 1005	1971	1972	1973	1970 1005	1971	1972	1973
Number of items	30	40	40	40	30	40	40	
Raw Score average	16.3	21.78	22.96	24.5	16.0	21.4	23.8	24.1
% correct average	54.3	54.5	57.4	61.2	53.3	53.5	59.5	60.3
Standard Deviation raw score	5.7	7.29	7.27	7.7	5.9	7.7	8.1	7.8
Percentage Score	18.0	18.2	18.2	19.2	19.7	19.2	20.2	19.5
Reliability	.83	.86	.87	.89	.84	.87	.87	.89
Difficulty	54.3	52.5	57.3	61.2	53.3	53.1	59.7	61.2
SPEEDEDNESS - in terms of percentage of students who reached								
the last item	88	87	86	90	90	87	90	95
the 3/4 point	97	96	96	97	98	97	97	99
Discrimination		.48	.52			.51	.55	
Equated mean scores	50.0	50.5	50.6	50.9	50.0	50.0	50.0	50.3
Item difficulty scores	12.5	12.7	12.1		12.6	12.6	11.9	

APPENDIX C

TEST QUESTION ATISTICS (ITEM ANALYSIS)

Item Analysis is a detailed statistical description of how a particular question functioned when it was used in a particular test. The analysis provides information about the difficulty of the question for the sample on which the analysis is based, the relative attractiveness of the options, and how well the question discriminated among the examinees with respect to a chosen criterion. The criterion most frequently used is total score on the test of which the item is a part. However, the criterion may be the score on a subtest, on some other test or, in general, any appropriate measure that ranks the examinees from high to low.

The portion of a typical item analysis that a committee member is likely to meet is reproduced below, followed by an explanation of each of the designated entries. The analysis is based on a sample of answer sheets carefully selected to be representative of the total group that took the test.

FORM MEP	BASE N 370	OMIT 30	A 209*	B 41	C 50	D 13	E 23	M_{TOTAL} 13.1	$\hat{\Delta}_E$ SCALE JBP	$\hat{\Delta}_E$ 12.6	CRITERION IS45
TEST CODE MATH	ITEM NO. 26	M_O 10.6	M_A 14.4	M_B 13.0	M_C 10.7	M_D 11.4	M_E 10.2	P_{TOTAL} .99	P_+ .57	$\hat{\Delta}_C$ 12.4	r_{bis} .49

The Four Boxes to the Left of the Doubles Lines:

Form and Test Code: These identify the test. The first letter designates the year (M is 1964, Q 1968); BP means Advanced Placement. Test Code gives subject.

Base N: Number of answer sheets in the sample.

Thus: These four boxes say that this is the analysis of question 26 on the 1964 Mathematics Advanced Placement Examination, based on a sample of 370 papers.

The Twelve Boxes Between the Two Sets of Double Lines:

On the top row, the box labeled OMIT shows the number of individuals in the sample who omitted this question but answered a subsequent one and the boxes labeled A-E show the number who chose each option.

Thus, OMIT 30 means that 30 examinees skipped this question but answered at least one question later in the test.

(Note: a person is considered to have dropped out after the last answer he has marked; dropouts are not included in the analysis of a question. Omits are assumed to have considered the question and are included with incorrect responses in computing the difficulty index for a question.)

The Key is marked with an asterisk. In this case 209 individuals chose A, the correct answer.

On the second row, M_O, M_A, M_B, etc. indicate the average ability level or mean criterion score of the examinees who chose each option. This mean criterion score is an index describing the average ability level of the candidates on a scale which has a mean of 13 and a standard deviation of 4.

For example, if the criterion being used is the score on the total test, then the average score of the entire sample on the test is assigned the index 13.0, which is considered t be the mean criterion score of the total sample. If the average of scores of the group choosing an options is above the sample average, their mean criterion score will be greater than 13.0; if their average is below the sample average, it will be less than 13.0. These criterion scores are related to percentile rank of performance on the criterion approximately as follows:

Criterion Score	Relative Rank on Criterion
20 or above	highest 5%
18 or above	highest 10%
16 or above	highest 20%
14 or above	highest 0%
above 13	upper half
below 13	lower half
12 or below	lowest 40%
10 or below	lowest 20%
8 or below	lowest 10%
6 or below	lowest 5%

In the example, the mean criterion score M_A of those choosing the correct answer A was 14.4, higher than that of any other group. The weakest group on the average were the 23 who chose E and whose mean criterion score of 10.2 puts them in the bottom quarter of the sample.

The Eight Boxes to the Right of the Double Lines:

The meanings of the labels in this group of boxes are as follows:

P. TOTAL is the per cent of the sample still answering questions; i.e., 99% answered question 26 or a subsequent question. (The dropout at this point is 1%.)

M TOTAL is the mean criterion score of the P TOTAL. In this case the dropouts tended to be below average in ability so that M TOTAL is 13.1, which is slightly higher than the 13.0 for the complete sample.

P+ (per cent pass) is the per cent of the P TOTAL that answered the question correctly. In this case, 57% of 366, or 209.

Clearly, r itself is not a very stable statistics and one can expect it to vary from one use of the item to another. The r biserial of an item will be affected by such things as the extent to which the item measures what the test as a whole measures, the appropriateness of the question for the particular group of examinees, and the amount of variability within the examinee group with respect to the ability being tested.

BIBLIOGRAPHY

1. Michigan Department of Education, Purposes and Procedures of the Michigan Assessment of Education. Bureau of Research, Michigan Department of Education, September 1969.

2. Michigan Department of Education, Activities and Arrangements for the 1969-70 Michigan Assessment of Education. Bureau of Research, Michigan Department of Education, September 1969.

3. Michigan Department of Education, Objectives and Procedures: of the Michigan Educational Assessment Program 1970-71. Bureau of Research, Evaluations, and Assessment, Michigan Department of Education, December 1970.

4. Michigan Department of Education, Objectives and Procedures: The First Report of the 1971-72 Michigan Educational Assessment Program.. Research, Evaluation and Assessment Services, Michigan Department of Education, October 1971.

5. Michigan Department of Education, Objectives and Procedures: The First Report of the 1972-73 Michigan Educational Assessment Program. Research, Evaluation and Assessment Services Unit, Michigan Department of Education, October 1972.

6. Michigan Department of Education, Levels of Educational Performance and Related Factors in Michigan., Michigan Department of Education, Assessment Report no. 4, 1970.

7. Michigan Department of Education, Distribution of Educational Performance and Related Factors in Michigan, Michigan Department of Education, Assessment Report no. 5, 1970.

8. Michigan Department of Education, Distribution of Educational Performance and Related Factors in Michigan: A Supplement. Michigan Department of Education, Assessment Report no. 8, July 1971.

9. Michigan Department of Education, Levels of Educational Performance and Related Factors in Michigan. The fifth Report of the 1970-71 Michigan Educational Assessment Program, Michigan Department of Education, June 1972.

10. Michigan Department of Education, Distribution of Educational Performance and Related Factors. The Sixth Report of the 1970-71 Michigan Educational Assessment Program, Michigan Department of Education, June 1972.

11. Michigan Department of Education, Staff Reply to the MAPEA Task Force Papers, An undated paper prepared under the auspices of the Michigan Department of Education.

12. Michigan Department of Education, A Response to the Major Issues in the Paper Entitled "A Critical Assessment of Michigan School Assessment Program", MDE April 1971.

13. Michigan Department of Education, Individualized Pupil Report: Explanatory Materials, The Second Report of the 1972-73 Michigan Education Assessment Program, MDE, March 1973.

14. Michigan Department of Education, Minimal Performance Objectives for Mathematics Education in Michigan. Michigan Department of Education, 1973.

15. Educational Testing Service, Technical Report of Selected Aspects of the 1969-70 Michigan Educational Assessment Program. Michigan Department of Education, August 1971.

16. Educational Testing Services, Technical Report: The Ninth Report of the 1970-71 Michigan Assessment Program. Michigan Department of Education, June 1972.

17. Educational Testing Services, Technical Report: The Fifth Report of the 1971-72 Michigan Educational Assessment Program. Michigan Department of Education, November 1972.

18. Educational Testing Services, The Equating Report: Year to Year Analysis of the Cognitive Tests of the Michigan Educational Assessment Program 1970-72. Michigan Department of Education, March 1973.

19. Educational Testing Services, Manual of Information for Test Committee Members. ETS, Test Development Division, September 1967.

20. Fawcett, Harold P. "Teaching for Transfer," The Mathematics Teacher, XXVIII (December, 1935) 465-71.

21. Gagne, Robert M., "On the Learning of Mathematics. A Product Orientation," Readings in Secondary School Mathematics, Edited by Douglas B. Aichele and Robert E. Reys. Boston: Prindle, Weber & Schmidt Inc., 1971.

22. Kearney, C. Phillip, Needs Assessment at the State Level in Michigan: Why and How? Remarks to an NASE Seminar, Denver Colorado; October 2, 1972.

23. Michigan Association of Professors of Education Administration, Task Force in Educational Assessment and Accountability, The Michigan Educational Assessment Program: An Abstract. A Report Adopted by MAPEA, April 1972.

24. Oakland Intermediate School District, A Critical Assessment of the Michigan School Assessment Program. Oakland Schools, Pontiac, Michigan, March 1971.

25. Remmus, H. H., Gage, N.L. and Rummel, J. Francis, A Practical Introduction to Measurement and Evaluation, Harper & Row Publishers, New York, 1960.

26. Suppes, Patrick, "Some Theoretical Models for Mathematics Learning," Journal of Research and Development in Education, I (Fall, 1967), 5-22.

27. Thorndike, Robert L., Educational Measurement, American Council on Education, Washington D.C., 1971.

28. Wells, David and Willoughby, Steven S., "The Forum: Accountability," Mathematics Teacher, November 1972, 588, 589, 655-663.

29. Womer, Frank B., Working Paper for Establishment of Long Range MEAP Goals, September 1973.

The following were members of the Guidelines Committee for Quality Mathematics Teaching during the development of this monograph:

Maja S. Barr	Saginaw
Richard Debelak	Iron Mountain
Theresa Denman	Detroit
Jacqueline Dombroski	Petoskey
Herbert Harmon	Kalamazoo
Lou Henkel	Grand Rapids
Jean Houghton	Weidman
Daniel Korman	West Branch
Evelyn Kozar	Detroit
Bea Munro	Ann Arbor
Charles Schloff	Dearborn Heights
Geri Westover	Bay City
James K. Bidwell (Chairman)	Mt. Pleasant

Your comments and criticisms of this monograph as well as suggestions for other monographs (or manuscripts for them) can be sent to:

Albert Schulte
Oakland Schools
2100 Pontiac Lake Road
Pontiac, MI 48054