

Submission to Joint Dissemination Review Panel

- I **NAME OF PRODUCT:** Comprehensive School Mathematics Program (CSMP)
- II **DEVELOPED BY:** CEMREL Inc., 3120 69th St., St. Louis, MO 63139
Key people: Burt Kaufman, Director of CSMP through 1979
Clare Heidema, Director of CSMP from 1979
Martin Herbert, Director, Mathematics Research and Eval Studies
- III **SOURCE AND AMOUNT OF FUNDING:** National Institute for Education, approx \$7 million
- IV **SOURCE AND AMOUNT OF EVALUATION FUNDING:** NIE, approx \$1 million
- V **YEARS DEVELOPED AND EVALUATED:** 1972-1982
- VI **SUMMARY:**

Students in CSMP are better able than comparable Non-CSMP students to apply various problem solving processes, such as using patterns and relationships. This claim is based on comparative testing at each grade level from grades 2-6 involving an average of about 60 classes per grade, using Analysis of Covariance on class means. Additional analyses at the school, district and student level, and by sex and ability of students, support this claim. CSMP students also perform at least as well on the traditional arithmetic skills, a claim based on Analysis of Covariance data from large numbers of classes in grades 2 to 6.

VII **DESCRIPTION OF PRODUCT**

The impetus for this program was the need to improve several shortcomings in mathematics education: the static content of the curriculum, the rote method in which it is usually taught, and the lack of materials for teaching mathematical thinking skills to students. CSMP is an elementary school mathematics program intended for regular classroom usage, which features new content, the use of special pictorial devices, a spiral approach and an emphasis on problem solving through student materials and detailed lessons in the teacher's guides. The main materials associated with the program are as follows:

- ° Teacher's Guides at each grade contain a master schedule of activities and a detailed lesson plan for each activity. There are between 2 and 6 guides per grade level ranging in length from about 500 pages in Kindergarten to about 1900 pages in sixth grade.
- ° Student materials consist of worksheets to accompany individual lessons and workbooks, which are 16-page booklets covering larger units of work. There are between 100 and 200 worksheets per grade and between 4 and 16 workbooks per grade, depending on grade level.

Claims of Effectiveness

1. CSMP students perform at least as well in traditional arithmetic skills as comparable Non-CSMP students.
2. CSMP students are better able than comparable non-CSMP students to apply the mathematics they have learned to new problem situations using processes involving:

Relationships and Number Patterns
Production of Multiple Answers
Mental Arithmetic
Word Problems
Estimation
Number Representations
Pre-algebra
Prediction

Intended Beneficiaries The program is intended for use in regular, heterogeneously-grouped classrooms and is now the mathematics curriculum for about 55,000 students in over 100 school districts, including use with gifted, Chapter 1, and non-English speaking students (though no special claims are made for these populations).

Characteristics of Development Group Materials were developed on a day-to-day basis in regular classrooms in an inner suburban St. Louis school district. The classes were near the national average in achievement scores and in racial composition.

Resources Required The program is to be taught by a regular classroom teacher and to be supervised by a locally-designated coordinator, most often a district mathematics supervisor. No other personnel are required, nor is any special equipment or facility beyond the normal classroom. Depending on grade level, between 6 and 30 hours of training are highly recommended (although not required) and training arrangements are determined by the local district. A network of qualified "turnkey" trainers is available to adopting sites if desired.

Typical personnel training costs range from \$0 per teacher (for example, when the coordinator conducts the training in two regularly scheduled staff development days, followed by monthly two-hour in-school sessions) to approximately \$350 per teacher (for example, when a consultant conducts a one-week workshop for teachers who are paid a daily stipend).

The approximate costs-per-student of all materials, based on present, moderate-sized printing runs, are shown below for kindergarten, grades 1-3 (average) and grades 4-6 (average).

Table 1
Materials Costs Per Student.

	<u>Installation (Year 1)</u>			<u>Subsequent Years</u>		
	K	1-3	4-6	K	1-3	4-6
Teacher Materials	\$1.50	\$2.30	\$3.70	—	—	—
Student Materials	\$3.70	\$6.20	\$7.10	\$3.50	\$5.60	\$6.70

In addition, beginning in fourth grade, one hand calculator is recommended for every two students. Calculators can be drawn from existing school supplies, purchased separately, or provided by students themselves.

VIII DESCRIPTION OF EVALUATION DESIGN

General Evaluation Activities

The evidence presented was generated by CEMREL'S Mathematics Research and Evaluation Studies (MRES) project, which operated and was funded independently of the CSMP development group. Its activities were monitored by an external Evaluation Panel chaired by Dr. Ernest House. A 50-volume Evaluation Report Series describes the complete set of evaluation data.

The initial phase of the development cycle of CSMP materials at each grade level culminated in a printed Experimental Version of the curriculum. The materials were then tried out for two years in that grade in what were called "Extended Pilot Tests". The first year of each Extended Pilot Study focused on a small number of classes in the St. Louis region. This trial was used to obtain preliminary evaluation results and to develop evaluation procedures and instruments. In the second year of the Extended Pilot Test, larger numbers of classes in many geographic locations were tested.

Experimental Design

During the second year of the Extended Pilot Tests the curriculum was used in regular classrooms under normal conditions. Materials and training costs were borne by participating districts who agreed to cooperate in data gathering activities.

Participating schools began using CSMP materials in the lower grades. The most common strategy was to begin all their kindergarten or first grade students in CSMP; in each succeeding year those students advanced one year in the curriculum while new groups started CSMP from that first level. Thus, in the later grades, teachers did not volunteer for the program but more or less "inherited" it and their CSMP students from the previous grade level. For most of these teachers, teaching in the Extended Pilot Test was their first experience with CSMP. They received training during the summer or early fall, through either a CSMP-run workshop or a workshop conducted on site by the local CSMP coordinator.

The design of the testing program was comparative in nature. Control classes were selected jointly by CEMREL and the participating districts. Since CSMP was being used at a given grade level throughout the school, control classes were chosen from another nearby school with similar students and teachers. In some cases, particularly in sixth grade, the program was being used district-wide and control classes were not available from within the district. In such cases they were selected from CSMP schools in other districts, but where CSMP was being used only at lower grade levels and had not yet reached the grade level being tested (that is, the control classes had no previous CSMP experience).

CSMP students usually had been studying CSMP since at least first grade while at the same time the Non-CSMP classes had been using their district's regular mathematics curriculum, which was almost always a commonly used math series from one of the large text-book publishers.

Testing took place in May each year, using standardized math tests and/or the MANS Tests (see next page). Included in the testing program was a standardized reading test whose scores were used as a covariate in the analysis. Class mean scores were calculated and an Analysis of Covariance was performed on the class means. Students who had entered the program after October, whether CSMP or Non-CSMP, were excluded from this main analysis.

A total of 27 school districts participated in these comparison studies, at least 9 per grade level with some districts participating at more than one grade level. These 27 districts were distributed as follows:

<u>Type of Community</u>	<u>Geographic Location</u>
7 large city	7 east
12 suburban	8 central
4 medium city	6 upper midwest
4 small city /rural	3 south
	3 west

The number of classes participating at each grade level is shown in Table 2.

Table 2
Participating Classes by Grade Level

	Number of Classes		Mean Reading Percentile Rank	
	<u>CSMP</u>	<u>Non-CSMP</u>	<u>CSMP</u>	<u>Non-CSMP</u>
Second	57	50	51	50
Third	42	33	56	55
Fourth	30	21	64	62
Fifth	31	25	61	60
Sixth	26	37	77	78

It can be seen that the CSMP and Non-CSMP classes were well matched in reading ability, there being no significant differences between the two groups in any year.

IXa EVIDENCE OF EFFECTIVENESS FOR CLAIM 1

Table 3 summarizes all of the available data from mathematics computation tests in comparison studies, grades 2-6. The adjusted class means were calculated using an Analysis of Covariance on the class means with reading score as covariate. Separate studies were conducted in individual districts in grades 2 and 3; districts were combined in grades 4-6.

Table 3
Comparison Data, Computation Tests

Grade	Test	Number of Classes		Number of Items	Adjusted Means		Signif at .05	In Favor of
		CSMP	Non-CSMP		CSMP	Non-CSMP		
2	CTBS	13	13	28	21.6	20.3	*	CSMP
		15	13		20.6	19.3		CSMP
		6	6		20.5	18.6		CSMP
	ETS Coop Prim ¹ Stanford Ach	6	6	55 percentile ranks	36.2	35.2		CSMP
		6	6		47	52		Non-CSMP
3	CTBS	12	12	48	36.1	35.4		CSMP
		15	12		35.8	36.8		Non-CSMP
		7	8		37.2	33.3		*
	ETS Coop Prim ¹	6	12	55	42.3	41.8		CSMP
4	SAT + MANS ²	30	21	48	27.0	28.2		Non-CSMP
5	CTBS	31	25	48	34.9	34.3		CSMP
6	MANS Computation (modelled after standard. tests)	26	37	34	27.0	26.7		CSMP

¹Total Math; This test does not have a separate computation section
240 items from the Stanford Ach Test + 8 items on fractions from the MANS Tests

CSMP classes had higher scores in 9 of the 12 studies, including the only 2 significant results. This supports Claim 1, that students in CSMP perform at least as well as Non-CSMP students in traditional arithmetic skills.

IXb EVIDENCE OF EFFECTIVENESS FOR CLAIM 2

The MANS Tests

Introduction

The MANS Tests (Mathematics Applied to Novel Situations) are a series of short tests, different at each grade level, designed to assess some of the underlying thinking skills taught through CSMP. They were developed by CEMREL because suitable standardized mathematics tests for measuring such skills are not available. Development of such tests has been recommended by both the National Assessment of Educational Progress (NAEP 1983):

"The very things that are difficult to teach are often difficult or expensive to test. Educational leaders need to pressure test developers to include items that reflect the higher level objectives of the curriculum."

and by the National Council of Teachers of Mathematics (NCTM 1980):

"The evaluation of problem-solving performance will demand new approaches to measuring. Certainly present tests are not adequate."

The MANS Tests use standard terminology and do not contain any of the specific language or typical problem activities of CSMP. The tests use straightforward language and most of them present mathematical situations which are unfamiliar to CSMP and Non-CSMP students alike.

At each grade level, the MANS Tests consist of several short tests, each with its own standardized directions which a specially trained tester uses in explaining the task and sample items to the class. Liberal time limits allow almost all students to finish. For most tests, students produce their own answers instead of selecting from given alternatives. The reading requirements are kept intentionally low relative to grade level.

Reliability and Validity

Developmental Procedures At each grade level, there were two years of activities including outside review; pilot testing in at least 5 local classes; test and item analysis; and revision.

Coverage Standardized mathematics tests usually have 3 sections. Two of these, computation and word problems, are explicitly covered in MANS, partially through the "rental" of standardized achievement subtests from publishers. The third section, concepts, is integrated throughout MANS. The average number of mathematics items in seven leading standardized tests (CAT, CTBS, ITBS, MAT, SAT, STEP and SRA) is shown below. There are at least three times as many non-computation items in the MANS Tests as in the standardized tests.

	Number of Computation Items		Number of Other Items	
	Standardized	MANS	Standardized	MANS
Grade 2	31	18	38	142
Grade 3	38	22	34	138
Grade 4	39	48	53	198
Grade 5	41	54	54	237
Grade 6	41	34	55	339

Outside Review During test development, all tests were reviewed by the external CSMP Evaluation Panel which included distinguished scholars in mathematics, assessment and evaluation, and mathematics education. There were also reviews by education practitioners.


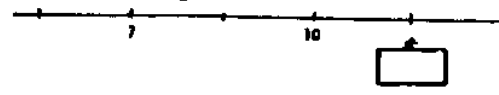

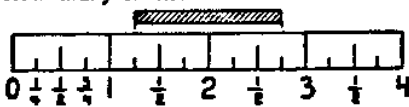

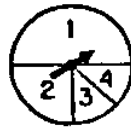

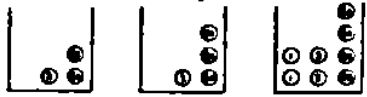
Reliability The reliability/internal consistency (KR20 corrected by Spearman-Brown for an equivalent 20-item test) was calculated for each of the 85 individual Mans tests. The reliability was above .80 for 72 of these tests; between .75 and .80 for 10 tests; and below .75 for 3 tests (.68, .71, and .72). The median KR20 was .86. Correcting for an equivalent 30-item test, a more usual number for standardized tests, produced KR20's above .80 for 83 of the 85 tests.

Correlations with Other Measures The median correlations between Reading scores and Total MANS scores were .60, .57, and .56 for grades 2, 4 and 6 respectively. The median correlations between standardized computation scores and Total Mans scores was .63. The median correlations between Total Mans and teacher estimate of student's problem solving ability was .59.

Student and Teacher Ratings Mean teacher ratings of importance of individual MANS tests, collected in 4th and 5th grades, were 4.3 and 4.1 on a 5-point scale. Mean rating of how well students liked individual MANS Tests, collected only in 4th grade, was 3.0 on a 4-point scale.

MANS Categories Individual MANS Tests are grouped into categories according to mathematical process considered by the CSMP Evaluation Panel to be generalized processes appropriate to problem solving at the elementary grades. Several of the basic goals espoused by both the National Council of Teachers of Mathematics and the National Council of Supervisors of Mathematics, including "using mathematics to predict", and "estimation and approximation" are included as MANS categories but do not appear in standardized tests. The next page lists the MANS categories and shows sample items from each category. The actual student format was much more extensive and was preceded by standardized directions and sample items explained by the tester. Items in the Estimation category had short time limits. A few item types were repeated, with different items, in two or more grades.

Sample of MANS Items by Grade Level

Category	Second Grade	Fourth Grade	Sixth Grade					
Number Patterns, Relationships	<p>2 - 10 8 - 4 10 - 5 6 - ?</p> <p>Which is Larger? 585 + 250 or 580 + 290</p> <p>What is the missing number? 28, 25, ?, 19, 16, 13</p>	<p>Will 700 be in any of the boxes? .. <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td>60</td><td>67</td><td>74</td><td>81</td><td>88</td></tr></table> ..</p> <p>What are the missing numbers? ?, 50, ?, 200, 400, 800</p> 	60	67	74	81	88	<p>Which is larger? $\frac{5}{2}$ or $\frac{5}{4}$ 0.9 or 0.11111</p> <p>Name a fraction (decimal) that is: larger than $\frac{1}{3}$ but smaller than $\frac{1}{8}$ larger than 0.2 but smaller than 0.3</p> <p>Label the missing number</p> 
60	67	74	81	88				
Multiple Answers	<p>Write # Sentences about 8</p> <p>8 = 9 - 1 8 = 3 + 4 + 1 8 = 2 x 4</p>		<p>Take out 3 balls together Add to get total score Give all possible scores</p> 					
Mental Arith	<p>300 - ? = 250</p>	<p>Hit = gain 5, miss = lose 1 start with : 3 below zero end with : 5 above zero # of misses : 2 # of hits : ?</p>	<p>12 x 75 = 900 13 x 75 = ?</p> <p>$\frac{1}{2} \times ? = 40$ 0.75 - 0.5 = ?</p> <p>scratch work not allowed</p>					
Word Problems	<p>One-stage, with pictures, and read to the students</p>	<p>Two-stage Miscellaneous data Word problem approximations</p>	<p>Three-stage Word problems with fractions, decimals Novel word problems</p>					
Estimation	<p>90 - 12 is in which interval 0 - 10 - 50 - 100 - 500?</p>	<p>602 is about ? as large as 298? 2, 5, or 10 times</p>	<p>$\frac{1}{2} + 8$ is: (<1 or =1 or >1)?</p> <p>Which interval contains 1,002.5 + 21.5? 0 - 1 - 10 - 20 - 50 - 100?</p>					
Number Representations	<p>Write "two thousand, eleven"</p> <p>100 more than 901 is ?</p>	<p>How many inches? </p> <p>Name the 2nd largest 4-digit number using only 2, 5, 7, 8?</p>	<p>Which are equivalent to $\frac{1}{3}$? $\frac{2}{6}$ $\frac{11}{31}$ $\frac{3}{15}$ $\frac{4}{12}$ $\frac{50}{150}$</p> <p>Which are equivalent to $\frac{3}{4}$? 0.750 0.075 0.75 7.5 75.0?</p> <p>Put an arrow at 1.35 in.</p> 					
Predicting	<p>(4th and 6th grades only)</p>	<p>Spin 100 times How often is the score 2?</p> 	<p>Spin both spinners together How often is the total > 9?</p>  <p>From which box should you make a blind draw</p> 					
Pre-algebra	<p>(sixth grade only)</p>		<p>$(1)(3) = 1 + 2 + 3$, so $(1)(99) = (2)(99)$</p> <p>If $k + 2 + k + 7 = 13$, then $k = ?$</p> <p>Apply transformations - geometric rotations and/or symbol reversals - to various figures</p> <p>If $q = 5$, then $2 \times q^2 = ?$</p>					

Sixth Grade Results

Data will be presented in detail for the sixth grade in order to show effects on students at the end of the CSMP curriculum. Then there will be a shorter presentation of data from the earlier grades to show the consistency of the findings across districts and grade levels.

Analysis of Class Means

Because the treatment, CSMP, was administered at the classroom level, class means were the primary unit of analysis (though student level data is also shown, next page). Table 4 shows adjusted means across the 26 CSMP and the 37 Non-CSMP classes for each MANS category from the Analysis of Covariance on class means, with the Gates-McGinitle Vocabulary Test as covariate. The adjustment in means due to differences in vocabulary scores between CSMP and Non-CSMP classes was always small, less than 1%. Also shown is the effect size (the difference in class means divided by the standard deviation of the control means).

Table 4
Summary Class Mean Data, Sixth Grade

MANS Category	Adjusted Means CSMP	Adjusted Means Non-CSMP	p-value (1,60)	Effect (difference in adj. means - Size stand dev of control means)
Relations, Patterns	46.1	40.3	.01	1.00
Multiple Answers	38.8	31.9	.01	.91
Mental Arithmetic	31.5	28.3	.01	.63
Word Problems	15.1	13.6	.01	.56
Estimation	24.4	22.5	.01	.41
Number Representations	28.8	26.3	.01	.38
Pre-Algebra	30.0	27.5	.01	.47
Predicting	15.2	13.9	.01	.52
Total MANS	229.9	204.3	.01	.63

It can be seen that CSMP classes had higher scores than Non-CSMP classes on all categories and this difference was significant at the .01 level each time. Figure 1 shows the performance of these 63 classes in graphical form. Each entry represents a class, with average MANS score plotted against vocabulary score. The regression line on the graph is the best linear predictor of MANS score for as given Vocabulary score.

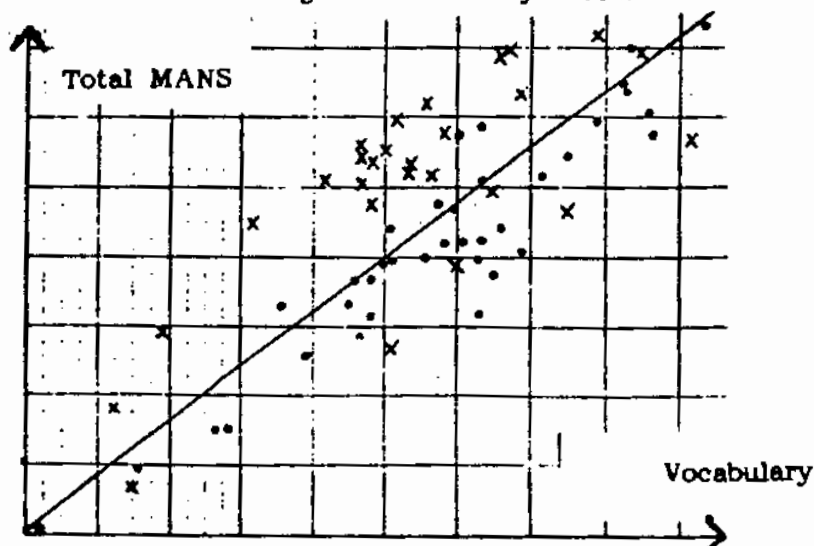


Fig 1. 6th Grade Class Means
(x = CSMP class, o = Non-CSMP)

Other Levels of Analysis

Figures 2 and 3 show school and district means. Each entry represents a school or district, with MANS score plotted against Vocabulary.

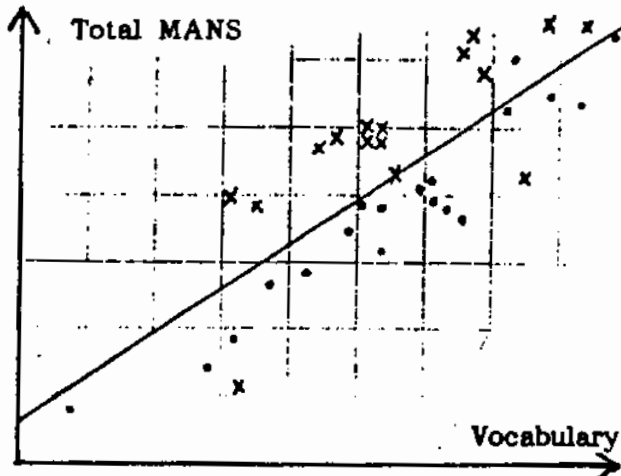


Fig 2. 6th Grade School Means
(x = CSMP school, o = Non-CSMP)

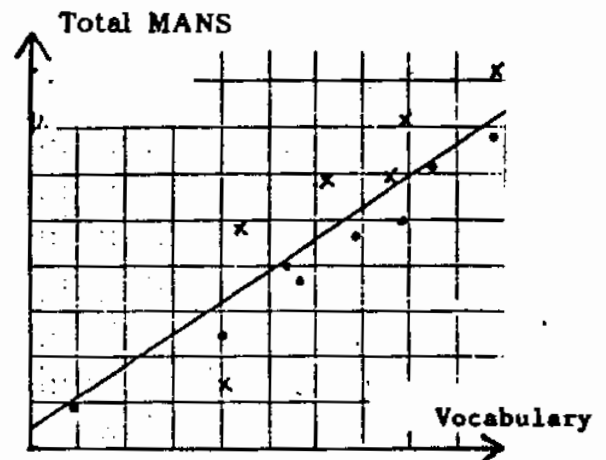


Fig 3. 6th Grade District Means
(x = CSMP district, o = Non-CSMP)

Figure 4 and Table 5 show student level data. In Figure 4, students are grouped into quartiles according to their percentile rank on the Vocabulary test. Average MANS scores are shown separately for each quartile of CSMP and Non-CSMP students. Table 5 shows MANS scores according to sex of student. These various data show the advantage of CSMP for classes and students at various ability levels and regardless of sex.

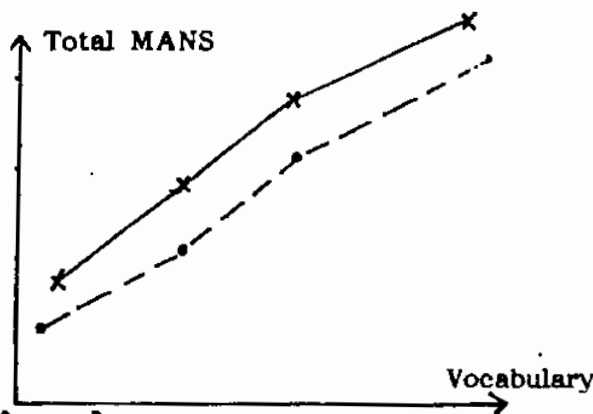


Fig. 4. Student Means Grouped by Reading
(x = CSMP students, o = Non-CSMP)

Table 5

MANS Scores by Sex of Student

Boys:

CSMP 160.4

Non-CSMP 147.3

Girls:

CSMP 154.3

Non-CSMP 140.0

CSMP "Graduates"

Since the CSMP development has only recently been completed, there have not been many "graduates". However, one study was conducted in the largest CSMP site in the St. Louis area. Seventh grade math teachers, inexperienced in teaching CSMP, were asked to rate their students, who were mixed former CSMP and Non-CSMP students. Students were rated on: participation, motivation, creativity and problem solving, and practical applications. The mean adjusted composite rating for the 55 former CSMP students was 12.1 versus 11.3 for the 210 former Non-CSMP students. This difference was significant at the .20 level, a suggestive difference given the usual unreliability of such subjective measures.

In addition, former CSMP students had higher mathematics grades for each quarter, the adjusted means being 3.9 versus 3.6, 3.8 versus 3.5, and 3.7 versus 3.5 (where A = 5, B = 4, etc.). The first two differences were significant at the .05 level, the other at the .10 level.

Educational Significance

In order to assess educational significance, CSMP students' performance on the MANS Tests was compared with similar gains on standardized tests. Using effect size on student level data, the CSMP advantage was .37 raw score standard deviations. On the five leading standardized tests for which this data was available, an increase of 1/3 of a raw score standard deviation corresponds to an improvement from the 50th percentile to an average of the 61st percentile, and from the 75th percentile to about the 85th percentile. If one translates the results into simple percentage terms, the gain is from the 50th to about the 63rd percentile.

The size of the CSMP advantage on the MANS Tests is also roughly comparable to two findings of national significance. First, the 40-point decline in the Mathematics section of the Scholastic Aptitude Test from 1963 to 1970 is equivalent to about 5 items on a 60-item test, or less than 1/2 of a raw score standard deviation. Second, the "most salient finding" of the recent national assessment in mathematics (NAEP 1983) was that "13-year-olds have improved dramatically between 1978 and 1982" (the improvement was about 3 percentage points) and that "of particular significance is the 8 percentage point gain for 13-year-olds in heavily minority schools."

Thus the CSMP advantage on the MANS Tests is an educationally significant result in itself but more so because of the nature of the MANS Tests which are based on applications of mathematics to novel situations. Also described in the 1983 national assessment report is the difficulty of making improvements in this area (NAEP 1983):

"With one exception, there was very little change in problem solving performance between 1979 and 1982. The one exception is that 13-year-olds showed significant growth in solving routine problems - i.e., word problems of the type usually found in textbooks and practised in school...Most of the routine verbal problems can be solved by mechanically applying a computational algorithm...Even the 13-year-olds, who made significant gains on routine problem solving, showed no change in their performance on non-routine problems."

From the same report, in a discussion of the major implication of the findings:

"Schools are doing a good job of teaching mathematical topics that are relatively easy to teach ... there was very little change in topics that are relatively difficult to teach, such as non-routine problem solving...Changes at the higher cognitive levels will occur only when higher-level cognitive activity becomes a curricular and instructional focus."

Results from Other Grade Levels

Table 5 shows summary MANS data for grades 2-5. Adjusted means are given; the size of the adjustment due to differences in reading ability was always small, the largest being 1.1 %.

Table 6
Summary Class Mean Data, Grades 2-5

Grade	Adjusted Means		Signif ¹ at	Number of Categories Tested	Number of Categories Significant (p < .05)
	CSMP	Non-CSMP			
2	75.3	66.5	.01	7	6
3	96.2	85.4	.01	7	5
4	112.7	96.1	.01	7	6
5	147.8	131.7	.01	8	7

1 F-tests with 104, 74, 48 and 53 degrees of freedom respectively.

Figures 5-8 show graphs of class means for grades 2-5; x = CSMP class, o = Non-CSMP.

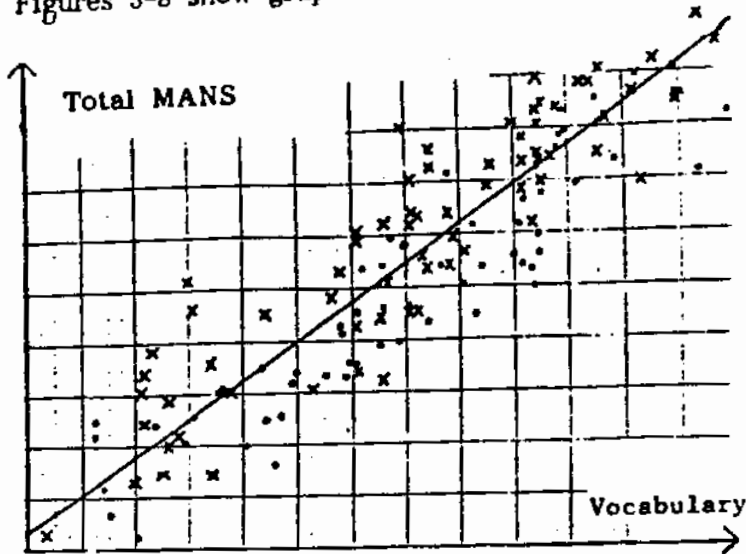


Fig. 5. Second Grade Class Means

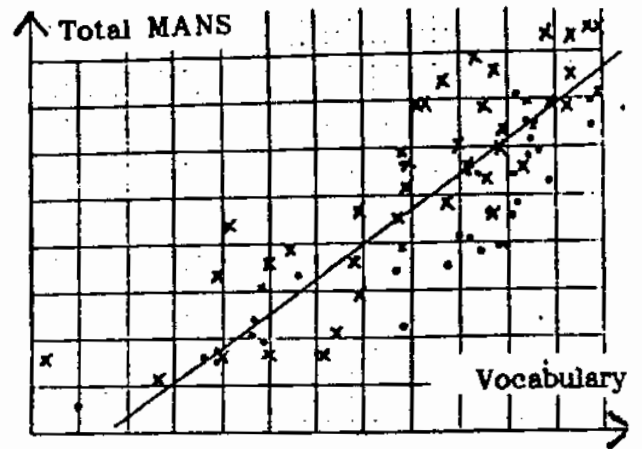


Fig. 6. Third Grade Class Means

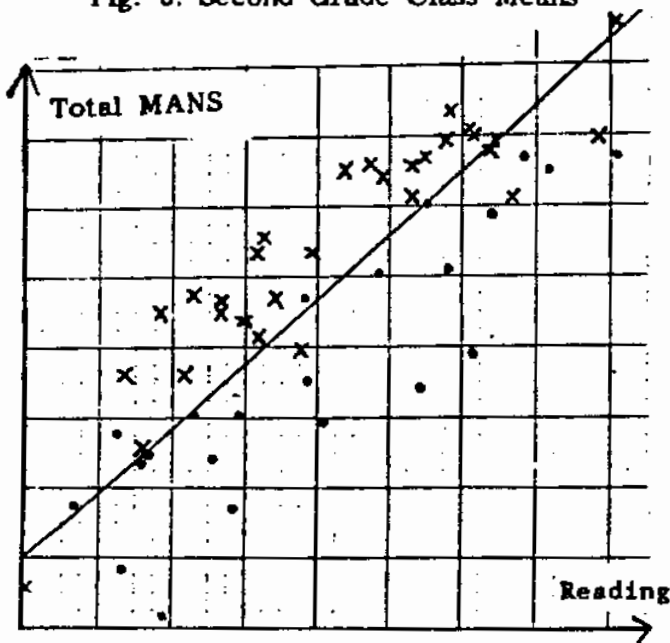


Fig. 7. Fourth Grade Class Means

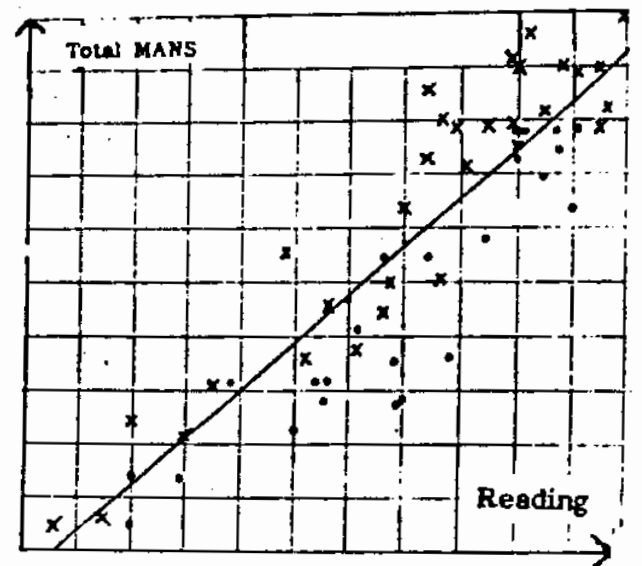


Fig. 8. Fifth Grade Class Means

Figures 5-8 illustrate the consistency with which CSMP classes outperformed comparable Non-CSMP classes. Other analyses at these grade levels show a similar consistency when the data are analyzed at the individual student level by reading score, sex of student and teacher estimate of student's problem solving ability. Joint research studies between CEMREL and individual districts, conducted at various other times, uniformly produced differences in favor of CSMP, with the difference reaching significance in 7 out of the 9 studies.

References

National Assessment of Educational Progress (NAEP). The Third National Mathematics Assessment: Results, Trends and Issues. 13-MA-01. Denver, Colo: Educational Commission of the States, 1983.

National Council of Teachers of Mathematics. An Agenda for Action: Recommendations for School Mathematics of the 1980's. Reston, VA.: The Council, 1980.