Comprehensive School School Mathematics Program:

Final Evaluation Report

Martin Herbert
Director, Mathematics Research & Evaluation Studies

McREL

April, 1984

This report was prepared by the Mid-continent Regional Educational Laboratory, a not-for-profit educational laboratory. This project reported herein was performed pursuant to a contract from the National Institute of Education, Department of Education. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the National Institute of Education should be inferred.

EXECUTIVE SUMMARY

The Comprehensive School Mathematics Program

CSMP is a grades K-6 mathematics program intended for regular classroom use with students of all ability levels. The program was developed by CEMREL, Inc., an educational laboratory funded by the National Institute of Education.

Distinctive Program Features

Three representational languages - the Minicomputer, string pictures, and arrow diagrams - are used frequently throughout the curriculum, both to convey mathematical ideas and to pose problems. The curriculum is highly structured in a spiral organization with each lesson described in detail in the Teacher's Manual, including a "script" complete with sample question-and-answer dialogue. The lengthy lesson development extends the time the teacher normally spends in whole group instruction. CSMP emphasizes "mathematically rich" situations and builds entire lessons around such situations. There are no behavioral objectives nor are tests built into the curriculum, although student workbooks - 16-page booklets which are assigned once a week - give the teacher one method of evaluating student progress. The CSMP curriculum provides much less time for practicing computational skills and much more time on new content in probability and geometry. It introduces decimals, fractions, negative numbers, and the concept of multiplication earlier than usual, but does not stress computational mastery in these areas.

Special Requirements

From 1 to 5 days of training need to be provided for prospective CSMP teachers. This training can be done by a consultant from the CSMP network of turnkey trainers, or by district personnel, presumably the local CSMP Coordinator, who would also be responsible for monitoring the program, ordering materials, planning implementation, and general trouble shooting. The program costs about as much to begin using as a regular textbook program, but costs more than a textbook to maintain because of its consumable student materials. Because CSMP is not a textbook, it is not likely to be approved in formal textbook adoption procedures.

Program Implementation

CSMP is being used in at least some classes in over 100 school districts and by about 55,000 students. It is used in both urban and rural settings, and as both a gifted and a Chapter I program. Most districts began using CSMP in a few kindergarten and first grade classes and gradually expanded to other schools and other grades, though this expansion seldom reached district wide use through grades K-6, except in small districts.

The most important factor in a successful CSMP implementation is the existence of a skilled and committed CSMP coordinator with district-wide responsibilities. Coordinators report that the two biggest obstacles in implementing CSMP are the training of teachers, especially the change in teaching philosophy required by many teachers, and the lack of computation practice in the program.

CSMP teachers report spending more time in math class than comparable Non-CSMP teachers, and they spend a higher proportion of the time in teacher-led instruction. Teachers supplement the program with computation practice, using about as much time as Non-CSMP teachers do in supplementing their program (usually with "enrichment" activities). This supplementation is most commonly done a few minutes at a time or as homework. Many teachers, particularly in the upper grades, drop lessons from the Geometry and Probability strands in order to complete the schedule.

Program Evaluation

CSMP was evaluated by a special group within CEMREL which operated and was funded independently of the development team. This group produced 50 evaluation reports over the 10 years of the Extended Pilot Tests of CSMP materials. These pilot tests involved 23 districts; subsequent Joint Research

Studies, initiated and supported by local districts, involved 11 other districts. The evaluation effort was monitored by an independent 5-member Evaluation Panel chaired by Dr. Ernest House.

Student Achievement on Standardized Tests

Based on over fifty comparative studies, in over 600 classes, it is clear that CSMP students perform very much the same as Non-CSMP students of comparable ability on a variety of standardized tests in computation, concepts, and applications (or "problem solving"). In computation, there is a slight tendency for CSMP students to do better than Non-CSMP students in grades K-3 and worse in grades 4-6. CSMP students do not do as well in the multi-digit algorithms, like long division, though teacher supplementation in these skills seems to improve performance.

Student Achievement on the MANS Tests

The MANS tests (Mathematics Applied to Novel Situations) are a set of short tests, different in each of grades 2-6, designed by the CEMREL evaluation unit to assess CSMP students' performance in problem solving. The tests were needed because there were no good standardized problem solving tests available. Many of the MANS tests present mathematical situations unfamiliar to both CSMP and Non-CSMP students and none of the tests contain any of the specific CSMP terminology or representational languages. Most items are open-ended and problem oriented. The tests have been used by over 20,000 students during this evaluation.

CSMP classes did better than Non-CSMP classes at every grade level and every ability level. The results were statistically significant throughout, based on Analysis of Covariance of class means, adjusted for reading scores. They were also educationally significant because of the importance of problem solving, the usual difficulty in improving students' problem solving abilities, and the size of the CSMP advantage (typical average percentage correct: 57% versus 50%, or alternatively, effect sizes of 1/3 to 1/2 of a standard deviation in favor of CSMP).

CSMP students were particularly good in process oriented tests, especially solving and using number patterns and relationships, doing mental arithmetic problems (such as ? - 250 = 140), and producing multiple answers to problems. They were also very good in the special topic areas of Algebra and Probability. CSMP students had a more modest advantage in the MANS processes of Estimation, Number Representations, and Word Problems and there was no different between CSMP and Non-CSMP students in the special areas of Geometry, Logic, and Organization of Data.

Other Findings

- Students, who completed CSMP K-6, were rated slightly higher by their seventh grade mathematics teachers than former Non-CSMP students, and received significantly higher mathematics grades, though this advantage decreased with time.
- Students who transferred into the program, provided there were only a few per class, scored slightly below their veteran classmates on the MANS tests but above comparable Non-CSMP students.
- At every grade level, boys outscored girls on all MANS categories except Computation and Elucidation of Multiple Answers. In Estimation and Mental Arithmetic the advantage for boys averaged more than a quarter of a standard deviation (a little less for CSMP), a surprisingly strong result considering the ages of the students.
- In schools where CSMP was started K-4 in the same year, rather than K-1 followed by a new grade each year, second grade classes appeared to gain the full benefit of CSMP after one year while third and fourth grade classes made about half of the normal gain over Non-CSMP performance.

Teacher Reactions

Over half the teachers queried (over 500 in all) gave unqualified approval to CSMP, often describing it in glowing terms. About 10%-15% thoroughly disapproved of the program. The remainder liked the program overall, but had minor or major reservations about some aspect of the program. CSMP was slightly more popular with teachers of grades K-2 than in the upper grades. Teachers considered CSMP's challenging thinking skills and high student motivation as the best aspects of the program. They rated CSMP far better than their previous math program in overall quality, student interest, reasoning ability, and appropriateness for high ability students. On these last two items, their ratings were over a full point higher, on a 5-point scale, than Non-CSMP teachers' ratings.

Many teachers (about half) thought the program was poorer than their previous math program in student achievement in computation and in appropriateness for low ability students. These two common complaints surfaced in many ways, but were not well corroborated by test data which showed only small, easily remediable, computational deficits and very few instances of poor performance by low ability students.

Conclusion

CSMP is a difficult program to implement; it requires more money, a strong coordinator, training and additional preparation by teachers, and a change in teaching philosophy on the part of many teachers. The program does not seem to have much effect on standardized test scores. The program elicits a strong reaction from teachers, mostly favorable. The most important evaluation result is the improvement that CSMP makes in students' ability to deal with various kinds of novel, problem oriented, situations. Most mathematic educators consider this ability to be very important, very hard to bring about and very often ignored in favor of easy-to-measure computational skills.

Table of Contents

I. INTRODUCTION

The Comprehensive School Mathematics Program began in 1966. Lyndon Johnson was president, the Great Society had peaked, and curriculum development in mathematics education was in its glory days with the recent work of projects like SMSG, UICSM, The Madison Project, and SSMCIS. Gradually, CSMP became the last great dinosaur of federal mathematics curriculum development, an anachronism whose momentum (investment) and promise carried it through one challenge after another.

In 1969, one of CSMP's senior advisors remarked to the author, in effect, "Oh, it (CSMP) will never be widely used. This is an experiment to see how good a program we can develop (without regard to implementation considerations)." That remark turned out to be an overstatement.

Amazingly, CSMP did get finished, after many years and millions of dollars, threats to cut off funding and changes in the national setting, such as the financial squeeze in local districts, the return to "basics", and the testing/ accountability movement, all of which hindered the program's dissemination effort. Its completion and quality are a tribute to the determination and talent of Burt Kaufman, CSMP director for most if its existence.

But the remark illustrates one of the central problems in CSMP's development, which was finding a balance between, on the one hand, the philosophy and spirit of what the developers thought mathematics education should be (with "good mathematics" guiding all) and, on the other hand, considerations of practicality and implementation. Neither CEMREL, the educational laboratory that housed and managed CSMP, nor the National Institute of Education, which funded CEMREL and CSMP, understood this problem clearly or formulated a policy to deal with it.

The developers, however, had a clear vision of what they wanted and maintained this vision as they developed materials, often at the cost of reducing the size of the potential market for their materials. The curriculum is also viewed by many mathematics educators as "extreme" - repetitious, idiosyncratic, inefficient, and lacking key elements and varieties of approaches. Thus the program offends some educators right away and for those who like its approach, it presents problems in implementation. Given the resources invested, the amount of time provided for development, and the brilliance of many of the ideas in the curriculum, it's hard not to come to the conclusion that a golden opportunity was missed.

Nevertheless, the program has been used in about 150 school districts and is now being used by about 55,000 students. It is one of the very few viable alternatives to the "national" curriculum exemplified in virtually all available textbooks in elementary school mathematics. The formal evaluation of CSMP is now complete and shows generally positive results, including some hard to achieve student learning gains in certain areas of problem solving. The curriculum has been approved by the Joint Dissemination Review Panel and is eligible for NDN support. Camera ready copies of the final, revised curriculum materials have been completed. A set of extensive training materials is available and a national group of turnkey trainers, the CSMP Network, is in place. Estimates for future sales revenues are slightly higher than for costs of future printing, so that, with the creation of an inventory for the revised grades 4-6 materials, the program can be self sufficient.

The supreme irony of the CSMP experience is that, after all this, after putting millions of dollars and over twelve years into CSMP, the National Institute of Education is now apparently unwilling to spend an additional 2% of that money to create this inventory, a one-time expense which would ensure the future availability of CSMP at precisely the time that a mathematics/problem solving curriculum like CSMP is most needed and likely to be in most demand. Within a year the CSMP curriculum may no longer be available.

This report summarizes CSMP evaluation results. The evaluation effort, like the development effort, has been long running and of wide scope. Testing has been conducted in over 30 districts and 600 classes. Questionnaires have been received from 500 teachers; 250 teachers have been interviewed. The most notable accomplishment has been the development of the MANS Tests (Mathematics Applied to Novel Situations), a series of innovative tests that have been used to compare CSMP and Non-CSMP students in grades 2-6.

Two circumstances helped the evaluation effort immeasurably. First, the evaluation operated and was funded independently of the development group for most of its existence. Without this arrangement, the integrity and quality of evaluation work would have disappeared as would the program itself long before development was completed. Second, the evaluators were lucky to have an extraordinary group of advisors to work with, expecially the five-member Evaluation Panel from 1974-1983 consisting of Len Cahen, Bob Dilworth, Peter Hilton, Ernie House and Stan Smith. They were helpful, talented, diverse in experience, and always prodding, in the nicest way, for the work to be done better. The author wishes to acknowledge the work of Knowles Dougherty, who was part of the evaluation team during most of the Extended Pilot Test and was co-developer of the MANS Tests, and Gail Marshall, another team member who wrote some sections of this report. It was a good group.

The author has been the senior evaluator since 1968. He has fought the usual battles with the developers and the sponsors and has somehow managed to survive to the end. Victors in war get to write the history books; evaluators who survive get to write the final report. In the case of CSMP, both the history and the data are complex and interesting; the author is grateful to NIE and McREL for the chance to finish the job.

II CSMP: DEVELOPMENT AND DESCRIPTION

Philosophy and Goals

Like SMSG, UICSM, the Madison Project and other mathematics reform projects that preceded it, CSMP was designed to teach students mathematics and not merely arithmetic. One of the key aspects of CSMP has been its dual emphasis on both mathematical content and pedagogy designed to support mathematical reasoning. As the program was developed, piloted and revised, both content and pedagogy were modified to reflect classroom experiences.

One of the basic tenets which CSMP developers have often stated is that elementary school mathematics should not unduly stress drill and practice in arithmetic computation but should introduce children to what the developers term "mathematically important ideas".

To present those "mathematically important ideas" to students, three basic principles guided the developers. These principles, which differ from those on which "traditional" text book mathematics programs are based, are the following:

> Mathematics is a unified body of knowledge and should be organized and taught as such, so that, for example, the artificial separation of arithmetic, algebra and geometry should not be maintained.
>
> Mathematics as a body of knowledge requires certain ways of thinking and cannot be done by the exclusive use of memory.
>
> Mathematics is best learned by students when applications are presented which are appropriate to students' levels of understanding and to their natural interests.

CSMP's point of view is also illustrated in the following description of the curriculum, excerpted from materials prepared by the developer for promotional purposes:

> "An underlying assumption of the CSMP curriculum is that children can learn and can enjoy learning much more math than they do now. Unlike most modern programs, the content is presented not as an artificial structure external to the experience of children, but rather as an extension of experiences children have encountered in their development, both at the real-life and fantasy levels. Using a "pedagogy of situations", children are led through sequences of problem-solving experiences presented in game-like and story settings. It is CSMP's strong conviction that mathematics is a unified whole and should be learned as such. Consequently, the content is completely sequenced in spiral form so that each student is brought into contact with each area of content continuously throughout the program while building interlocking experiences of increasing sophistication as the situations become more challenging.

"A feature unique to CSMP is the use of three nonverbal languages that give children immediate access to mathematical ideas and methods necessary not only for solving problems, but also for continually expanding their understanding of the mathematical concepts themselves. Through these languages the curriculum acts as a vehicle that engages children immediately and naturally with the content of mathematics and its applications without cumbersome linguistic prerequisites. These languages include: the Language of Strings (brightly colored strings and dots that deal with the fundamentally useful and important mathematical notion of sets); the Language of Arrows (colored arrows between pairs of dots that stimulate thinking about relations between objects); and the Language of the Papy Minicomputer. The Minicomputer, a simple abacus that models the positional structure of the numeration system, is used both as a computing device and as motivation for mental arithmetic. Its language can be used to represent the nature and properties of numbers. CSMP is flexible enough to facilitate whole-group, small-group, and personalized instruction, and is appropriate for all children from the "gifted" to the "slow learners". It recognizes the importance of affective as well as cognitive concerns and has been developed and extensively tested in classrooms nationally. Thus, unlike many approaches to mathematics which believe that students need to have mastered their own language before they can handle logical mathematical tasks, CSMP uses these precise, pictorial modes rather than relying exclusively on verbal instruction to express the abstract concepts embodied in CSMP content."

The Development of CSMP

Brief History of CSMP

Comprehensive School Mathematics Program stands for both the name of a curriculum, CSMP, whose evaluation is the subject of this report, and the name of a project which was responsible for developing curriculum materials. Two major curricula were developed under CSMP project auspices: CSMP, a K-6 mathematics program for regular classroom instruction, and the Elements of Mathematics (EM) program, a grades 7-12 mathematics program for gifted students. EM treats traditional topics rigorously and in depth. It includes much of the content generally required for an undergraduate mathematics major. These two curricula are unrelated to one another but certain members of the CSMP staff contributed to the development of both projects.

The CSMP Project was established in 1966, under the direction of Burt Kaufman, who remained director until 1979. It was originally affiliated with Southern Illinois University, Carbondale, Illinois. It was originally affiliated with Southern Illinois University, Carbondale, Illinois. After a year of planning, CSMP was incorporated into the Central Midwest Regional Educational Laboratory (later CEMREL, Inc.), one of the national educational laboratories funded at that time by the U.S. Office of Education.

By 1968, CSMP had a staff of about 15 teacher-writers, artists, and evaluators, as well as a large and active group of consultants. Also involved with CSMP's development was a program advisory committee, chaired originally by Robert Davis and later by Peter Hilton and Gerald Rising, and a CEMREL-wide advisory committee for evaluation, chaired by Dr. Michael Scriven.

During the initial development work, third grade lessons were written. In this early development phase, the emphasis was on "activity packages" in which several class-length activities on a single topic were grouped together to form a single "package". Most of the students' work was on an individual basis, and a management system was devised which included pre- and post-testing and remediation strategies. Under this system, students occasionally worked in pairs, and many activities were accompanied by audio-tapes which helped students work through the exercises. Most of the teacher's time was spent working individually with students while a teacher aide handled the management details.

In 1970-71, an experimental comparison was made using third grade classes in the Carbondale, Illinois public schools. This program, as it was developed to that time, was used in one group of classes, and a "stripped down", less expensive version of the same content was taught in a traditional way in another group of classrooms. The comparison showed there were virtually no differences in achievement between the two groups of classes.

At about the same time, CSMP staff also became aware of the work of Frederique Papy in Belgium. The staff began to develop kindergarten and first grade activities based on her work, which used arrow diagrams, string pictures and the Papy minicomputer to convey mathematical ideas. These circumstances eventually led to the decision by the development staff to abandon the individualized approach used in the third grade materials in favor of the pedagogical and substantive innovations of Dr. Papy thus placing the teacher in the more traditional role in group instruction. Meanwhile, development continued from second grade, a year at a time.

In 1972, the Office of Education conducted a review of all lab and center development programs in anticipation of their transfer to the newly created National Institute of Education. The review recommended a phasing out of CSMP. However, a subsequent site visit by a three-person review team led to a recommendation that the almost-completed K-2 materials be given a pilot trial. They also recommended that development work be restricted to planning activities, pending the results of the pilot trial. Early in 1973, a contract through 1975 was signed with NIE to conduct pilot trials and to complete curriculum development through third grade. Then, according to that contract, a decision about funding for further development work would be made. Thus began the extended pilot studies (1973-74), conducted by the evaluation staff directed by Martin Herbert, and monitored by an evaluation panel chaired by Dr. Ernest House, University of Illinois. Subsequently, an external review by a three-person team, chaired by Dr. Gail Young, recommended in 1975 that NIE continue funding for development of lessons for grades 4-6.

In the meantime, the development activities came more and more under the direction of Dr. Papy, who had joined the staff as Director of Development, and the curriculum gradually took on its present form and philosophy. In 1975, after the voluntary departure of most of the development staff, and faced with a strained relationship with the Carbondale schools, the project moved to St. Louis and was housed in a single facility with other CEMREL programs.

In the fall of 1975, the development staff was rebuilt and developmental work began in two fourth grade classes in the University City Public Schools, a racially integrated school district of inner-suburban St. Louis. In 1977 pilot testing of the fourth grade curriculum in regular classes was undertaken by the evaluation staff as part of the sequence of Extended Pilot Tests.

In 1979, Clare Heidema became director of CSMP and supervised the completion of development as well as the final revision of materials. This final revision occurred at each grade level in the year following the Extended Pilot Tests. The testing of sixth graders was completed in fall, 1982, and the final revised versions of all materials will be completed in early 1984.

Development Cycle

By 1973 a four-stage process of materials development had been established and this procedure was followed in subsequent years.

1. Writing and Teaching Lessons (1 year). The CSMP staff, led by Dr. Papy, generated short sequences of lessons around a topic and then taught the lessons to two or three classes. The overriding criterion for selection/development of lessons was always whether the lesson themes or "situations" were mathematically rich, i.e. could easily lead in several ways to important mathematical ideas or ways of thinking. Also guiding development was the need to maintain a grade-by-grade correspondence with the arithmetic skill development that is so well established in American schools. Several observers watched the lesson being taught, occasionally worked with individual students, and contributed to decisions on lesson revisions. These observations also affected future decisions about what to teach and how to do it. Overall, the classes were of average ability though they contained a higher proportion of both high and low ability students than most classes do.

This first stage of development distinguished CSMP's development process for two related reasons. First, there was no overall master plan describing what content would be taught at which grade levels. The content was gradually filled in as time progressed. Second, the daily teaching of lessons allowed for rapid, even overnight, changes based on student reactions to the lessons. Thus, development was at the same time fluid and empirically based.

2. Local Pilot Test (1 year). This stage was carried out by the development group and was more a further stage of development than a pilot trial. The previous year's lessons were revised and organized into a year long sequence and taught in 6-8 classes in the St. Louis area. Regular classroom teachers taught the lessons and were observed and assisted by CSMP staff at least twice a week.

The process was still informal and fluid during the local pilot test. Occasionally CSMP staff would write lessons after observing a class period and then bring those lessons with them on their next visit. CSMP staff often conducted lessons themselves. There was no provision for end-of-year testing of student achievement, though student performance on workbook assignments was systematically reviewed.

3. Extended Pilot Test (2 years). The first two stages resulted in a set of materials, the Final Experimental Version, that included both student materials and detailed Teacher's Guides for that grade level. In the first year of the Extended Pilot Test, about 10-12 classes in the St. Louis area used the curriculum in a more or less "hands-off" manner. Students in these classes had used CSMP in earlier grades, but teachers were usually inexperienced with CSMP and were trained in summer workshops lasting one or two weeks and conducted by CSMP staff. Classes were observed by both development and evaluation staff, test instruments were developed, and experimental comparisons were made between CSMP classes and similar non-CSMP classes in the same district. Evaluation related activities were the responsibility of a special unit within CEMREL which was independent of the development group. All expenses for materials and training were paid for by CSMP. Thus the first year of the Extended Pilot Test provided a vehicle for trying out the materials in a small controlled experimental trial, and for developing training and evaluation procedures for use the following year.

In the second year of the Extended Pilot Test, a much wider test of the materials was conducted in school districts nationally as well as in the St. Louis area. No conditions were placed on the number or location of pilot classes and participating school districts were free to choose teachers and classes in ways consistent with their own pilot needs. However, participating districts were required to provide evaluation data as required by the evaluation staff (questionnaires, access to classes, student testing) and to cooperate in providing appropriate control classes for the comparison of student achievement. This comparison was accomplished mainly through the use of the MANS Tests, a special series of tests developed by the evaluation staff.

Local districts were also required to pay the cost of instructional materials and to provide a coordinator responsible for several tasks: overseeing the implementation of the program, acting as a liaison between CSMP and the district, attending a summer training workshop conducted by CSMP, and subsequently training teachers as needed in their districts.

Approximately 30-50 CSMP classes used the materials in this second year of the Extended Pilot Test. The comparative studies of student achievement involved about 60 classes altogether, and formed the main source of data for the summative evaluation of the program.

4. Final Revisions. Based on various evaluation data, including classroom observations, teacher reactions and student achievement, the Final Experimental version was revised and a Final edition prepared for nationwide availability. During this stage, extraneous lessons were eliminated, lessons were shortened or lengthened to reflect time limits at typical sites, and Teacher's Guides were revised to incorporate teachers' suggestions or to clarify lessons.

The years in which these stages were completed are shown below for each grade in Table 1.

TABLE 1
Completion of Development Stages, By Grade

	Pre 1973	73-74	74-75	75-76	76-77	77-78	78-79	79-80	80-81	81-82	82-83
Development:	K-2	3		4	4,5	5,6	6				
Local Pilot:	K-2		3		4	4	5	6			
Extended Pilot:		K,1	K,1,2	2,3	3	4	4,5	5	6	6	
Final Revision:				K		1,2	3		4	4,5	5

Aside from student and teacher materials, CSMP has also documented its goals and procedures in reports, articles, and program materials such as Coordinator's manuals, workshop manuals, and preview packets. A list of current documents is given in Appendix D.

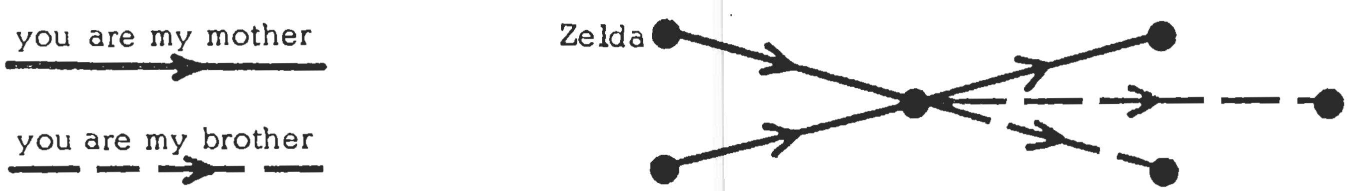
CSMP Representational Languages

The next section describes the CSMP representational languages, with several examples of their usage. The reader may wish to skip ahead to the next section, CSMP Content, page 19.

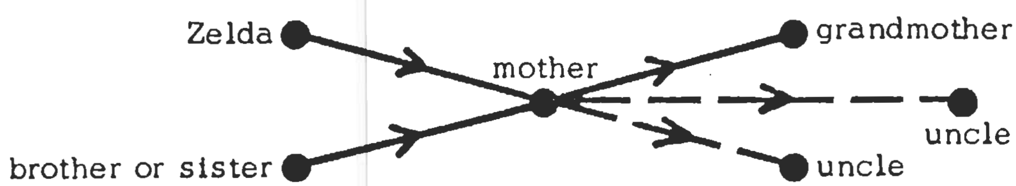
Three pictorial devices - Arrow Diagrams, String Pictures, and the Minicomputer -are used extensively throughout the curriculum as vehicles for presenting and working with a wide variety of content and mathematical processes. CSMP calls these devices "languages". The examples used below to illustrate these languages are taken from the Teacher's Guides and student workbooks. If the reader will take the time to think about these examples, a good deal of the CSMP philosophy may become accessible.

The Language of Arrows.

CSMP uses colored arrows to represent relations among numbers or objects represented by dots. In the examples of arrow diagrams in this section, it has been necessary to use solid and dotted arrows for different relationships. CSMP uses color to distinguish arrows, a much more effective and striking visual device. An example showing non-numerical relationships is given below, where each dot represents a different person.

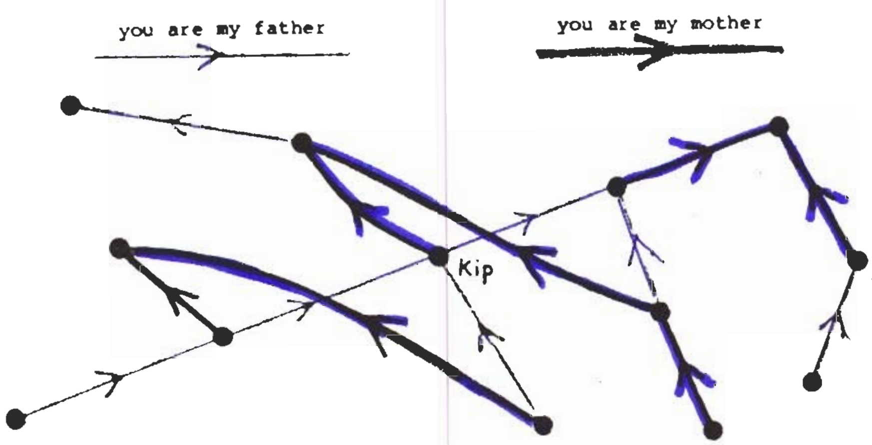


Labelling all dots to show their relationship to Zelda, gives the following:



Note the indeterminacy of the lower dot and fact that some arrows are missing, for example from the uncles to the grandmother.

In a more complicated example, shown below, every dot can be labelled "Kip's____.

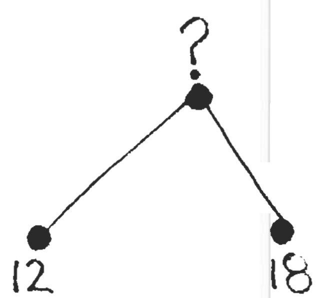


Arrows are most often used to represent numerical relationships. The picture below represents 2 + 3 = 5

Note the "key" in the previous diagram to show what the arrow stands for. The diagram below represents the equation 3A + 5 = 14.

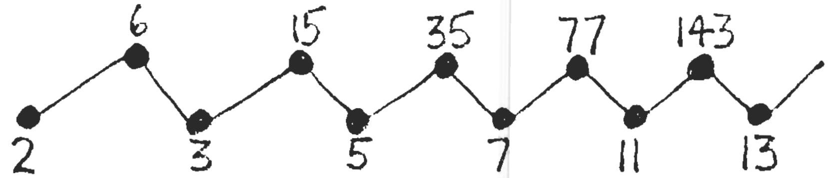
If an arrow can be drawn in either direction between two dots, then the dots can be connected by a chord, as illustrated in the following highlights from a 35-page Story-Workbook for third graders.

> The principal of a certain school who was the number 0 and the vice principal, the number 1, made up the following rule in their school to reduce the amount of talking: "Two numbers will be allowed to talk to one another only if one of them is a multiple of the other." Very gradually some interesting things are developed in the story book:
>
> > 0 and 1 are the only ones who can talk to <u>everybody</u>.
> >
> > each number can talk to itself.
> >
> > Two friends, 12 and 18, can't talk to one another but one of them has the following idea:

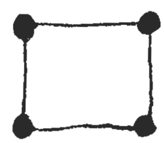


> > Some numbers - common multiples of 12 and 18 - can be intermediaries for these two friends.
> >
> > The same thing happens to three good friends; 4, 10 and 15.
> >
> > Some numbers, for example 24, can talk to several friends who are smaller than they are. Others (prime numbers) can't talk to any smaller friends.

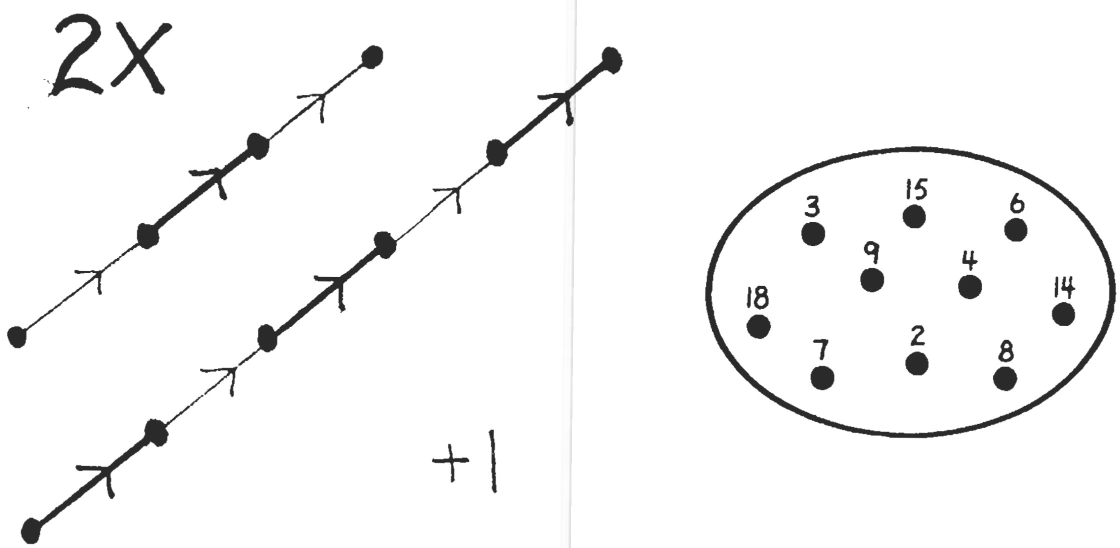
An interesting parade took place:



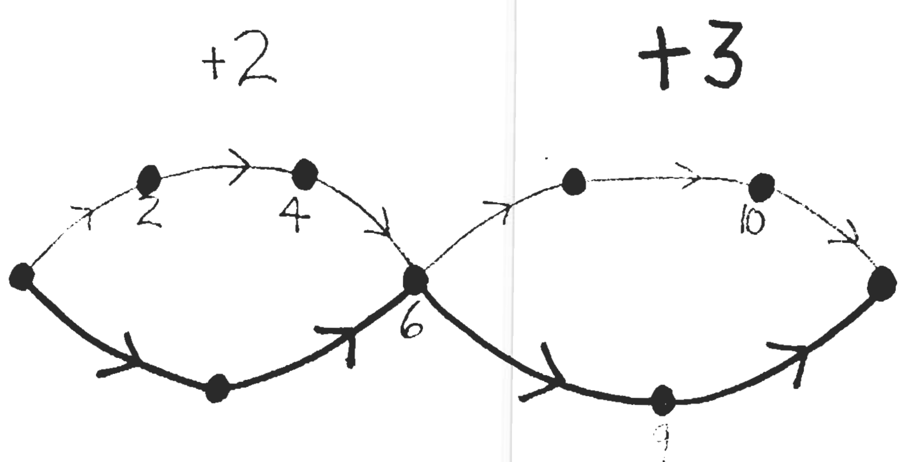
Find four numbers who communicate as follows:



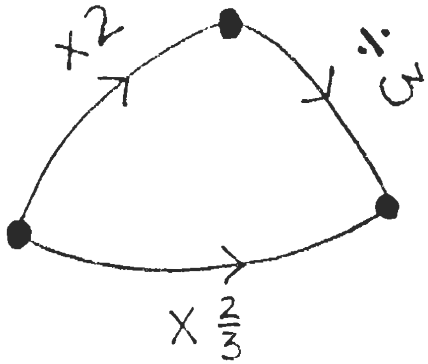
In the problem below, third grade students have to label the arrow diagram with exactly the dots shown in the string.



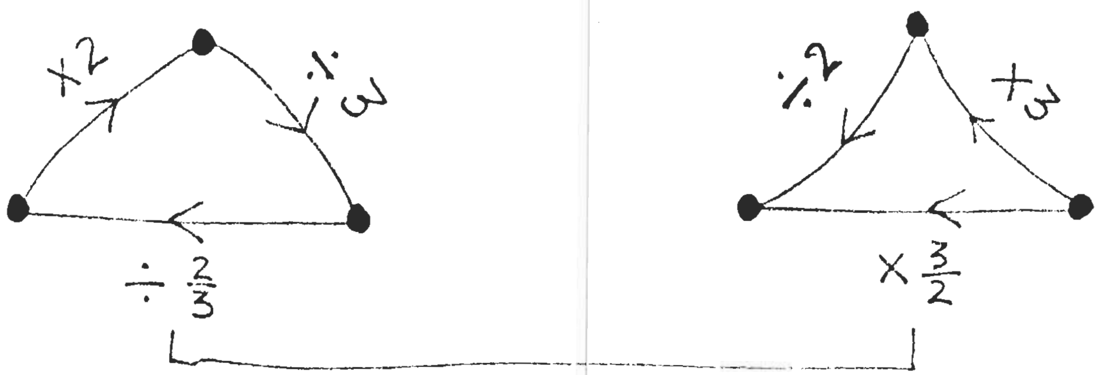
Common multiples of 2 and 3 appear naturally in the partially labelled diagram below.



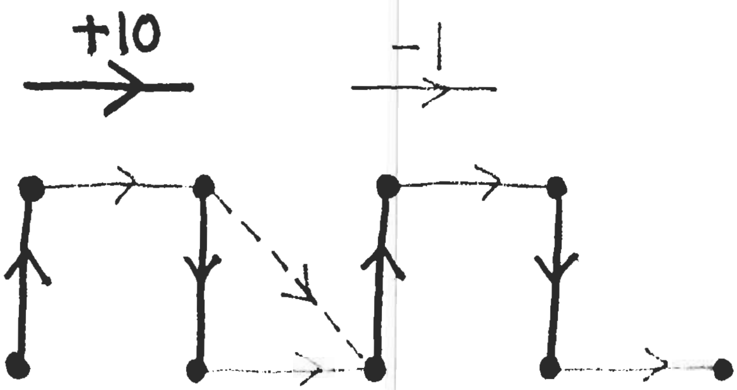
Return arrows are used frequently. One effective use is in showing the relationship between multiplication and division of fractions, introduced in fifth grade. Multiplying by 2/3 is split into two steps as shown in the diagram below.



Return arrows express multiplication and division as inverse operations. Hence, the dotted return arrow must represent ÷ 2/3 (bottom left). Alternatively, return arrows for the upper arrows could be drawn first (bottom right), in which case the dotted arrow represents the composition of a X 3 arrow and a ÷ 2 arrow, which is X 3/2. So ÷ 2/3 means the same as X 3/2.

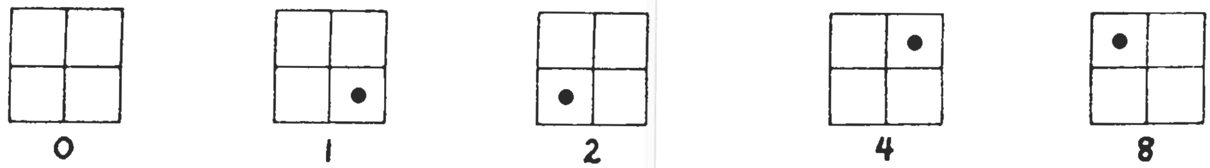


In the diagram below, it is possible to determine which dots represent the largest and smallest numbers and what the dotted arrow stands for, without actually labelling any of the dots.

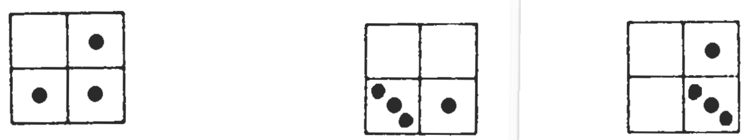


The Language of the Minicomputer.

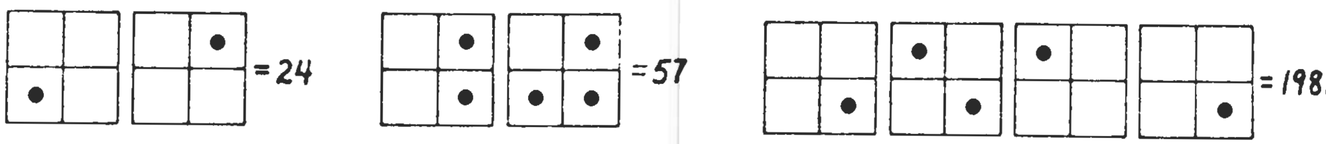
Minicomputer lessons use one or more square boards, each divided into four squares and colored according to the Cuisenaire values so that checkers placed on it assume the values 1, 2, 4 or 8.



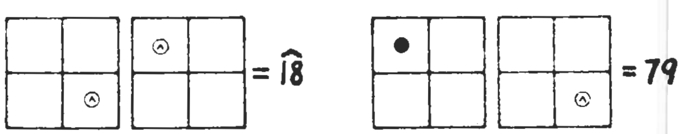
Numbers can be shown in several ways. The number 7 is represented below in three different ways but the standard representation is the one on the lower left, where there is no more than one checker per square.



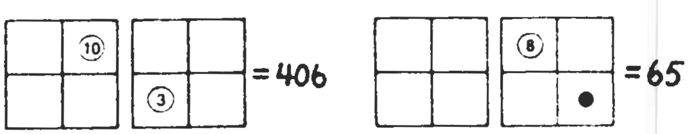
Several boards placed side-by-side correspond to the 1's, 10's, 100's, etc. places of the normal positional notation.



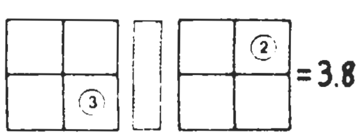
Negative numbers can be shown with special negative checkers used alone or in combination with regular checkers.



Weighted checkers have numerals written on them to represent that many checkers.

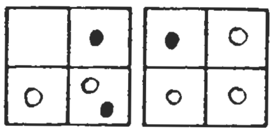
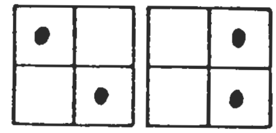


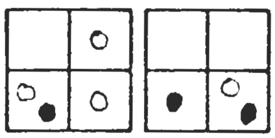
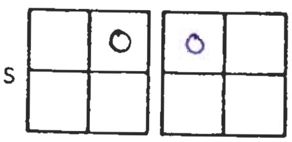
A green bar placed between two boards represents a decimal point.

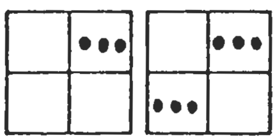
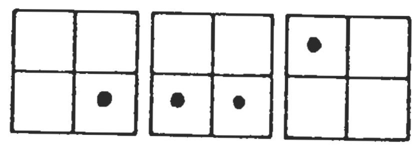


"Plays" are made by replacing two checkers on a square (e.g. two 4's) with a single checker on the next highest square (one 8) or vice versa. A special play, replaces a checker on the eight square and a checker on the two square with a checker on the one square of the next board: this special play, and its reverse, allow plays from board-to-board. The actual boards may be large demonstration boards used for teacher-led instruction, smaller paper "boards" for individual student use, or pictures of boards in workbooks allowing students to give paper and pencil responses.

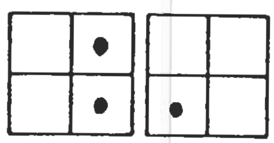
The Minicomputer can be used to calculate with each of the four standard numerical operations.

37 + 58 =  making plays simplifies to  = 95

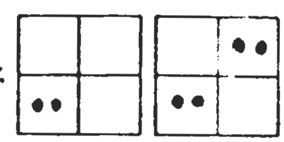
71 - 23 =  cancelling solid checkers leaves  = 48

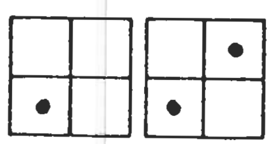
3 x 46 =  making plays simplifies  = 138

52 ÷ 2

First, 52 = 

Then, making plays to get two checkers per square: 52

= 

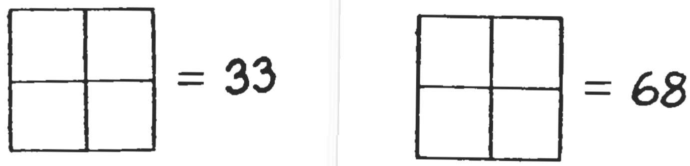
So 52 ÷ 2 =  = 26

While the Minicomputer has an obvious value for representing and calculating with numbers, it is also used as a device to stimulate mental arithmetic and to pose problems, particularly in the upper grades.

The following examples show a few of the kinds of problems that can be posed.

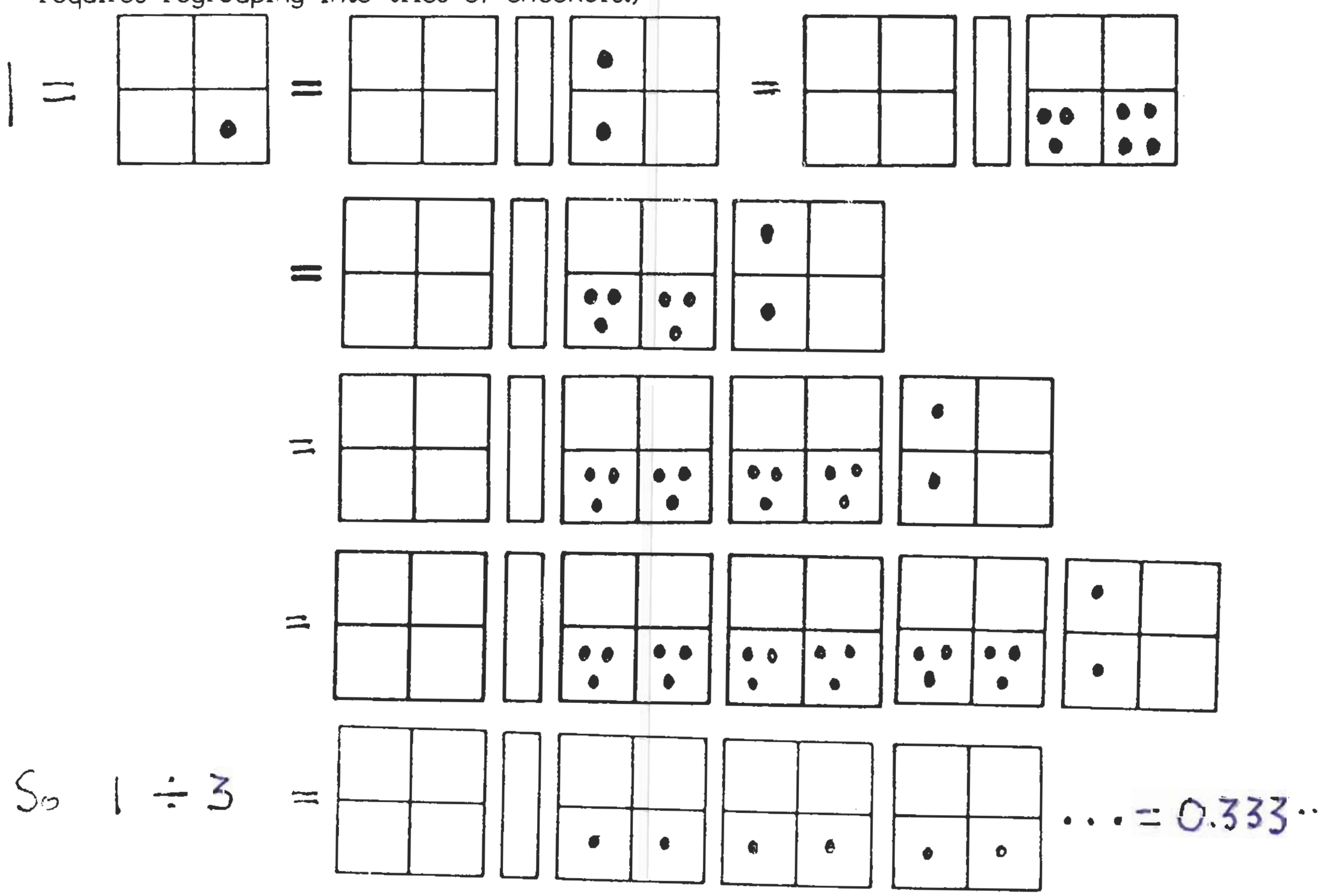
List the numbers that can be shown on the one's board using exactly 3 regular checkers.

put each number on the Minicomputer using a 5 checker and exactly one of these checkers: ② ③ ④ ⑥ ⑦ ⑧ ⑨

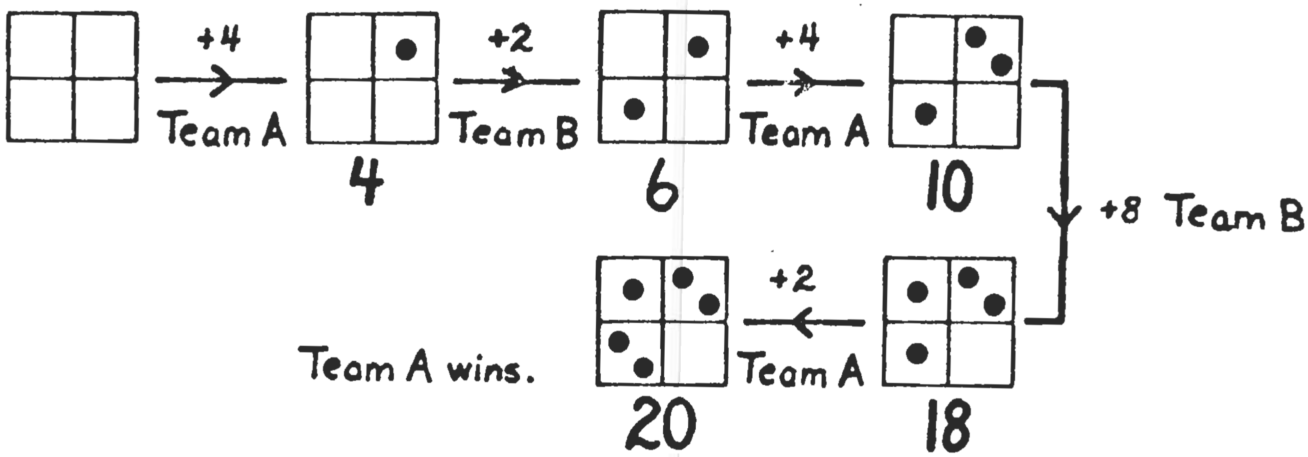


Use any two weighted checkers to show 26.

Infinite repeating decimals can be illustrated by the following sequence for 1/3. (Students will already have learned that division by 3 on the Minicomputer requires regrouping into trios of checkers.)

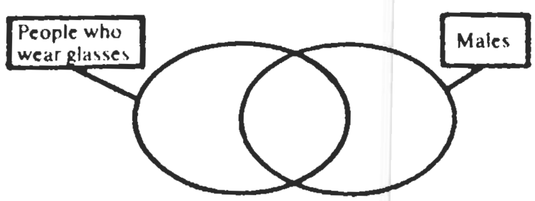


An example of a strategy game starts with the Minicomputer is shown below. The game starts at zero and uses a single board. Teams take turns adding a single checker to the board. The first team to reach 20 (without going over) wins. Below is a hypothetical sequence of plays in a game won by Team A.

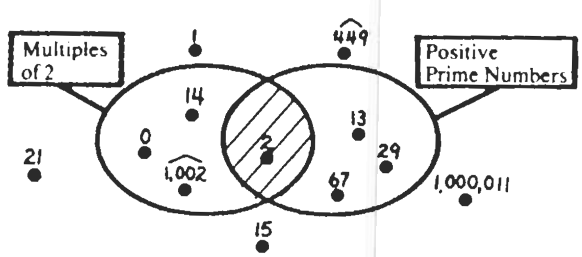


The Language of Strings

Strings are used to show the classification of objects according to certain attributes. Young students might be asked to put dots for themselves in one of the four regions of the following string picture. Note that a girl without glasses would be represented by a dot outside both strings.

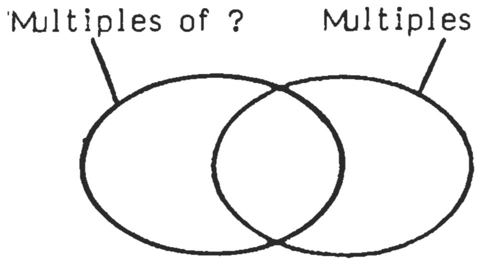


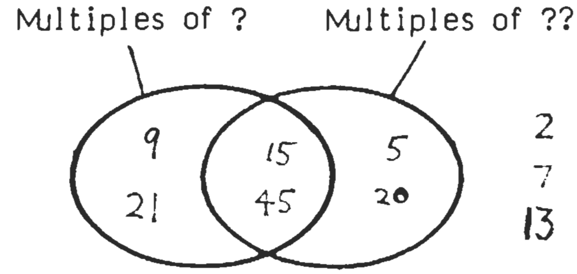
Strings are most often used in CSMP to classify numbers. The following string picture shows that 2 is the only positive prime which is a multiple of 2, duly noted by the cross hatching of the intersection to indicate that all elements of that region are shown.



Starting with the diagram, lower left, the teacher might proceed a follows:

Ask the class for numbers, and place and label the numbers in the diagram.
Give numbers to students to place and label.
Ask the class for numbers which belong in a particular region.
Ask the class to name the intersecting region of the diagram, which by now might look like the one lower right.

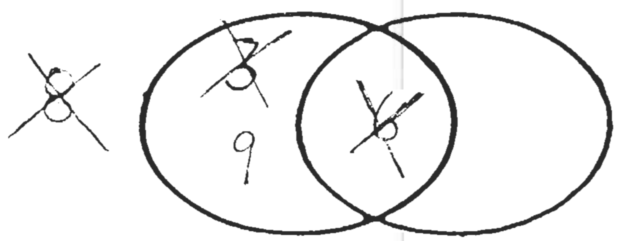




Often pieces of information are given one at a time, allowing inferences to be made with each additional piece of information. In the diagram at the bottom of the page, the problem is to try to figure out the labels for each string. The labels to be chosen from the following list:

multiples of 4,
odd numbers,
smaller than 10,
positive divisors of 12,
positive divisors of 18 or
positive divisors of 24.

After a few numbers have been tested, the information now available might be as represented in the diagram. (A crossed out number means that number does not belong in that region.) As it happens, all but one of these possibilities for each string can now be eliminated.



Hand calculators.

Although not really a CSMP language, hand calculators are used in many lessons for investigating numerical properties and patterns and in various games which require strategic planning.

An example of the use of hand calculators in a problem solving contest is to assume that some of the keys on a hand calculator are broken leaving only the following keys:

$\boxed{5}$, $\boxed{6}$, $\boxed{8}$, $\boxed{9}$, $\boxed{+}$, $\boxed{-}$, $\boxed{\times}$, and $\boxed{=}$

Try to display the number 54. Two relatively easy sequences of keys that produce 54 are 6 x 9 = and 5 9 - 5 = .

The number 540, however, is much more difficult to get on the display. There are many solutions, some requiring many fewer buttons to be pressed than others.

CSMP Content and Curriculum Organization

The previous examples illustrate the difficulty of separating CSMP content from pedagogy. Probably the easiest way to describe the content of CSMP is to show how it differs from what is usually taught in the traditional K-6 program. Listed below are topics in the CSMP curriculum that are not typically found in most programs. It is important to note that describing these topics as unique content of CSMP does not mean that, as a result, students would ordinarily learn (know) a body of content in the usual sense. Each of the topics listed below is to be taught in one, or a few, teacher-led lessons in which a situation is developed through gradual extensions and problems. There is no body of facts or theorems to learn, students are not specially tested on the topics, and mastery of concepts is not usually required for the next topic. (Pedagogical considerations are described in more detail later in this chapter.)

Geometry. Construction of figures with translators, angle templates, compass and straightedge; properties of shapes independent of distance; parallelism and parallel projections; reflections and symmetry; generalized distances other than Euclidian (for example, one-digit distance, map distance, taxi distance); tesselations; the "map" of a cube; the triangle inequality.

Probability. Predicting and comparing results of probabilistic experiments simulated by marbles, spinners and dice; probability concepts such as randomness, equally likely events, fairness, selection with and without replacement; combinatoric analysis of probabilities; the multiplication principle in multi-stage events; geometric solutions to multi-stage random experiments.

Numbers and Number Theory. Prime factoring, modular arithmetic; various abaci and positional notations (binary, base 3, 4, 8, 2); codes and decoding in combinatorics; representation of fractions by infinite series; introduction to approximation and relative magnitudes; relations, functions, operations as functions, converses and compositions; negative numbers.

Logic. Negation of attributes; terminology (every, at least, at most, exactly); strategic thinking in special games.

Conversely there are a number of areas in the traditional curriculum that the CSMP curriculum does not cover (or emphasizes less).

In the early grades there are virtually no lessons dealing with telling time, calendars, common English measures, and coins/money. The Teacher's Guides inform teachers of this and ask them to teach those topics in their own way at the appropriate times.

Although CSMP students spend considerable time working with string pictures, the associated set terminology which appears in some fifth and sixth grade textbooks is not used (e.g. set, intersection, union, brackets, etc.). There is little usage of certain terminology in geometry, such as isoscelese, equilateral, circumference, and pi.

The curriculum calls for very little emphasis on cancelling with fractions, and on multiplying or dividing of mixed numbers.

The division algorithm (e.g., a 2-digit number divided into a 4-digit number) is not developed as fully as is traditionally done in elementary schools.

There are very few word problems of the kind typically found in text books and standardized tests.

Several topics are introduced at an earlier level in CSMP, for example, fractions, especially taking one-nth of a number, and partitioning a set of discrete objects into equally numerous subsets, decimals, and the process of multiplication. CSMP students learn about the concept of multiplication in first grade and are exposed to several representations of basic facts and how to calculate them. There are also many instances in first grade of multiplying a larger number by 2 or 3, such as 2 x 37.

At the same time, numerical skill development proceeds more slowly, so that, for example, the subtraction, multiplication and division algorithms are not practiced as early or as often as in traditional programs. The subtraction algorithm is developed later and in a different way then is traditionally done. Rote memorization of multiplication facts is not emphasized and the multiplication algorithm of 2-digit by 3-digit numbers does not get introduced, let alone mastered, until fifth grade. Very little time is set aside for developing skill in the division algorithm. Though fractions are introduced early, the curriculum devotes less time to adding and subtracting mixed numbers and common fractions, especially those with unlike denominators.

Grade Level Organization. The curriculum is divided into four levels:

Kindergarten.

Grade One, Parts I and II for first and second semester respectively.

Upper Primary Grades, Parts I and II (second grade) and III and IV (third grade).

Intermediate Grades, Parts I to VI for the six semesters in grades 4-6.

Content Organization. In kindergarten and first grade the content is organized and presented as a single sequence of lessons emphasizing elementary arithmetic concepts and their exemplification in the CSMP languages. In the other grades, content is organized by four strands:

The World of Numbers

The Languages of Strings and Arrows

Geometry and Measurement

Probability and Statistics

The Probability and Statistics strand begins in fourth grade. The Strings and Arrows Strand is concerned with logical thinking and reasoning skills though it also contains a good deal of number work, either directly or as required during the course of lessons primarily concerned with other objectives, such as strategic thinking.

Sample Sequence of Lessons. Within each strand there are blocks of lessons dealing with the same idea, which is developed further with each lesson. An example of an unusually long block of lessons with arithmetic development is given below (most blocks are 2 or 3 lessons long). The lessons are from the sixth grade geometry strand.

The sequence begins with two lessons about circles. Students collectively draw many circles, all passing through a fixed point, but whose center is always on another given circle. (First the fixed point is outside the given circle, then on it, then inside it, each time producing striking results.) Various questions are asked about smallest circles and about the effect of moving the fixed point slightly, or even all the way to the center of the given circle. The second lesson concludes with construction of lattice by successively drawing new circles whose centers are previous points of intersection and then joining these points of intersection.

Then there is a sequence of nine lessons in which the teacher helps students to:

draw perpendicular lines using a paper square as a corner,

draw closed shapes containing only right angles,

construct perpendicular lines with compass and straightedge, first from points on a line, then from points outside the line,

construct equilateral 8-sided polygons with compass and straightedge,

draw all the possible quadrilaterals each of whose sides must equal one or the other of two given lengths,

do the same thing for triangles,

determine where a stick of fixed length could be broken twice and so that a triangle could be formed from the pieces.

three lessons on estimating the probability of two random breaks of a stick producing pieces that could form a triangle.

Included in the strands are sets of lessons, on various probability, reasoning, and number games; the games have rules which can be changed to make the game more challenging or to feature some new mathematical idea.

Schedule Organization. The schedule is organized in a spiral fashion by days of the week. On two days of the week, e.g., every Monday and Thursday, lessons come from the World of Numbers Strand. One day of the week is devoted to workbooks which provide practice and problems from recent lessons in all strands. The other two days of the week are devoted to the other strands: in grades 2 and 3 a day each for the Geometry and Measurement strand and the Strings and Arrows strand and in grades 4-6 roughly equally divided among these two strands plus the Probability and Statistics strand.

A suggested sequence of lessons for the last nine weeks of second semester, third grade, is shown on the next page as an example of schedule organization. Each column is for a day of the week.

10	N19 Decimals on the Number Line N p. 151	L10 I AM NOT MY NAME L p. 78	N20 Return Arrows #2 N p. 164	W10 GALAXY OF PROBLEMS #3 (Lesson 2) W p. 75	G7 Touring Houses #1 G p. 64
11	N21 Subtraction Algorithm #2 N p. 170	ADJUSTMENT DAY	N22 Composition of Functions #2 N p. 176	W11 Spinner Game Detective Story W p. 85	G8 Length of Spirals G p. 71
12	N23 Comparing Prices #1 N p. 185	L11 Multiplication Modulo 10 #1 L p. 84	N24 Composition of Functions #3 N p. 198	W12 GALAXY OF PROBLEMS #4 (Lesson 1) W p. 96	G9 Touring Houses #2 G p. 81
13	N25 Subtraction Algorithm #3 N p. 206	L12 Talkative Numbers #1 L p. 94	N26 Roads #3 N p. 211	W13 GALAXY OF PROBLEMS #4 (Lesson 2) W p. 99	ADJUSTMENT DAY
14	N27 Mystery Arrows N p. 217	L13 Who is Max? L p. 102	N28 Division Problems #3 N p. 225	W14 Who are Zin and Zan? W p. 109	G10 Rectangles of a Given Area G p. 91
15	N29 Multiplication Problems N p. 232	L14 Three Boxes #2 L p. 112	N30 Division Problems #4 N p. 241	ADJUSTMENT DAY	G11 Rectangles of a Given Perimeter G p. 102
16	N31 Hand-Calculator Games #2 N p. 249	L15 The String Game #3 L p. 122	N32 Division Problems #5 N p. 256	W15 FISHING FOR NUMBERS (Lesson 1) W p. 116	G12 Touring Houses #3 G p. 113
17	N33 Comparing Prices #2 N p. 261	L16 Talkative Numbers #2 L p. 129	N34 Roads with Cords N p. 271	W16 FISHING FOR NUMBERS (Lesson 2) W p. 119	ADJUSTMENT DAY
18	N35 Minicomputer Golf #2 N p. 278	L17 Multiplication Modulo 10 #2 L p. 137	ADJUSTMENT DAY	W17 A Short Story about Three Friends W p. 133	G13 Pencil Tracings G p. 122

Fig. 1. Sample 9-week schedule of lessons, third grade

During any given week, students will encounter a wide variety of activities. In reading down any column, it can be seen that many topics appear more than once (and will reappear many times over the course of the full 36-week sequence for the whole year). This is especially true of lessons whose names are followed by "#". Each lesson reviews the previous one and takes the idea a step further or into a new direction.

CSMP believes that different children learn at different times and at different rates and since learning is not necessarily a linear process, this spiral organization gives each student a new chance to work with an idea at each turn of the spiral. Thus, according to developers, when students return to a topic a week or two later, some who did not understand the concepts the first time around may now be better prepared to work on the ideas.

Each grade level has its own prescribed schedule of lessons which is presented in the Teacher's Guide. The Kindergarten schedule has a linear sequence of 108 lessons spread over 9 months (i.e. about 3 per week). At other grade levels, lessons are grouped on a days-of-the-week basis, like the portion of the third grade schedule shown previously, and range from 150 to 180 lessons depending on grade level. Included in the schedule are "Adjustment Days" to allow for holidays, snow days, etc.

One other feature of CSMP's curriculum sets it apart from most other curricula. Many other programs have built-in testing programs (for example, tests in the teacher's edition) which specify behavioral objectives for each unit. CSMP has no behavioral objectives and no program devised tests. Instead, teachers are encouraged to gauge students' progress and assign grades to students on the basis of classroom responses and performance on the weekly workbook assignments.

Materials

The Teacher's Guide and Individual Lessons.

The Teacher's Guide contains a multi-page description of each lesson. Each lesson description has several parts: an overview which describes the lesson's purpose, a capsule summary of each part of the lesson, a "script" for each lesson which includes accompanying diagrams and examples, and assignments for additional student work. A lesson from the third grade Teacher's Guide is shown below. The lesson can be found in the schedule shown on the previous page.

N15 Minicomputer Golf #1

CAPSULE LESSON SUMMARY

Explore the effect of moving various checkers in a configuration on the Minicomputer — after a move, is there a larger, a smaller, or the same number on the Minicomputer? Introduce a version of the game "Minicomputer Golf", in which checkers are moved from a starting configuration in order to reach a specific goal.

MATERIALS

Teacher: Minicomputer set; colored chalk

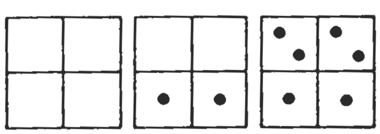
Student: None

DESCRIPTION OF LESSON

Exercise 1

T: I am going to put a number on the Minicomputer. See if you can figure out which number it is.

Gradually put this configuration on the Minicomputer, starting with the checkers on the squares of largest value. Pause frequently so your students can do the mental calculations.



Let the students whisper the number to you before letting one of them answer aloud.

S: 57.

Invite several students to explain how they knew this number was 57.

Write these words on the board.

More

Same

Less

T: I am going to move one of these checkers to another square. Tell me if the new number is more than, less than, or the same as the number on the Minicomputer now.

Move a checker from the 2-square to the 1-square.

Point to each of the words on the board in turn and ask the students to hold up their hands when you point to the word which describes the new number. The students should indicate that the new number is less than before.

T: How much less . . . ?

S: 1 less.

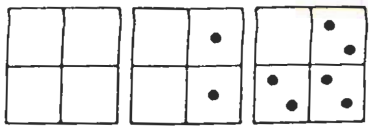
Repeat this activity several times as suggested below. Do not return checkers to their original positions. Each move will start from a new number on the Minicomputer.

Move a checker:

- from the 4-square to the 1-square (3 less)
- from the 20-square to the 40-square (20 more)
- from the 10-square to the 2-square (8 less)
- from the 1-square to the 10-square (9 more)
- from the 8-square to the 4-square (4 less)
- from the 8-square to the 2-square (6 less)

Fig. 2. Sample Lesson from Teacher's Guide, Third Grade.

Put this configuration on the Minicomputer.



T: Who can move exactly one checker and make the number 2 more than it is now?

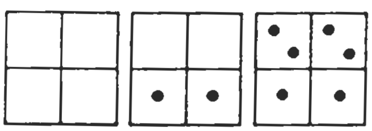
A student moves a checker from the 2-square to the 4-square.

Continue this activity by asking for volunteers to make these changes. Again, do not return checkers to their original positions; otherwise, some changes may be impossible.

- 9 more (from the 1-square to the 10-square)
- 19 more (from the 1-square to the 20-square)
- 10 less (from the 20-square to the 10-square)
- 3 less (from the 4-square to the 1-square)
- 30 less (from the 40-square to the 10-square)
- 6 more (from the 4-square to the 10-square or from the 2-square to the 8-square)
- 99 more (from the 1-square to the 100-square)

Exercise 2: Minicomputer Golf

Put this configuration on the Minicomputer.



T: What number is this?

S: 57.

T: Today we are going to play a game called "Minicomputer Golf."

Draw and label a dot for 57.

T: Our goal is to reach 200 by moving the checkers.

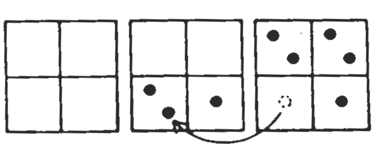
Draw and label a dot for 200.

T: Do we need to make the number on the Minicomputer larger or smaller?

S: Larger.

Invite a student to move exactly one checker from any square to any other square of the Minicomputer. When the checker has been moved, ask the student how much larger or smaller is the new number. If the student is unable to tell you, replace the checker in its previous position and ask the student to make another move. Continue in this way until the goal is reached. The move which reaches the goal is the winning move. We describe a sample game.

The first volunteer moves a checker from the 2-square to the 20-square.



T: Is the new number larger or smaller than the number before?

S: Larger.

S: 18 larger.

A +18 arrow starting at 57 records the increase.

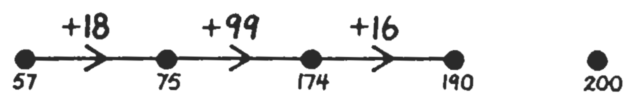
T: What number is 57 + 18? (75)

Some students might look at the Minicomputer to calculate the number (75); others might do the addition mentally. If necessary, write the addition problem on the board and solve it collectively. Label the dot with the answer.

The next volunteer moves a checker from the 1-square to the 100-square and tells the class that the number is now 99 larger. The class calculates and finds that 75 + 99 = 174, and so decides that the new number is 174.

T: Do we need to make the number on the Minicomputer larger or smaller? (Larger)
How much larger . . . ? (26)

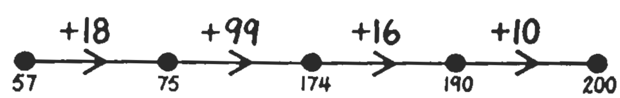
A student moves a checker from the 4-square to the 20-square.



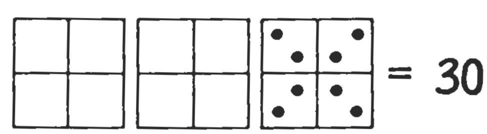
T: What is the number on the Minicomputer now? (190)
What is the distance from 190 to 200?

S: 10; we need to make the number 10 more.

A student moves a checker from the 10-square to the 20-square and the goal is reached.



Play the game a second time but start with this configuration . . .

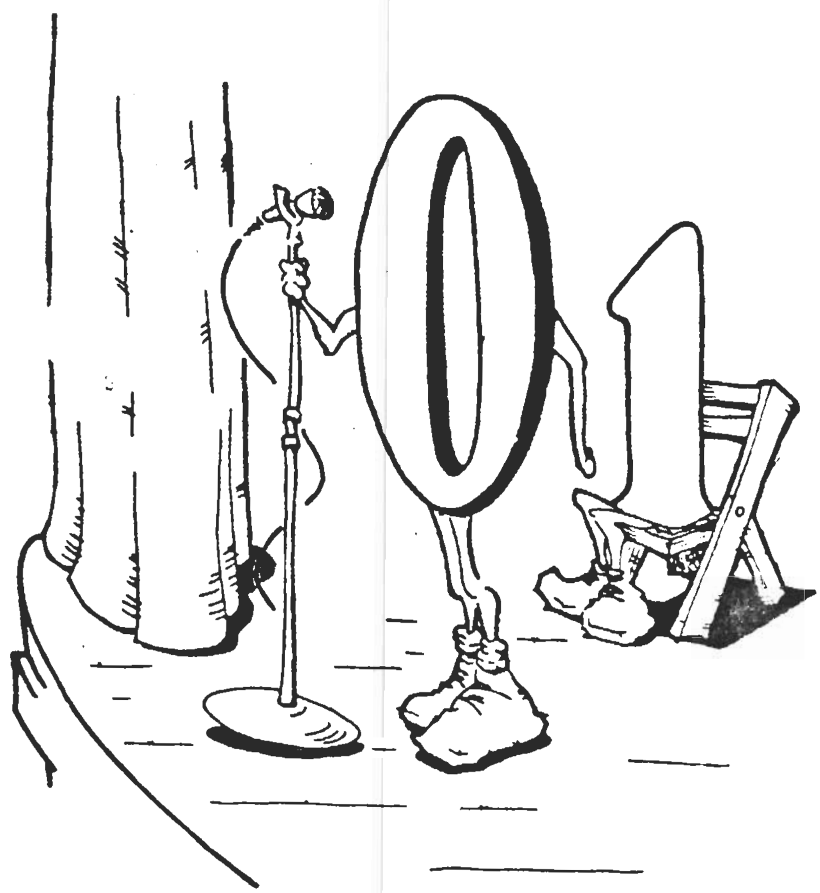


. . . and set the goal at 200.

If your class wishes to play other games of Minicomputer Golf, we suggest that you start with any number represented by eight positive checkers on the Minicomputer and choose any multiple of 100 up to 1,000 as your goal. Of course there are other choices, but they may be too difficult for your students or they may require calculations which slow the game down to the point of becoming boring.

As this sample shows, CSMP is very teacher-directed. Teachers are encouraged to follow the Teacher's Guide fairly carefully until they become comfortable with the kinds of questions and procedures intended. Because of CSMP's highly structured schedules and lessons, the Teacher's Guide is the crucial program vehicle. It provides support and instruction to the teacher from training, through practice of the lessons, and on to eventual mastery of the content and pedagogy. The guides tend to be long; the Kindergarten Guide is 514 pages, while the guides in other grades average about 1500 pages and are divided into several volumes by strands and semesters.

The lessons are based on a "pedagogy" of situations which are designed to feature both real world and fantasy situations. Numbers may be imbued with personalities and fantasy roles which support their mathematical properties. Two numbers, zero and one, are shown below as they appeared in a storybook (as the principal and vice-principal in the story about talking numbers, described earlier in this chapter).



In presenting the lessons, teachers can use a variety of materials designed by CSMP to illustrate key concepts, for example, a string game set (strings, colored geometric shapes, and score pads) or a large minicomputer with magnetic checkers (plus smaller sets for students to use).

The Teacher's Guide prepares teachers for ways of questioning that do not frequently occur in most teacher-student classroom interactions. For example, ways of eliciting multiple answers to the same problem are often modeled, as shown below from a portion of a second grade lesson plan.

Allow the discussion to continue for a while if the students remain interested. Encourage a variety of observations. You may also wish to guide the discussion by asking questions of your own.

T: Michael told me that he gave the most May baskets. Where is Michael? How many May baskets did he give?

S: Four.

Have a student point to Michael's dot; label this dot "Michael".

T: Can you find a child who gave exactly one May basket and received exactly one May basket?

Encourage the students to find several such children.

T: Michael's friend Peggy received the most May baskets. Where is Peggy? How many May baskets did she receive?

S: Five.

Call on a student to point to the dot for Peggy; label this dot "Peggy".

T: Find pairs of children who gave May baskets to each other.

There are several such pairs; encourage the class to find them all.

T: What do you think about the child with the loop?

S: He gave a May basket to himself.

T: How many May baskets were given all together? (24)
How many May baskets were received all together? (24)

Student Materials

The main student materials are consumable workbooks and worksheets. Workbooks are typically 16 pages in length, and are intended to complement in a general way the various teacher-led lessons. There are between 12 and 16 of these workbooks per grade. They are graded in difficulty, from one star (all students should be capable of doing the problems) to four stars (only the best students will be able to do them) and are to be assigned individually by the teacher according to the ability and progress of the student. Workbooks are assigned once a week and are often preceded by a teacher-led lesson to give students a preview. The schedule specifies when each workbook should be assigned; normally two or three consecutive workbook days are alloted for each workbook. Teachers grade the workbooks according to their own criteria and needs.

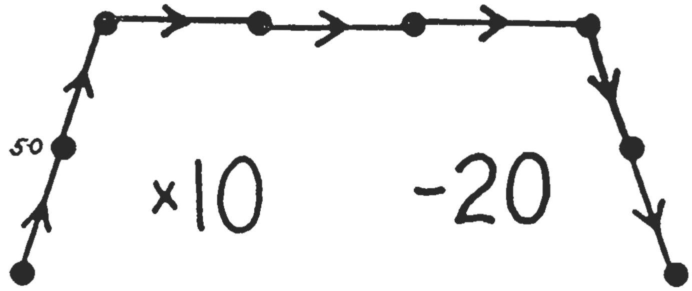
Another frequently used kind of student material is the worksheet. Worksheets are usually assigned for individual student work after each teacher-led lesson, usually one or two per lesson, according to directions in the Teacher's Guide. These worksheets appear altogether in a single, bound, consumable book, containing between about 100 and 200 individual worksheets per grade.

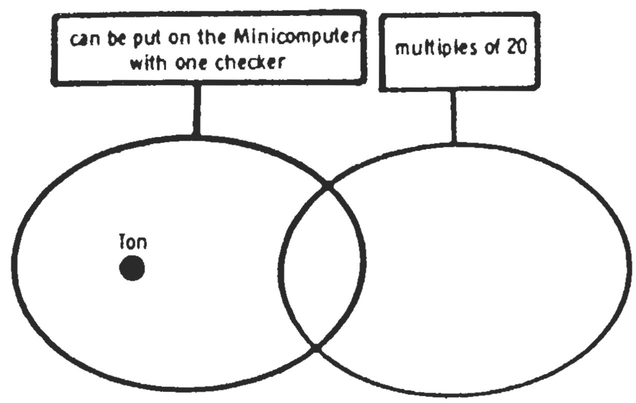
The program's emphasis on problem solving is also fostered by "Detective Stories", like the one shown below from a fourth grade student workbook, which encourage students to form hypotheses, consider alternatives, and test conclusions.

Ton is a secret number.

Ton is in this arrow picture and in this string picture.

Who is Ton? ____________





Cost of Materials

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

Table 2
Materials Costs Per Student

	Installation (Year 1)			Subsequent Years		
	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

Entry Modules

Special sequences of lessons have been developed for use with new-to-CSMP classes who are beginning third, fourth or fifth grades. (For new second grade classes there is a review built into the curriculum, which teachers can use in somewhat expanded form.) These lessons are intended to give students a rapid, intense introduction to the CSMP languages so that classes can move into the regular sequence with a delay of no more than 4 or 5 weeks at the upper grades, less at the lower grades. These modules make it possible for school districts to begin CSMP in several grades at the time, instead of implementing the usual yearly grade-by-grade advancement from K-1.

Training and Coordination

During CSMP's development and evaluation, most teachers were trained in their local district by the CSMP coordinator. Some were trained directly by CSMP staff and others by coordinators from another district. Recently, a network of Turnkey Trainers, trained by CSMP staff, has been established to assist local districts.

The mechanism for training/implementation was a cooperative agreement, the Memorandum of Understanding. Once the agreement was signed, the district was asked to appoint a local coordinator and to send that person to St. Louis for 3-10 days of training. This training usually occurred during the spring or summer prior to the first year of implementation. In turn, the coordinator assumed responsibility for training all new-to-CSMP teachers before the start of school. The smallest permissible adoption unit was one teacher in one classroom.

A Coordinator's Manual and individual training kits for teachers are available for use in teacher training. The manual presents formats for two workshops: Primary and Intermediate. Both syllabi contain an introduction to the program and to the CSMP languages. Workshop participants are expected to experience the program in much the same way as students would, i.e., they study the same problems and exercises that appear in the curriculum.

The workshop schedules are arranged in five 6-hour blocks. At the primary level it is recommended that first grade teachers attend for the first three days of the workshop, that second grade teachers attend the first four days, and that third grade teachers be present for all five days. (The number of days can be reduced by one in each case, if necessary.) It is acceptable for these numbers of days to be reduced by one each. Grades 4-6 teachers are expected to attend all five days of the Intermediate workshop. The schedules allow time for participants to look through lessons and workbooks, practice making large diagrams on the blackboard, practice using materials like the minicomputer, solve problems in the lessons, and share ideas and problem solutions with one another.

The primary workshop agenda is shown in Figure 3, next page.

DAY 1	DAY 2	DAY 3	DAY 4	DAY 5
Introduction and Opening Discussion	Computation on the Minicomputer	Composition Games	Minicomputer Tug-of-War	Minicomputer Golf
Introduction to Strings	Building Arrow Roads	Individual Minicomputer Practice	Introduction to Decimals	Decimals
Introduction to the Minicomputer	Negative Numbers	Detective Stories and Hand-calculators	A Subtraction Algorithm	Multiples and Divisors
Introduction to Arrows	A-Block Games	The String Game with A-Blocks	Probability	The String Game with Numbers
Introduction to Detective Stories	Permutations	A Multiplication Relation	Composition of Multiplication Functions	Division Problems
Order Among Integers	Multiplication by 10 on the Minicomputer	Probability	Graphs	Workbooks
Minicomputer Dynamics	Area and Perimeter	Arrow Pictures	Modular Arithmetic	Exercises for Logical Thinking
Taxi Geometry	Workbooks	An Addition Algorithm	Discussion	Games with Hand-calculators
Discussion	Discussion	Mental Arithmetic		Closing Discussion
		Discussion		

Fig. 3. Primary workshop agenda from the Coordinator's Manual.

External Review of CSMP Materials

In 1974, an external review of CSMP was conducted. CSMP materials available at that time were curriculum materials through second grade plus plans and samples from third and fourth grades. These materials were sent to the five reviewers listed below, out of a group of seventeen people recommended by the Mathematical Association of America for this task.

Professor Shirley Hill
University of Missouri at Kansas City

Professor Dan E. Christie
Bowdoin College

Professor Leonard Gillman
University of Texas at Austin

Professor George Springer
Indiana University

Professor Sherman Stein
University of California at Davis

Reviewers were asked to evaluate the "soundness and appropriateness of the materials" and "the relevance of the mathematical content." A summary of the five reviews, prepared by one of the reviewers, Dr. Shirley Hill, is given in Appendix H. Evaluation Report 1-A-2 gives all the reviews in full. Most reviewers were favorably impressed, as the following summarizing quotes show:

> "It should be stated at the outset that the CSMP materials which I have examined are impressive."
>
> "On balance, I find the materials very impressive."
>
> "On the whole I am impressed by the CSMP materials."
>
> "The authors have certainly done some good things, but their gains may be offset by other innovations which, in my opinion, should be dropped."
>
> "My opinion is that it is indeed "more of the same" (though) most of it is more skillfully written than the SMSG materials."

Reviewers generally liked the early inclusion of probability and the materials on relations and functions, graphing and arrow diagrams, and combinatorics. On the other hand all were negative toward the minicomputer, as the following quotes show:

> "a horrible aberration"
>
> "a disaster of the first magnitude"
>
> "it represents a diversion rather than a step forward"
>
> "seems a bit of a gamble and the investment is great"
>
> "I wonder whether the investment in Minicomputer skills really pays off adequately in understanding"

Summary

CSMP is a dramatic curricular innovation. During its development, conscious decisions were made about elementary school mathematics. The most important of these were the following:

> Mathematically important ideas should be introduced to children early and often. The concepts of set and relation should have a pre-eminent place in the curriculum.
>
> Mathematically rich problem solving activities should be prominent and should generate topics, guide content sequencing, and provide computation practice.
>
> The curriculum should be organized in a spiral with integration of different topics from day to day.
>
> Training for teaching CSMP should be made available to teachers as should a set of highly detailed lesson plans.

These beliefs were translated with remarkable integrity into the eventual curriculum materials and resulted in a curriculum with very distinctive features. Each of these features was a response to some aspect of mathematics education that many mathematics educators believed to be weaknesses in traditional instruction. These are outlined below in what might be called "the case for CSMP".

1. Authoritative mathematics education groups, then and now, have recommended that new content such as probability and statistics be introduced into the curriculum.

   CSMP introduces a considerable amount of new content, especially in the intermediate grades. Most topics are introduced in an informal way, with emphasis on developing teacher-led situations, and contain processes found in new content areas, such as linear programming, combinatorics, probability and statistics.

2. It is generally agreed that arithmetic skill development should be based on an understanding the processes, thus making for better recall later.
   CSMP presents numerical skills and concepts in a slightly different sequence from traditional programs. The <u>concept</u> of multiplication is introduced earlier than usual, as are decimals, fractions and negative numbers. On the other hand, many of the <u>skill</u> algorithms, such as long division, subtraction with borrowing, adding fractions and multi-digit multiplication are developed more slowly. Mastery of these skills is not intended to occur until somewhat later in the curriculum.

3. Higher order thinking skills and problem solving in general are hard to develop and teachers generally are not well prepared to teach them. Hence, they are seldom taught.

   CSMP is filled with mathematical situations which are rich in possibilities for good thinking and problem solving.

4. Recent NIE-sponsored research has indicated that teacher-led instruction which actively engages students may be more effective than assignment of individual work to students.

   The CSMP lessons extend the length of time normally spent by teachers working with the whole class, and reduce the time students spend on individual work.

5. Most elementary mathematics teachers have little formal mathematic education beyond a year or two in high school.

   CSMP has extensive training programs and materials for turnkey trainers, local coordinators and classroom teachers.

6. Many traditional programs devote long blocks of time to a single topic, such as the multiplication algorithm, before proceeding to the next block. This bores students, resulting in less positive attitudes towards mathematics in the upper elementary grades when this rote skill development is at its peak. In addition, mathematics becomes perceived as a set of disjointed, unrelated topics.

   CSMP uses a spiral approach in which a topic is taught one day but then left for a week or two and in which the same concept reappears briefly in several contexts over a long period of time. Consequently, there are few points in the sequence at which mastery is required, there is less pressure on students and the sequence of varied lessons is more interesting to them.

7. Many students enter school with very limited verbal skills and consequently have trouble understanding new mathematic concepts.

   CSMP uses various representational "languages" which are able to convey rather complicated mathematical concepts, relationships and patterns in simple ways. This reduces the verbal load on students; fewer technical words are needed and ideas that are difficult to explain verbally can be introduced earlier.

8. Mathematics education groups have called for a reduction in the huge investment of time spent by students in learning computation skills, for example, the months of student time needed to learn long division in an age of universally available calculators. National assessment data show that computation skills are being maintained far better than application skills.

   In CSMP, this investment of time is deliberately reduced, particularly on the long algorithms, leaving more time available for other topics.

9. Problem solving skills are notoriously hard to teach. Many teachers, though willing, have not learned the basic question-asking techniques that should be used in attacking a mathematical problem with a group of students.

   CSMP provides very detailed lesson guides containing sample "scripts" where good question-asking techniques are highlighted.

10. Traditional student materials are boring, and often filled with repetitive drill and practice of computation skills.

    CSMP's student workbooks are attractive, colorful and amusing. They provide an interesting variety of problems that students can solve directly on the workbook page.

In any curriculum development of the scope and vision of CSMP, hard decisions must be made about comprising the integrity of the program in order to increase the attractiveness and marketability of the product. The problem was well stated ten years ago by one of the CSMP External Reviewers (see previous section):

> "...in order to sustain a group of authors over the years of developing and testing such a gigantic endeavor, the leader of the group must sustain an esprit de corps, a dedication, a self confidence among his colleagues that borders on the ecstatic. Such enthusiasm is necessary, and it is dangerous. The group must zealously believe in the uniqueness and value of its creation, yet keep an open mind. It must blend religious dedication with scientific neutrality."

Whether or not CSMP succeeded in maintaining this delicate balance is open to question. At least one reviewer thought not: "It (CSMP) has its own private religion, complete with rituals, which often become obsessions." No doubt this is an extreme and minority view, but the uncompromising stance of the developers did result in a product which was viewed by some educators as too radically different. No other curriculum has such a detailed and extensive Teacher's Guide, introduces as many topics at as early an age (for example, decimals, multiplication, negative numbers), makes such extensive use of representational devices, devotes as much time to probability and geometry, has as "loose" a spiral as CSMP's spiral organization of content, and devotes as little time to rote computational skills and algorithms as CSMP.

Each of CSMP's distinctive features, desirable though they were thought to be by most mathematics educators, created problems in one way or another for districts wishing to implement CSMP. In addition, although the perceived weaknesses in traditional mathematics instruction have continued to exist during CSMP's long development, the context in which the program was implemented changed continually. At the national level, there is a long list of factors which have changed the way school districts operate. The list includes:

the move toward mastery learning,

increased use of computers,

an emerging consideration of teacher accountability,

the recent re-examination of American education,

the growing number of state and locally mandated tests,

the national shortage of mathematics teachers,

increased financial pressure on most school districts,

changes in textbook adoption procedures,

the push for better problem solving by professional teacher organizations, and

the emergence of the National Diffusion Network.

Each of them altered somewhat the rules of the (CSMP implementation) game.

III. CSMP IMPLEMENTATION

Overview

This chapter will describe CSMP's implementation and how the characteristics of the program and decisions by the adopting sites affected the success of the program. The implementation of CSMP will be presented chronologically from adoption onwards, concluding with the experiences and reactions of teachers.

Before beginning a chronological description, it will be useful, as an overview, to consider the relationships among three different aspects of the implementation process, and how these effect what eventually takes place in classrooms.

1. Program Requirements. CSMP has been implemented in many different ways. The stated requirements were often compromised in practice, but there are four considerations that any adopting site must attend to:

> Costs of materials. Start-up costs for CSMP are slightly higher than for other programs because of the extensive teacher materials, but are well within the normal range. Maintenance costs for CSMP through third grade are roughly comparable to other programs; both use consumable students materials. However, beyond third grade CSMP continues to use consumable materials at a cost of about $7.00 per student per year. This is considerably more than other programs using textbooks which last several years.
>
> Teacher training. Although a few teachers were capable of learning the program at the same time as their students, it was necessary for the districts to establish training programs. Coordinators had to be trained at CSMP workshops, and teacher training required either direct stipends for summer training or payment to substitute teachers if training occurred during the school year. In some districts, professional development days were available, thus reducing training costs. Personnel were needed to conduct the initial training and to assist teachers when they returned to the classroom. In succeeding years, training had to be extended as new teachers joined the district and as new grades or schools began the program.
>
> Program management. In addition to overseeing or conducting teacher training, the local coordinator was also responsible for ordering and distributing materials, describing the program to district staff and parents, troubleshooting in areas such as testing and funding, and planning further implementation of the program.
>
> CSMP pedagogical characteristics. The distinctive features of CSMP, summarized at the end of the last chapter, all had ramifications for adopting districts. They made CSMP different from what districts were used to in a mathematics program. Of course, it was in the classroom that these characteristics had their most dramatic effects, but because there are so many characteristics and they are so distinctive, they also affected events at the district level.

2. Local Setting. There were some relatively fixed conditions at each site at the time of implementation. Size and location of the district, average class size, and type of student population had some effect on quality of implementation. Some less clearly defined factors also affected the program, such as the role of building principals and the district's reasons for adopting CSMP. But in retrospect, existing local conditions had a relatively small effect on the program, except for two factors which were very important to the program's success: the existence of a skilled and influential coordinator for the program at the district level and the availability of continued funding support for the program. This second factor was often a result of the first.

3. Local Decisions and Events. The way a district chose to implement CSMP and the way it dealt with CSMP's special characteristics, i.e., decisions made when adopting and implementing the program, largely determined how successfully the program was implemented in a district. Some of the local decisions concerned how to respond to general district events which could affect the program such as a change in the testing program.

4. Classroom Effects. There was a surprising consistency in teacher reactions to CSMP regardless of grade level, teaching experience, ability of students, and pre-service training. For example, a significant minority of teachers thought the program was less appropriate for low ability students, primarily because of its de-emphasis on computational skills, but the proportion of teachers holding this belief seemed relatively unaffected by these factors. However, local decisions and events, such as high level support for the program, amount of training provided, accountability constraints, and pattern of adoption by grade and school had a significant bearing on how faithfully CSMP was taught.

Extent of CSMP Use During the 1981-82 school year, the last year for which reliable data is available, CSMP was being used by about 50,000 students in over 100 school districts. Of these school districts, 6 were large urban districts, and 17 were rural or small town. The remainder were about evenly divided between suburban districts and medium-sized cities. Most districts were public school districts but 23 of the districts were private or parochial.

Most of the districts used CSMP as the regular mathematics program, but 12 districts used it primarily with gifted students. In 14 districts it was a Chapter I program or remedial program.

From the beginning of the Extended Pilot Tests in 1973-74 through 1981-82, the program was used in 134 sites. Many sites have been either in the midwest, especially the St. Louis area and Michigan, Wisconsin, Illinois and Kentucky, or in the east, especially New York, Pennsylvania and Maryland. There have been relatively fewer sites in the west, northwest and plains states.

Initial Implementation of CSMP

There were several stages in the initial implementation of CSMP adoption: awareness; follow-up to awareness; decision to adopt; and strategies for first year implementation.

Awareness.

Districts learned about CSMP in several ways. At some sites the mathematics educators' grapevine spread the word about CSMP to local administrators or teachers, who then brought it to the attention of district decision-makers and lobbied for its adoption. Alternatively, school district personnel read about CSMP in educational journals or through presentations sponsored by groups like the National Conference of Teachers of Mathematics. Occasionally, someone who had been a CSMP Coordinator at one site would move to another school district. More recently, after approval by the Joint Dissemination Review Panel, awareness was fostered through the National Diffusion Network. Adoption of CSMP by Chapter I sites has been attributable in large measure to NDN-sponsored awareness sessions, since CSMP is one of the few Chapter I eligible projects in mathematics. Quite often, a local administrator found out about the program from an administrator in a nearby district (this was particularly true for what have been termed "lighthouse" sites described below) or from the same specialized area (such as a fellow coordinator of Gifted programs). In a survey of 55 coordinators whose district started using the program since 1978, personal contact was listed as the most popular method of finding out about the program (15% of the respondents). But eight other methods were listed by 6% to 11% of the coordinators surveyed: CEMREL contact, Gifted and Talented Conference, literature, NDN conference, university course, Chapter I conference, CSMP used in the area, and CSMP-sponsored awareness workshop.

Follow-Up.

When a district learned about CSMP, district personnel usually contacted either the CSMP staff or another district where CSMP was already being used. They arranged to watch CSMP being taught, interviewed teachers and administrators, reviewed curriculum materials, and learned about the adoption-training-implementation process. Occasionally, an interested school district would request a CSMP staff member to visit the site and conduct an awareness session for school personnel, board members, and even parents. Alternatively, district personnel visited CEMREL in St. Louis and discussed the program with CSMP staff.

The presence of nearby CSMP sites was very helpful for prospective adopters. Adoptions in the first few years were sufficiently far-flung that districts in many regions of the country could more conveniently visit a relatively nearby site and see the program in action rather than traveling to CEMREL. At certain "lighthouse" sites, coordinators were so convinced of CSMP's value as a mathematics program that they took the initiative in persuading neighboring school districts to watch it being taught, to adopt it and to push for its implementation.

Figure 4 shows the distribution of several lighthouse sites, as well as the sites which adopted the program based on visits to those sites.

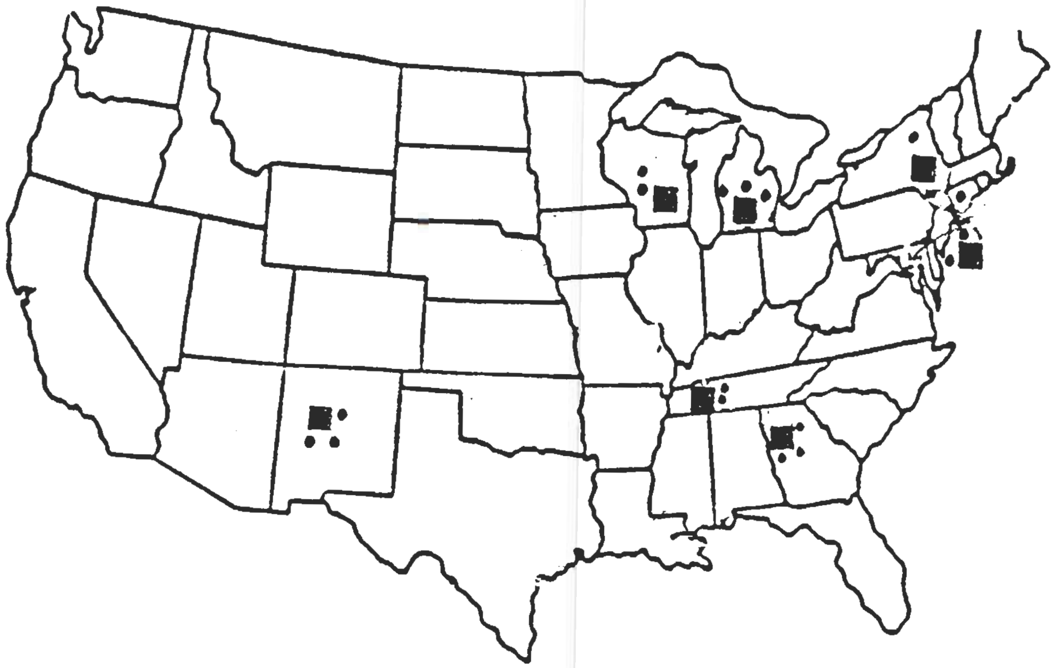


Fig. 4. Distribution of "Lighthouse" sites (■)
and subsequent adopters (•)

Decision to Adopt.

The most common reason given for deciding to use CSMP was dissatisfaction with the present curriculum on the part of a mathematics supervisor or other district personnel. A lack of materials for teaching problem solving or thinking skills, and the consequent dreary emphasis on computational skills, were cited as weaknesses in their present program. Hence, the detailed CSMP Teacher's Guides, with their heavy emphasis on the discovery approach and on question-asking techniques, were particularly attractive to these educators. Visits to existing sites, where they could observe students' responses to the materials, were often persuasive. Many districts were looking specifically for a math program for either gifted students or Chapter I students. Adoption of CSMP by Chapter I schools has increased recently attributable largely to heightened awareness of the program through the National Diffusion Network. Gifted sites chose CSMP because it provides the type of problem solving deemed appropriate for higher ability students and it contains more mathematics and more different topics in mathematics than most commercially available projects.

But occasionally ulterior motives were prominent:

> a desire to be innovative and make change for change sake when federal or state dollars were available to support the start of the program, with no long range goal of total local financial responsibility.
>
> an opportunity to provide badly needed general mathematics training for teachers which might improve instruction regardless of eventual CSMP implementation.
>
> a desire to raise test scores in general.
>
> the appeal, for kindergarten and first grade teachers, of CSMP's manipulatives, stories and games.
>
> the availability, to programs for gifted students, of genuinely challenging mathematics without the need for acceleration through grade levels.

To get CSMP adopted, coordinators-to-be first had to persuade the school administration and/or school board to try out the program. This meant addressing two primary issues: how to pay for the program and how to evaluate it (at least informally) after some period of trial. At the same time, coordinators had to persuade school principals and teachers to use the program. They used several methods: active participation by teachers and principals in the adoption decision, complete voluntarism, gentle arm twisting, and administrative decree.

In order to begin using CSMP, a school district had to sign a Memorandum of Understanding with CEMREL. In this memorandum the district formally named its coordinator and agreed to provide CSMP teachers with the recommended amount of training.

Initial Implementation Strategies.

Selecting CSMP Classes School districts did not usually begin using CSMP at all grades at the same time, since it was difficult for students to plunge right into the CSMP curriculum without previous experience, especially in the upper grades. The most common starting points were K-1 and K-2, and occasionally K-3. It was also unusual for a district to begin using the program at all schools at the same time, unless it was a one-school district. To begin with, such an undertaking would have required a massive training effort by district personnel with no previous CSMP experience. In addition, districts felt they needed time to get the inevitable bugs out of the program, get it publicized within the district, and find out how students and teachers reacted to it.

Two strategies were used most often: either select a judiciously chosen school and implement CSMP throughout K-1 or K-2, or ask for volunteers in those grades at two or three schools. These strategies were used about equally often. During that pilot phase, while everyone scrutinized CSMP, the coordinator encouraged other teachers, other grade levels and/or other schools to participate. Sometimes CSMP never moved beyond second or third grade and sometimes never moved beyond one or two schools. But in most cases, the district went from volunteers at the start to selection of teachers/grades/schools at a later date.

Whatever the start-up strategy, the school usually became the eventual unit of implementation; some schools were CSMP schools - all classes used CSMP through a certain grade -while other schools didn't use any CSMP. Coordinators usually found it impossible to continue the program in a school where, at some grade levels, some students did have previous CSMP instruction while others did not. If only a single school in a large or medium-sized district adopted CSMP, the program was not likely to be continued, either in that school or in the district as a whole. CSMP was likely to get lost amidst all the other district-wide policies and practices. The only exception to this pattern, and it is a major exception, was when CSMP was adopted by a single school in a parochial school system. There is more autonomy for individual schools in those systems and so CSMP was more likely to survive as an adoption.

Rapid Implementation Some school districts decided to start the program in all classes K-3, K-4, or K-5 of one or more schools, rather than beginning K-1 or K-2 and advancing a year at a time. Altogether, 15% - 20% of CSMP districts used this rapid implementation model. In most cases, the model was used in a single school, sometimes the only school in the district.

A little more than half of these districts began K-3, the others K-4 or K-5. Altogether 19 of the districts used CSMP long enough to have a track record. Of these nineteen:

> nine had a very successful implementation which eventually went K-6: three in single-school districts, three in multi-school districts and three in one school of a multi-school district,
>
> six either grew at a slower pace or stayed the same,
>
> four were unsuccessful, two reaching K-6 status and then dropping the program and two maintaining the program on a much reduced basis.

In using the rapid implementation model, coordinators chose to put a very concentrated effort into a single year. There turned out to be several advantages and disadvantages to this decision. The biggest advantage was that after the first hectic year, the implementation settled down with confidence. Many of the uncertainties associated with start up (training, parent awareness, resistance of teachers to begin a new program, rationalization of CSMP with district guidelines) had been overcome. Financially, it was sometimes advantageous to get a sizeable one-year grant for teacher training rather than smaller amounts for several years. Psychologically, it was easier to motivate the whole teaching staff together in one year; upper grade teachers were less likely to feel like outsiders and common problems could be attacked by all staff.

On the other hand, the first year was very hectic. The coordinator had to be in a position to fully support the teachers over the course of the year in addition to providing solid training before school started. Teachers beyond second grade had to use special entry modules to prepare students in the CSMP languages. There were no colleagues with hands-on experience who could provide moral and practical support. Coordinators had to be able to anticipate negative teacher reactions about some aspects of the program; those which normally grow in importance from grade-to-grade would be full blown without the usual warning signals from the lower grades.

The rapid implementation model was a gamble, but turned out to be fairly successful. This was probably because it was usually undertaken only by coordinators who did their homework about CSMP, worked very hard, and were able to marshal some special resources for a year or two.

Special Adoptions. A unique feature of CSMP's implementations history is the diversity of sites which adopted the program. Several Indian Reservations adopted CSMP and used it with varying degrees of success. Aides were often called upon to translate CSMP's special vocabulary into the students' native language with some degree of apparent success. Teachers had mixed reactions to CSMP. However, their influence in the CSMP decision was limited, since these schools were administered by the federal government and tended to be centrally operated.

Title I sites were attracted to the CSMP curriculum because of its motivational characteristics for younger children. Though most Title I teachers were well satisfied with this aspect of CSMP, many standardized test scores did not improve. Since these test scores play a major role in Title I evaluation, adoption has been lower at the higher grades where motivational characteristics are less persuasive.

Where CSMP was used as a gifted program, the situation was quite different. Standardized test scores were less important as districts searched for more appropriate instruments. Teachers were pleased by the challenging nature of parts of the curriculum and its emphasis on problem solving. Coordinators saw CSMP as one of the few alternatives to acceleration.

In all three of these special types of sites, CSMP costs were less crucial than in regular implementations because special money was available over and above the usual textbook allotments. Administration of the program was easier because it was part of a centrally administered division.

Training CSMP Teachers. The original Memorandum of Understanding called for CSMP teachers to be provided with a certain amount of training, roughly a week for teachers of primary grades and two weeks at the intermediate level. Later these numbers were reduced to between one and five days, depending on grade level. Schools tried to accomplish this in one of two ways: a solid block of time in the preceding summer or one or two days before school plus odd days or afternoons during the year. In either case it was difficult for most districts to achieve the recommended amounts of time; well over half of all CSMP teachers did not receive the mandated levels of training. In many school districts, especially the larger ones, there were very precisely defined union agreements about what teachers could and could not be asked to do outside of the regular teaching hours (i.e. 8:30 - 3:00), and coordinators had to grab an hour here or there with a few teachers as best they could. At other sites, many teachers willingly gave up part of their summer vacation for an unpaid week of training.

Many districts, especially in the metropolitan St. Louis area, were able to take advantage of CSMP's training program and sent a few of their teachers to a St. Louis workshop. A helpful factor in some sites was the presence of an experienced trainer in the area. Since the CSMP-CEMREL staff could not visit all potential sites and could not train all potential adopters, unofficial "turnkey" trainers trained at CEMREL were able to train teachers in their region. The distribution of turnkey trainers and the sites they visited is shown in Figure 5.

Fig. 5. Distribution of CSMP turnkey trainers (▲) and adopting sites (●)

Funding. Only about one-third of the districts supported the program from the start entirely with regular district funds. Special funding of one kind or another - state, local or federal -was usually used to start the program. The most common support, used in about one-third of the districts, was through federal Title IV assistance. Other sources of support were state or local grants, usually for gifted or remedial programs, and federal Basic Skills grants. In any case, these special funds were not intended to be permanent endowments for the program.

Problems associated with funding were related to several other factors. Some districts did not anticipate the true costs of CSMP - costs associated with first time and on-going teacher training, and costs associated with replacement of students' consumable materials. For some districts, the problem of anticipating budgets was compounded by the fact that they were initially attracted to CSMP for short term reasons rather than to meet long term goals. Since "soft" money was available, several school districts elected to give CSMP a try knowing that the teacher training component would provide needed mathematics inservice training. A few districts had fallen into a pattern of adopting one or more innovations each year in a fairly hit or miss fashion. For those districts, CSMP was just one of many curriculum programs tried, all of which couldn't be afforded at the same time for very long. For all districts, funding became a more important consideration with the increasing financial pressure on schools that began in the late 1970's.

Continuing Implementation

Decision Making After Year One

During the first year that CSMP materials were used in a district, the program was the object of careful scrutiny and coordinators were on the spot. For almost all districts this first year went quite successfully and was helped by several factors. Initial implementation was most often in grades K-2, which turned out to be the grades in which the program was best liked by teachers. Teachers and schools either volunteered or were selected because of their probable receptivity to a program like CSMP. Money was often available from special funds for this pilot activity. Because the pilot was usually in only a few classes, the coordinators were able to monitor the program and help teachers on an individual basis. Participants had a natural enthusiasm for being part their district's lead group in an exciting, innovative program.

In spite of its early success, by the end of the first year most coordinators came to the realization that district-wide implementation would take more than a year or two. The logistics of teacher training were formidable. Implementing the program at more than one grade level at a time was difficult. Material costs were likely to be a problem without special funds which might not continue to be available.

At the same time some disturbing features began to appear, each of them destined to be a bigger nuisance with each successive grade level. The program required more bookkeeping than anticipated; materials were complicated to order, shipments had to be checked and distributed, and orders were late in arriving. At the classroom level, there was a bewildering array of materials. Teacher's Guides, workbooks, worksheets, demonstration materials, and manipulatives had to be stored and kept track of. As teachers came and went, there was a continual need for new teachers to receive teacher training. Most teachers were not able to complete the schedule of lessons in the required time. Pressures were developing in some districts for the program to prove itself on the district standardized tests while, at the same time, some teachers complained about the lack of drill and practice and began supplementing the program. A few teachers complained about the program not being good for lower ability students who couldn't follow some of the lessons, and didn't seem to be getting proficient at the CSMP languages. Some teachers did not like the spiral approach.

Overall the program was very well liked in the early grades and most of the teachers who made these complaints were nevertheless strong supporters of CSMP. But with a "second wave" of teachers to be introduced to the program, the difficulties worsened and sometimes proved insurmountable. Those teachers, often less venturesome than the first wave of teachers, and often less confident about their mathematical abilities, were reluctant to volunteer (or be volunteered).

In a few sites, these problems, and insufficient enthusiasm for the program, were enough for the district to put CSMP expansion on hold or even drop it. But most often, coordinators started to plan for an expanded implementation. Given the warning signals described above, they planned for a rather modest expansion, i.e., bringing another school or two into the program, consolidating it in the pilot school, and starting a few teachers at the next grade level with experienced CSMP students.

Often, after about the second year of implementation, the school district administration began to look carefully at the program. Up to this point, the administration usually had been content to approve a gradually increasing pilot stage. But as the implementation got larger, more visible, and inevitably controversial, senior administrators began to think about long range plans. There were four crucial considerations for the administration: test data became available, teachers' reactions were formally sought, total program costs could be fairly well projected, and the scope of the required teacher training effort could be determined.

These factors were judged in light of the present district mathematics curriculum. Obviously CSMP costs and teacher training demands were higher so that unless test data or teacher reaction indicated an improvement over the project curriculum, the administration was likely to be lukewarm to further expansion.

Typical Patterns of Change.

Once districts had decided to adopt CSMP and had decided on funding, training, and initial implementation, it became their responsibility to continue it, monitor it and make decisions concerning its implementation and continuation.

CSMP had at least three major patterns of adoption/continuation:

1. CSMP was adopted for a year or two, after a very limited trial, and then dropped. Often the adoption was on a limited basis such as at only one grade level or in only two or three classes.

2. CSMP was adopted for several years (3-10+) but there was an "ebb and flow" phenomenon associated with its implementation. From year to year the number of participating schools/grades/classrooms fluctuated with no stable pattern of consolidation or dispersion.

3. CSMP was adopted for several years (3-10+) and was successively adopted at each grade level and in more classrooms and schools each year.

Table 3 summarizes, for each year, the number of districts which began using the program and how long they continued to use it. The lower diagonal represents the numbers of districts continuing to use the program in 1981-82.

Table 3
Length of Adoption by Adoption Year

	Year of Initial Adoption								
	1973-74	1974-75	1975-76	1976-77	1977-78	1978-79	1979-80	1980-81	1981-82
Numb. of new sites	28	6	15	8	3	5	22	22	
Numb. of these sites continuing for:									
1 year	3	1	6	0	1	0	0	1	23
2 years	9	1	4	0	0	0	3	21	
3 years	1	0	1	1	0	1	19		
4 years	3	0	0	2	0	4			
5 years	0	0	0	1	2				
6 years	0	1	1	4					
7 years	0	0	3						
8 years	3	3							
9 years	9								

A total of 29 sites used CSMP for only a year or two before dropping it, 15 sites dropped it sometime after the second year, and 44 sites were still using CSMP after at least three years. The table indicates that most districts who began using the program since 1976-77 continued to use it in 1981-82; more recent data corroborates this finding.

Table 4 tells only part of the story because each year some districts were dropping CSMP, other districts were adopting it, and still others were maintaining it. Table 4 shows the number of sites dropping and adding each year. The percentage of sites continuing is also shown.

Table 4
Changes in the Number of CSMP Sites, by Year

Year	From previous year: # sites continuing	# sites dropping	Percent continuation	New Sites	Total
1973-74				28	28
1974-75	25	3	86	6	31
1975-76	21	10	68	15	36
1976-77	28	8	78	8	36
1977-78	29	7	81	3	32
1978-79	30	2	94	5	35
1979-80	34	1	97	22	56
1980-81	53	3	95	22	75
1981-82	65	10	87	23	88

After six years of fairly stable usage (always between 28 and 36 districts), there has been a steady increase, beginning in 1979-80, in the number of districts using CSMP.

The fluctuations in adoptions from year to year are attributable to several factors. Shifts in federal priorities and directives for educational laboratories affected the intensity of CEMREL dissemination efforts as well as the distribution of sites. Within CEMREL, the acquisition of staff with specific responsibilities for dissemination of program information increased the intensity of adoption efforts and the provision for program continuation. Outside of CEMREL, the establishment of the National Diffusion Network (NDN) facilitated awareness of CSMP and provided funds for adoption. Special monies, Title IV-C for example, served as an inducement for many sites to review their programs and select innovative programs designed to meet special needs.

CEMREL's own mandate from the government also affected adoptions. Over the years, the government first counseled CEMREL to look for a national audience for the program, then to focus on attracting large urban school systems to the program, and then to turn attention to potential adopters within the ten state region defined for CEMREL by the National Institute of Education, CEMREL's funding agency. These shifts in focus affected the dissemination staff's emphasis on adoption and implementation.

Implementation by Grade Level

The program was implemented more frequently at the primary level. Table 5 shows the number of sites which implemented it at each grade level.

Table 5
Number of Sites By Grade and By Year

	K	1	2	3	4	5	6
1973-74	29	31					
1974-75	28	29	18				
1975-76	31	29	24	18			
1976-77	24	29	27	23	16		
1977-78	25	31	28	27	22	12	
1978-79	30	34	32	29	22	15	7
1979-80	36	46	42	36	25	17	12
1980-81	49	51	46	40	27	19	16
1981-82	58	60	58	48	32	25	17

The table shows that CSMP implementation declined after third grade. There may be several reasons for this pattern. Since some sites adopted a gradual approach to implementation, and elected to begin using it at kindergarten the first year, first grade the next year, and so on, it would take a few years for CSMP to work its way up through the grades. But this can't be the only reason, since CSMP often was not used beyond third or fourth grade in sites where there was ample time for this to happen.

One factor in this lack of use in the higher grades was money; materials for the upper elementary grades cost more than schools are used to spending in those grades. Another reason was training, which is lengthier for the upper elementary grades. Also, the mathematics is more difficult and novel at higher grade levels and so teachers may have been reluctant to tackle the relatively difficult lessons. The physical materials (student booklets and Teacher's Guides) are also more voluminous after third grade. In some districts, there is a very real difference between what is viewed as appropriate mathematics for K-3 and what is viewed as appropriate for 4-6. In those districts, many teachers beyond grade 3 didn't view CSMP as "real" math; activities, games and mathematical stories were no longer as acceptable in the business of learning mathematics. Finally, in districts where there was a grade-by-grade adoption strategy, an "old-boy" network sometimes developed among teachers. Upper level teachers became increasingly isolated from the interchanges among CSMP teachers at the lower grade levels, hence, resistant to implementing it when their turn came. By then, many had adopted a defensive stance vis-a-vis "their" math program and efforts to recruit them for CSMP may have been less vigorous than they were for teachers at lower grades.

Adaptations

Adaptations at the district level took several forms. Some adaptations of the program took place even before the first implementation; districts knew in advance that it would use CSMP in special ways, for example, with gifted students or as a supplement to a regular textbook.

But most adaptations occurred after one or two year's experience with the curriculum and as usage was expanding. Below is a list of some of the common adaptations that were made on a formal basis by districts, or by a school within a district. Some of them were only extensions or additions to the programs; others changed it considerably.

Responsibility for various coordinator tasks were delegated; for example, someone else might be responsible for some of the schools, or one person had responsibility for training and another for ordering materials.

The length of the mathematics class was officially extended to take into account the longer CSMP lessons and/or the need to provide supplemental instruction in computation.

Various materials were xeroxed in order to reduce costs.

Lists of instructional objectives that were considered important but not covered fully enough in CSMP were prepared for teachers, who were then responsible for their students' attainment of those objectives.

Grading standards for CSMP worksheets were established, with remediation to be provided for students who did not reach the standard.

Students within a school were assigned to CSMP on the basis of parent decision or ability. In the latter case, CSMP became the upper track program.

The schedule of lessons in the Teacher's Guide was changed, either by deleting certain blocks of lessons or by collecting together groups of spread-out lessons into a single block, i.e., moderating CSMP's spiral approach.

Teachers were assigned as teams, with team members teaching either the upper or lower ability students of a pair of classes or teaching certain lessons to both classes.

CSMP tests were developed for periodic administration by all teachers, to be used as progress checks or for grading purposes.

Teacher training programs were adapted in every conceivable way.

Special materials and workshop formats were developed for use with parents.

Many of these adaptations were made in other districts by individual teachers, but never as successfully as when done on an official basis. Most of the changes described above were made in districts where CSMP went very successfully; they were sensible decisions made in reaction to concerns of teacher and administrators who liked the program, and they strengthened the program's standing in the district.

Failing to respond constructively to concerns about the program, or allowing a laissez-faire attitude toward teachers' individual (and sometimes idiosyncratic) adaptations, usually meant trouble later as the program came to be implemented in a less standardized way. Within limits, it was better to admit the problem and solve it than to ignore it.

The CSMP Coordinator

Kinds of Coordinators

In districts where CSMP was successful, the coordinators were a major factor because of their positions in the district, their belief in CSMP's goals and their degree of active sponsorship. Active sponsorship flowed from a firm belief in CSMP's goals, and was most effective when the coordinator was well-placed in the district's administrative hierarchy.

One of the key factors in the success of CSMP as a national program was its insistence that adopting school districts appoint a "coordinator" (usually a local administrator or teacher) who assumed day-to-day responsibility for the project by ordering supplies, conducting in-service and monitoring teachers as they taught CSMP lessons.

Districts had different strategies in selecting coordinators, and the choice affected the program at some sites. The adoption/innovation literature is full of case studies of adoptions which failed because sponsorship of a program was not well placed. The CSMP experience supports this literature. In a few cases a willing volunteer teacher espoused the program, pushed for its adoption, and was given coordinator duties but not administrative authority. In these cases, CSMP limped along, and was eventually dropped. The same was usually true when the principal of a school was the sponsor. It was difficult for the principal to get out of his or her own school into other schools, much less to effect a system-wide advocacy for the program. In contrast, a well-placed sponsor with district-wide responsibilities was a distinct advantage and in many cases protected the program when district leadership or goals changed, when standardized testing or accountability pressures mounted, or when new funding sources had to be found.

There were four different types of coordinators: outsiders, teachers, administrator custodians and administrator sponsors. Outsiders were typically math professors at local universities who volunteered to introduce CSMP to the district and support its implementation by conducting in-service and monitoring classrooms. They were generally able to galvanize teachers to adopt and implement the program, but they lacked the "clout" - the entree to decision-makers and sustained access to teachers - which was necessary to create a long-term CSMP commitment by the district. If a school superintendent changed, or policy shifts occurred, the "outsider" was usually not able to protect the program. When a decision regarding CSMP's future in the district was being made, the outside coordinator was not in a position to affect the decision.

At some sites, a teacher was the catalyst for adoption. Aroused by a CSMP awareness session or a report from a colleague in a neighboring district, a teacher would adopt CSMP in his or her own classroom or try to spearhead a building/district-wide adoption effort. These efforts, while successful in the short run, were unsuccessful in the long run. Teachers were not in a position to affect policy and couldn't secure funding needed to sustain the program. They lacked sufficient mobility within their own building, and from their building to other buildings, to create enough momentum for CSMP to take hold on a large scale. On the face of it, while they might seem to be a natural source of diffusion, teachers were not able to promote the program effectively. They were as impotent as outsiders when it came to advocating the program or protecting it in a district's budget.

Central office coordinators were more beneficial to CSMP's longevity. They were around when funding and staffing decisions were made; they had the visibility and the mobility to advertise the project within the district, and they had the authority to monitor and critique its implementation. At one site where the program was used with gifted students, the CSMP Coordinator was also the gifted coordinator. According to him, CSMP survived because the implementation effort kept a low profile, with little publicity and few demands on teachers or resources. The arrival of a new superintendent created a desire to reduce the visibility of the program further and to wait for the proper time to dramatize the program and its effects. So, even though the teachers in regular classes and the local math coordinator wanted to use CSMP district-wide, the coordinator's reading of the situation was to take a wait and see attitude. An outsider isn't as good at reading internal district politics and responding effectively to them.

There were two kinds of administrative coordinators. "Custodians" treated the program like any other project and merely carried out their duties as specified by the Memorandum of Understanding. "Sponsors", on the other hand, were firm advocates of the program. They were usually the ones who brought the program into the district, went to bat for its adoption, and acted as trouble shooters. When funds were low, they tried to find other ways to finance it; when teachers seemed to need more in-service they arranged for it, and when there were questions about the program's impact on students they went out and contracted for evaluations so the program could be considered on its merits. When CSMP was "in trouble" in a district, a sponsoring coordinator would often regard the difficulties as minimal while a custodial coordinator viewed the difficulties as yet one more obstacle to continuation.

Some of these district-level coordinators were math educators first and administrators second; for others the reverse was true. Being mathematically trained helped some to understand the goals of the program (which were not always spelled out). They were better prepared than their less mathematically sophisticated colleagues to present the mathematical content and processes during in-service. But others who, did not have a strong math background but who did understand the general conceptual development that CSMP aimed for, were also effective sponsors. Either a strong math background, or an understanding of the aims and the pedagogy to support those aims, was necessary for successful coordination. Otherwise, the program was a flash in the pan at some sites.

In 1981, eighteen coordinators were interviewed as part of a series of site visits. Seven of them were in central office staff positions, six had mathematics supervisory roles, three were school principals and two were classroom teachers. Not one had CSMP coordinating as the sole role. Thus, it is not surprising that three quarters of the coordinators reported that they attended to CSMP responsibilities "infrequently". For some coordinators, their CSMP functions constituted a second, almost full-time job. Acting on the specifics of the Memorandum of Understanding, they ordered materials for the district, attended CEMREL's in-service, conducted district in-service, monitored classes, critiqued and demonstrated lessons, met with parents, and arranged for CSMP's impact on students to be evaluated; all these were in addition to their other duties such as coordinating the district's gifted program or administering the curriculum division.

Other coordinators treated CSMP as a part time responsibility and delegated most work. They had teachers order the materials, let the math coordinator supervise the classroom teaching, recruited district research staff to gather evaluation data, etc. In many cases this was not from lack of interest in the program, but from lack of time to fill multiple roles.

Classroom visits were the most common activity undertaken by these coordinators (about 65% reported this activity), and evaluation activities were undertaken by half the coordinators. Only four of the eighteen conducted training; the rest turned that responsibility over to a turnkey trainer or others in their school district. While many of the coordinators interviewed in 1981 had direct personal involvement with CSMP and were responsible for initiating its adoption and participating in training, others inherited the job from the previous coordinator or from an interested advocate within the system but had no ownership involvement themselves.

Three-quarters of the coordinators viewed themselves as ultimately responsible for decisions specific to CSMP's day-to-day operations but were not the ones making decisions about renewed funding for CSMP. The majority of the coordinators reported funding the program out of their district's operating budget. A school's textbook fund or the district's operating funds were generally used for books and supplies. Thus, and unless prices for materials continued to rise dramatically, most of those coordinators thought they would be able to continue the program in spite of the fiscal problems facing their districts. That may be realistic, but data from previous years show that other sites which had adopted the program and intended to continue it were not able to because of program costs.

The intrinsic merit of CSMP was often named as the key factor in coordinators' efforts on behalf of the program. Several coordinators commented that they were looking for a program with a problem solving orientation and CSMP met those requirements. Those coordinators said CSMP was "the best program available", "way ahead of any other available text", "a <u>thinking</u> program", and "not a bandwagon approach".

The relationship between the coordinator and the building principals varied enormously. In most schools, principals were influential in adoption decisions, particularly when they had spending authority for textbooks and materials. Some principals were instructional leaders in their schools and greatly facilitated teachers' attempts to implement the program. This kind of active participation relieved coordinators of some of the day-to-day tasks that required school visits.

In other schools, especially large schools in large districts, principals took a managerial role instead. Though they cooperated with coordinators in logistic matters, they did not really learn much about the program. Their evaluation of the program was based mostly on their teachers' reactions to it, how smoothly it went, and how well their students performed on district-administered tests. If this information convinced them of CSMP's merit, they were very supportive. But such principals liked to run a smooth ship and differences of opinion about CSMP on the part of their teachers caused them great concern. Many of these principals were subjected to pressure from the central office to improve standardized test scores. Not really knowing the program, and the unmeasured learning that might result from it, they equated extra program cost with measurable achievement gains.

In summary, when the CSMP coordinator had a point of view that was similar to CSMP's , and held and continued to hold a position of responsibility in the district, the program was likely to survive in that district if funding continued to be available. In contrast, opportunistic adoptions, (where the reasons included "It sounded like a good idea" and "Money was available to do it so we did it") were likely to fade quickly.

Coordinator Concerns

During a Coordinator Roundtable at CEMREL in 1980, 26 coordinators completed a questionnaire in which they rated the likely effects of various potential problems associated with CSMP, both in their district and, hypothetically, in other districts.

Events that coordinators chose to define as "local" were easily the most critical factor for coordinators. Such events included changing school population, test requirements, lack of funds or the administration's lack of knowledge about CSMP.

Next in importance were the related issues of teacher training and change in teacher philosophy:

too great a change in teacher behavior or philosophy,
not enough time or authority to train/monitor teachers,
teacher training can't be done adequately,

followed by concerns about computation skills:

instruction on computation algorithms inadequate or too delayed,
lack of attention given to computation practice.

Least important were logistic matters of cost and organization of materials and lessons:

too much time needed for lesson presentation
organization of various materials too complicated in the schedule of lessons

Every issue on the list was rated by coordinators as more of a problem for other districts in general than for their own districts. Teacher training issues followed by computation concerns also topped that list and about half of the responses to the five statements listed earlier for these concerns were 4's or 5's, corresponding respectively to "High negative effect which is often decisive though sometimes possible to overcome" and "Decisive effect that causes rejection and is not possible to overcome".

Thus, one can assume that these coordinators believed that CSMP's teacher training requirement and low emphasis on computation skills would prevent the program from achieving widespread use generally, though they were rated as having only a "slight" or "moderate" effect in their own districts.

The main constraints in teacher training were time and money. In-service education is costly and the logistics of conducting in-service for special programs must compete with other school district priorities. Not only do teachers have to be paid for their in-service time, but that time has to be squeezed into (and often competes with) the district's plans for on-going in-service. Most districts allocate two or three days per year at most for in-service. During those days, all the in-service needs of teachers have to be met. Districts are often reluctant to release teachers from in-service sessions devoted to district needs in order to concentrate on special programs.

Another constraint in training was CSMP's uniqueness as a mathematics program as well as the complexity and sophistication of that mathematics. CSMP is unlike most of the mathematics that teachers learned in elementary school in pre-service training. For many teachers, the mathematics content and the distinctive languages were intimidating and contributed to teachers' reluctance to implement CSMP.

Several coordinators and teachers commented that a major drawback for CSMP was teachers' inability to see "what is going on". In their view, the workshops focus more than desired on individual lesson activities in the strands. Since many teachers have a restrictive definition of problem solving, thinking it to be only the heuristics involved in solving the usual word problems, merely calling CSMP a problem solving approach to mathematics did not help those teachers.

Regarding the computation problem, teachers, central staff, parents, and coordinators at all sites expected CSMP students to perform at least adequately on standardized tests, i.e., no decline in scores. Scores did decline occasionally on computation tests, though for the most part they stayed about the same or occasionally improved. But a result of "no change" generally did more harm than good, since some schoolboards and superintendents then had trouble justifying the increased training and material cost for CSMP. This effect was reduced in some cases where districts cooperated with CEMREL in conducting studies of student achievement using non-standard measures more appropriate to CSMP. CSMP students' improved learning on those tests persuaded some administrators to accept coordinators' claims about the program.

However, other administrators were not impressed. For them, the numbers that came back to them from their own standardized testing (for example, average percentile rank for each grade) determined their success or failure as administrators. This constricting influence of standardized tests, with its chain of accountability, public - schoolboard - superintendent - principal - teacher, places in jeopardy any program that deviates from the national curriculum.

Together, local and CSMP-related factors were constraints that most CSMP coordinators were able to overcome. They learned that a successful CSMP implementation was usually possible, but never automatic.

The CSMP Teacher and Classroom

Data in this section come from three sources. First, each year during the Extended Pilot Tests, CSMP teachers at certain grade levels were asked to respond to questionnaires. Altogether about 500 questionnaires were returned over the years. Proportionally more questionnaires were returned from the lower grades where the program has been available longer. The return rate was about 60% in the lower grades, higher in the upper grades. Second, about half that number of teachers were interviewed. The interviews were either extensive and wide-ranging when conducted locally, or briefer and more intense when conducted during a site visit to distant site. Third, teacher observations were conducted throughout the course of the evaluation. Locally they were much more extensive, the same teachers being visited frequently during the course of the year; in other sites they tended to be more frantic, a few minutes at a time. Teachers representing altogether about 40 school districts have been observed and interviewed.

Background and Experience

With two kinds of exceptions, CSMP teachers have been fairly typical elementary school teachers. Year after year, in comparative studies of student achievement, the responses of CSMP and Non-CSMP teachers were very similar in number of years of teaching experience, grade levels taught, and amount of preparation in mathematics.

One exception often occurred when a district first adopted CSMP and the coordinator had to develop an implementation strategy. A common way of doing this was to recruit a few kindergarten and first grade teachers from one or two schools. The presence in a school of particular teachers known for their excellence in teaching or for their openness to a CSMP-like instructional approach, was often a decisive factor in the selection of that school as a pilot school. Thus, during a district's first year or two of the program, CSMP teachers tended to be more able and open to new ideas. Later, as new teachers and grade levels started using CSMP, the overall composition of CSMP teachers in the school became more typical. Teachers at higher grades more or less inherited the program and their CSMP students, and the program became institutionalized.

The second exception occurred in some schools where the program was not monitored closely and was not officially mandated by the district as <u>the</u> mathematics program in the school. It therefore became fairly easy for teachers to avoid teaching CSMP if they wished. Many teachers began to teach it on a part-time basis and this led to one of two situations: either CSMP became voluntary, some teachers teaching it while others taught from the regular district textbook, (in which case the next grade's teachers would be faced with two groups of students: traditional and CSMP), or else teachers traded and a teacher who liked CSMP would also teach it to a colleague's class while the colleague reciprocated in a different subject. In either case, the CSMP teachers in those schools were <u>not</u> typical teachers; their teaching style and philosophy evidently agreed with CSMP. But this laissez-faire attitude usually led to the demise of the program in these schools.

Teacher Training

The training program developed for CSMP was designed to give coordinators and teachers a conceptual overview of the distinctive languages and content of CSMP as well as practical demonstrations and practice in teaching the lessons. The duration of the training was intended to be 8 hours for first grade, 16 hours for second grade, 24 hours for third grade, and 32 hours for fourth, fifth and sixth grades. CSMP recommended that all training be completed before school opened in the fall. These recommendations were seldom adhered to because of local constraints.

Sites had several options for training. Coordinators and teachers could attend sessions conducted annually at CEMREL. Alternatively CSMP staff members could sometimes visit a site and conduct training. A third option was the provision of a "turnkey" trainer who had been trained by CSMP staff, and was geographically proximate to the adopting site. The availability of a "turnkey" trainer was often a decisive factor in the adoption process.

It was the rare district that followed CSMP's specifications for training. From teacher survey data, between a quarter and a half of the teachers received less than 50% of the recommended number of training hours. Most teachers had no further training after they began teaching CSMP.

In several districts, teachers assumed a major training role by encouraging other teachers to observe their CSMP lessons, by conducting or assisting at district in-service days, and by arranging informal conferences within their buildings or across the district. At one site, a hot-line was established where teachers provided after-school hours assistance to their colleagues.

Although in most cases training did not meet CEMREL's specifications for intensity and duration, a majority of teachers surveyed thought they were adequately prepared to teach CSMP. Those teachers also said most other teachers in their schools could do an adequate job of teaching CSMP. Asked if they had any suggestions for improving the training, teachers made few suggestions for programmatic change but some recommended (not surprisingly) that the length of training be increased.

Where CSMP was most successful, teachers' involvement with CSMP has been a key factor. Surprisingly, length of training, intensity of training, and CSMP-conducted versus locally-conducted training played a relatively small role in this success and were not correlated very highly with student achievement. More important to success was the teachers' belief that they could learn the math, learn how to teach it, and that their students would profit from it. Thus, the skill of the trainer in imparting this confidence was very important. A willing group of teachers could overcome many in-service constraints. In fact, the program's impact on students made converts of many teachers who were initially reluctant. But teachers' resistance was not easily overcome and many adoptions foundered on that reluctance.

Daily Preparation and Materials Management

A common response in teacher interviews was that no amount of formal training could prepare someone for being a good CSMP teacher. Many teachers said, in effect, "You have to teach it for a year." This was meant in the dual sense of learning to teach it and learning to appreciate it. Day-to-day CSMP teaching was a relatively complex endeavor during the first teaching year. CSMP required daily planning according to a prescribed schedule, and access to two or three different volumes of Teacher's Guides during any single week. The teacher-led lessons took much longer than most teachers were accustomed to, often requiring 30 minutes or more and occupying seven or eight pages in the guide. Thus, to be successful, the teacher had to devote both time (for preparation) and energy (for the long lessons).

In comparing time required for daily CSMP preparation with time required for the previously taught mathematics curriculum, the most common response was "more at first but about the same after a year's experience". This response was given by at each grade level by between 50% and 60% of the teachers. The response "more at first and continues to be after a year's experience" was given by successively more teachers at each grade level, going from 9% of first grade teachers to 33% of sixth grade teachers. Fewer than 10% of the teachers reported that CSMP required less preparation time.

Logistics

The average amount of time reported by CSMP teachers for math class was about 45 minutes per day in grades 1 and 2, about 50 minutes in grades 3 and 4, and about 55 minutes in grades 5 and 6. Most teachers reported this amount of time to be longer than they previously took for math. It was also longer than reported by Non-CSMP teachers participating in the comparative studies of student achievement, grades 4-6. These Non-CSMP teachers reported spending an average of 3 to 8 minutes less per day depending on grade level.

Furthermore, lesson time was distributed in a different way. For CSMP teachers, nearly two-thirds of the time was spent in teacher-led activities; this was 50% more than Non-CSMP teachers reported. Conversely, CSMP teachers spent proportionally less time supervising and working with individual students or small groups. A sizeable proportion of CSMP teachers (nearly one-third) thought they spent too long in exclusively teacher-led instruction.

CSMP teachers spent an average of 20% of their math time supplementing the program with other activities. Most often this supplementation was in computation practice: the basic facts, whole number algorithms and, in the upper grades, practice with fraction and decimal operations. These items were most often cited (by one-third to one-half) of the teachers, as skills or concepts "that CSMP assumed students would know at the beginning of the year, which many did not know" or "that are not adequately covered by CSMP".

When similar questions were asked of Non-CSMP teachers, they reported spending virtually the same percentage of time supplementing, but this supplementing was much more diverse. Mental arithmetic, metrics, math labs and games, money, calculators, word problems and enrichment activities were most popular, but no single topic was listed by even one-third of the Non-CSMP teachers. These topics are often thought of as optional and done at the teacher's discretion.

The method of supplementing was also rather different. CSMP teachers tended to do it in very short stretches. The most common response to the question of when this supplementation occurred - "for a few minutes at a time" - was given by about half the teachers. Non-CSMP teachers' most common response was "for several consecutive math periods". This difference is compatible with the difference in what was supplemented, i.e., computation practice (CSMP) versus chunks of content that make for longer units of instruction (Non-CSMP). Teachers usually supplemented with teacher prepared or commercially available worksheets. Occasionally they assigned work from commercial textbooks that were in the school; frequent use of these textbooks was usually a sign of less than faithful implementation of CSMP.

Where do teachers find the time in the curriculum to spend an average of a day a week on these supplementary topics? In the case of Non-CSMP teachers, such topics may be part of the district curriculum but not in their textbooks. Also, it is not unusual for teachers generally to simply not cover the last one or two chapters in the text; such texts are written with this real possibility in mind and these chapters are not prerequisites for next year's work. CSMP teachers, on the other hand, did not consider their supplementation to be optional but there is little cushion in the CSMP schedule to allow for it. Hence, many CSMP teachers either omitted segments of the schedule or did not get through the schedule. In the upper grades, most CSMP teachers (75% - 90%) got pretty well to the end of the schedule but had to omit lessons to get that far. At the lower grades teachers were <u>less</u> likely to skip lessons but <u>more</u> likely not to get to the end of the schedule. For all CSMP teachers, the lessons most likely to be skipped dealt with probability and geometry, the content strands which are most different from the traditional curriculum, and least understood by teachers.

There were some other differences between CSMP and Non-CSMP classes. Student questionnaire data in fourth and fifth grades showed that CSMP students reported taking fewer tests and doing less homework; 10% - 20% fewer of them responded "a lot" to the questions about how often they did these tasks. CSMP teachers saw this as a weakness of the program; at every grade level, at least 70% of the teachers thought that periodic tests should be built into the curriculum for grading and general progress checks.

On the other hand about 25% <u>more</u> CSMP students reported that they played games a lot. These findings are unsurprising since the words "tests" and "homework" are virtually absent from the Teachers' Guides and many problems and lessons are presented in a game context. Also, high amounts of supplementation were associated with low amounts of game playing, i.e., supplementation <u>replaced</u> the game-playing part of the curriculum. For Non-CSMP, high supplementation was associated with high game playing, i.e., game-playing <u>was</u> supplementation.

CSMP and Non-CSMP teachers in fourth through sixth grade were asked to respond to pairs of statements about their math class. A five point scale was devised to show the relative balance between the two statements. The largest difference in mean scores between CSMP and Non-CSMP teachers occurred for a statement referring to lesson plans; CSMP teachers responded much more in the direction of "lesson plans are followed in great detail" versus "lesson plans serve only as a general guide". On two other pairs of statements, out of a total of eight, there was about a half point difference in responses. CSMP teachers were more likely to say that math class had a fun (versus businesslike) atmosphere, and that math class was oriented towards creative activities (versus solving specific problems).

Of particular concern for a curriculum like CSMP is the potential problem of new students transferring into the program. These students must become familiar with the special CSMP pictorial representations before they can even follow the lessons. This problem appeared to be most serious at the second and third grade levels, particularly with the minicomputer. But regardless of grade level, the number of new students, the time of year they entered, and their general ability level determined how big a problem they posed for the teacher. One or two new students of low ability or several of high ability could usually be brought into the program at the beginning of the year in a variety of ways. The spiral nature of the curriculum was undoubtedly helpful in many cases since students didn't have to master the content of one lesson in order to benefit from the next lesson dealing with that content.

However, when there were several low ability students and/or students entered periodically during the school year, teachers reported having problems. Test data showed that new students in general performed almost as well as veteran CSMP students of similar ability levels. Nevertheless, teachers' perceptions of the problem may have been a factor in some teachers' opinions that CSMP was not appropriate for low ability students. Also, it was probably a factor in a few schools where CSMP evolved into a program for upper track students. As new students entered those schools, the slower ones were sometimes targeted to the teachers who used CSMP on a more limited basis, thus accelerating the split between CSMP and Non-CSMP classes within a building.

Teacher Opinions about CSMP

For several years teachers at various grade levels were asked to compare CSMP with the mathematics curriculum they had previously used. The rate of return of these questionnaires was about 50%-60% in the lower grades; higher in grades 4-6. Mean scores were calculated at each grade level by assigning a score of 1 to the lowest rating ("much worse" than previous curriculum) and 5 to the highest rating ("much better"). Ratings are summarized below in Table 6.

Table 6
Mean Score by Grade,
Teachers' Comparison of CSMP to Previously-Used Curriculum

Grade Level	K	1	2	3	4	5	6
(N)	90	110	92	118	69	43	2.2
Overall quality	4.4	4.6	4.5	3.9	4.0	3.7	4.4
Student interest	4.5	4.8	4.2	3.9	4.0	3.9	4.4
Students' logical reasoning ability	NA	NA	4.3	4.0	4.3	4.4	4.6
Appropriateness for high ability students	NA	NA	NA	4.4	4.4	4.6	4.9
Students' facility with word problems	NA	NA	3.4	3.2	3.4	3.3	3.7
Student achievement in mathematical concepts				3.7	3.8	3.8	4.4
	4.4[1]	4.3[1]	4.0[1]				
Student achievement in computation skills				2.7	2.7	2.7	3.1
Appropriateness for low ability students	3.0	2.6	2.8	2.6	2.5	2.5	2.3

[1]In grades K-2, there was only a single item, "Students overall achievement".

The highest ratings were given for items 1 to 4, dealing with overall quality, student interest, logical reasoning, and appropriateness for high ability students. Each was rated, on average, between "better" and "much better" than previous curriculum. The lowest ratings were given in the last two items, dealing with computation skills and appropriateness for low ability students. Both were generally rated slightly worse than for their previous math program. Achievement in computation skills was rated at least a full point lower than achievement in mathematical concepts in grades 3-6.

The question regarding appropriateness of CSMP for low ability students drew the widest range of scores; there were relatively few "about the same" responses and many extreme responses, both positive and negative. For example, among fifth grade teachers, 55% of the teachers thought CSMP was less appropriate, but nearly 30% thought CSMP was more appropriate! It was not the case that low ratings came primarily from CSMP teachers who had many low ability students; if anything they came more from teachers with few low ability students. Non-CSMP teachers, however, were much more likely to rate their curriculum low on this criteria if they had many low ability students.

Teachers in grades K-2 gave more positive responses to CSMP than did teachers in grades 3-5, each grade level of which produced almost identical responses. The general increase in scores at sixth grade is probably because that group of teachers was small and happened to be teaching relatively higher ability students.

Fourth through sixth grade results were based on many fewer teachers. This was partly because fewer classes had reached those grades, and partly because questionnaires in some years were collected only from teachers of classes participating in a comparison of student achievement. Both CSMP and Non-CSMP participating teachers responded and their responses can be compared in Table 7. For CSMP, these responses are a subset of the responses from the previous table; they are not appreciably different from those of the larger group.

Table 7
CSMP and Non-CSMP Teacher
Comparing Present program to Previously Used Program

	Fourth Grade		Fifth Grade		Sixth Grade	
	CSMP	Non-CSMP	CSMP	Non-CSMP	CSMP	Non-CSMP
(N)	30	21	30	23	22	26
Overall quality	4.0	3.3	4.0	3.8	4.4	3.7
Student interest and involvement	4.1	3.4	4.0	3.4	4.4	3.8
Students' logical reasoning ability	4.4	3.0	4.6	2.9	4.6	2.8
Appropriateness for high ability students	4.4	3.4	4.6	3.9	4.9	3.4
Students' facility with word problems	3.4	2.5	3.3	3.0	3.7	3.2
Student achievement in mathematical concepts	3.6	3.3	3.9	3.4	4.4	3.5
Student achievement in computation skills	2.9	3.0	2.9	3.5	3.1	3.5
Appropriateness for low ability students	2.7	2.6	2.4	2.8	2.3	3.1

CSMP teachers gave higher ratings than Non-CSMP teachers on seven of the nine items. The average difference was between 1/2 and 1 point on five items and over 1 1/2 points on two items dealing with logical thinking and appropriateness for high ability students.

Non-CSMP teachers gave higher rankings on two of the nine items, those dealing with achievement in computation skills and appropriateness for low ability students. The average difference was less than 1/2 point. The CSMP discrepancy in teachers' perceptions of student achievement in computation versus concepts did not appear with Non-CSMP teachers, who rated them equally. Appropriateness for low ability students usually was rated lower by CSMP teachers, but Non-CSMP teachers also did not give their curriculum high ratings on this item.

When responding to questions about the most effective way to teach low ability students, CSMP and Non-CSMP teachers gave virtually identical responses to 7 out of 8 questions. The only difference between the two groups was that CSMP teachers were more likely to say that best learning takes place when a teacher can give individual help versus working with small groups. Special provisions for low ability students were reported to be available by 85% of both the CSMP and Non-CSMP teachers, and were usually provided through a resource teacher or room.

When teachers were asked to describe their overall evaluation of CSMP, responses could be fairly easily divided into three groups. About 65% of the teachers in grades K-2, and about 40% of teachers in grades 3-6, gave an unqualified positive response to the program, often describing it in glowing terms. At the other extreme, a steady 10-15% of the teachers' were thoroughly negative towards the program. The remaining teachers responses can best be described as positive but qualified, such as "I like the program overall but..." About half of these reservations dealt with minor issues or were not considered serious by the teachers, but two familiar issues were raised most frequently year after year and were of considerable concern to many teachers: the lack of attention in CSMP to the basics -basic arithmetic facts and the arithmetic algorithms - and the perceived difficulty of the program for low ability students.

Similarly, when asked to name the worst aspects of CSMP, teachers most often alluded to these two concerns. Non-CSMP teachers, however, thought coverage of the basics to be a positive aspect of their program. In naming best aspects, CSMP teachers almost always named thinking skills (problem solving, mental work, creativity, reasoning, challenging, etc.) or motivation/interest; these two areas were more likely to be named by Non-CSMP teachers as worst aspects of their programs.

Next most frequently named complaints by CSMP teachers were that lessons were too abstract, that too much of the lesson was teacher-directed, and that students did not have the prerequisite skills needed for some lessons.

One area in which CSMP teachers' opinions changed dramatically by grade level concerns the spiral approach. In giving free responses to a question about the spiral curriculum, 74% of first grade teachers were very positive and only 10% negative. These figures changed monotonically by grade level until at fifth grade there were 30% very positive and 30% negative; the other 40% expressed qualified approval.

Fifth and sixth grade teachers were asked to respond to a series of statements about the spiral approach. Three statements produced strong nearly unambiguous approval of CSMP: teachers agreed that the spiral approach was more interesting and students felt less pressured than in a mastery approach, and teachers did not agree that students never master the content. However, on four other statements, about half the teachers gave responses that were negative towards CSMP: teachers had to repeat a lesson because students didn't remember, the spiral approach only worked for some students, too much time elapsed before the class returned to a topic, and 2-4 consecutive days on a new topic would be preferable to the current schedule. These statements also appear in free response evaluations of the spiral approach and in teacher interviews, though less frequently in the lower grades.

CSMP, The Low Ability Student, and Computation

The most common complaints about CSMP are its perceived inappropriateness for low ability students and its lack of attention to developing the basic computational skills. These two complaints surfaced at all levels - teachers, principals, coordinators, central office staff, school boards and parents. No school was without at least one or two teachers who disliked the program for those reasons. In the upper grades the program is being used disaproportionately more often by districts or classes with higher ability students.

To what extent are these complaints justified? Data presented in the next chapter will show that CSMP students perform about as well as Non-CSMP students on computation tests and that CSMP low ability students perform nearly as well as CSMP students at other ability levels vis-as-vis their Non-CSMP counterparts. On the other hand, there are occasional instances of weaknesses in these areas. In the large Extended Pilot Tests of fifth and sixth grade classes, for example, the lowest ability CSMP districts happened to perform poorly compared to Non-CSMP districts of similar ability. When data were analyzed at the student level, low ability CSMP students as a group fared worst in comparison to Non-CSMP students in computationally oriented tests. CSMP classes whose teachers supplemented the program least, and who most agreed with the CSMP philosophy, tended to have the lowest computation scores. But the few findings of this nature are overwhelmed by most other findings. The data do not support the intensity felt by some teachers over these issues. It is worth considering why teachers felt this way, given the overall data on low ability students' success.

The computation issue seems the more straightforward of the two issues. Even a cursory review of the CSMP materials reveals that there is less computation practice of the paper and pencil, drill and practice variety. It is not likely that this difference is entirely compensated for in the teacher-led lessons, certainly not when it comes to the multiple-digit algorithms. Very few teachers rated CSMP better than their previously used math programs in student achievement of computational skills; most rated it a little lower. Teachers did supplement to the extent they thought necessary and this supplementation seemed to help.

Some CSMP users approved and supported this supplementation and did not feel it to be a particularly black mark against the program. Teachers generally know how to teach computation skills. They were able to fit the supplementation in with short bursts during class or as homework, had lots of practice materials around, and could easily check student skills. But many teachers were encouraged not to supplement by coordinators and by the Teacher's Guides whose spiral philosophy downplays the need for supplementation.

Regardless of whether this supplementation was done *surreptitiously or with approval*, it required additional time in an already crowded schedule. In some districts this was recognized and taken into account but usually the additional time burden fell squarely on the teachers' shoulders. Thus, this perceived weakness probably does exist, can be ameliorated fairly easily, and at a cost which seems high to some teachers and low to others depending on their view, and their district's view, about priorities in mathematics education.

The issue of appropriateness for low ability students is more complicated. Substantial though smaller numbers of teachers felt that CSMP was more appropriate for low ability students. In questionnaire and interview responses, many teachers said the program had positive effects on low ability students: "...seems like students working at all cognitive levels get something out of", "there's something for that child who isn't quite as fast...can still participate and be right and (the program) clues me into what they're thinking". Given these teachers' views and the generally positive test data, it is worth considering why so many teachers did not like this aspect of the program. A few reasons are offered here.

For many teachers, the issue was tied to the computation issue. They had some doubts about whether parts of CSMP, especially geometry and probability, really taught mathematics, and whether these areas had any practical value. They held these views even more strongly for low ability students, whose primary educational need was seen to be adequate computational skills. Higher ability students might or might not learn problem solving skills but one way or another would pick up the necessary computation skills. Low ability students could not be expected to learn many problem solving skills and without the teacher's help they also wouldn't develop adequate computation skills.

Teachers of higher ability classes, with only a few low ability students, were more likely to think CSMP inappropriate for low ability students than teachers of lower ability classes, with many low ability students. The gap in achievement seemed to widen for some teachers of high ability classes. This may be because CSMP gives the teachers many opportunities to see their children working at genuine problems and responding in class to difficult questions. Clearly some students show abilities that were previously masked in the traditional computationally oriented program. The three and four-star workbooks contain some genuinely challenging material which some students gobble up while others never even see. There are probably more occasions than formerly for good students "get it" and become enthused while the slower students appear lost.

Thus, even though low ability students may have benefited from CSMP (as test data suggest), teachers' day-to-day experiences suggested to them that these students were getting farther and farther behind. The CSMP curriculum does not contain progress tests, but teachers could easily check their students' computational skills against their own well-developed, experienced-based standards and find the program lacking. They did not have an easy way to measure students' thinking skills, nor a standard against which to compare it, so could not see any compensating gains.

Many teachers stated that the spiral approach didn't work for low ability students and that they had to reteach crucial parts of a previous lesson because students didn't remember from the last time. This led some teachers, sometimes with district support, to regroup lessons and teach several related lessons in a block, contrary to the recommended schedule of lessons. Observations suggest teachers may have been right in some instances, but it was sometimes hard for anyone to determine which elements of a previous lesson really were crucial. It was also difficult at times to predict whether or not students would somehow muddle through the new lesson in spite of only a hazy remembrance of the previous lesson.

CSMP places heavy emphasis on the "guided discovery" approach. This means asking questions that students haven't heard asked before, let alone know the answer to. The ratio of questions that students can readily answer to the total number of questions asked in a lesson is probably much lower in CSMP than in traditional programs. So teachers see many more instances then they are used to of low ability students not being able to answer a question.

Lower ability students who transferred into a CSMP classroom were faced with special catch-up problems because they had to learn the special CSMP representational languages. Again, test data indicate they did catch up but this undoubtedly requires special efforts by teachers which would not be necessary in a traditional program. The spiral approach, though helpful in this regard, may also stretch out this catch-up process.

In summary, some teachers' day-to-day experiences suggested to them that the program didn't work well with low ability students and this conclusion was not altered by abstract test data. This opinion was reinforced if they did not share CSMP views on decreased computation emphasis, the spiral scheduling approach, and guided discovery lessons. Most adapted the program in sensible ways to remediate this problem, and the adaptations may often have been warranted. Some made such extreme changes that the program became very different and gradually ceased to be taught.

Teacher Observations

Teachers at over 40 sites have been observed teaching CSMP. Most lessons observed followed the intended lesson in the Teacher's Guide at some level of correspondence, but there was wide variation in how faithfully, and how well, the lessons were taught. This variation did not seem to be related to objective factors such as size and ability of class, district circumstances, teacher experience and background, etc. It had more to do with teachers' general teaching skills and their understanding of CSMP.

General Teaching Skills. Most teachers had at least adequate classroom management skills; students were reasonably quiet and attended to the lesson, teacher and students could be heard, work was assigned and the assignment understood, materials were at hand for use. A minority of teachers, perhaps 10%, had management problems that were enough to disrupt the lesson seriously - sometimes temporarily, sometimes for the duration of the lesson. These problems had nothing to do with CSMP and no doubt affected learning in all subject areas.

But CSMP placed an added burden on poor managers because of the many student and teacher materials, the complicated schedules, the long lessons and the lack of closure (objectives) inherent in CSMP's spiral approach. It may be that such teachers could cope better with a very traditional program involving, say, 15 minutes of lecturing followed by 25 minutes of drill and practice in a very circumscribed, computationally oriented curriculum. In either case, the students would have to take on a larger burden of the learning for themselves; higher ability students can do so, lower ability students cannot.

In addition to having basic management skills, most teachers also had reasonably good expository skills, usually adequate for explaining the mathematical concepts and skills in CSMP, provided they themselves understood them. The teaching skills that were most important in CSMP had to do with asking questions and coping with what might be called CSMP's "guided discovery" lessons. Question-asking techniques needed for student learning include the following:

> asking for several answers to a question and asking "why" or "why not" questions,
>
> basing the next question on an evaluation of the previous response,
>
> waiting a few seconds after asking the question before naming the respondent,
>
> distributing questions widely,
>
> matching questions with ability of the respondent,
>
> following up on the consequences of an answer,
>
> when necessary, asking the next easiest question or a related question that has been previously answered.

Any good teacher should possess these questioning skills. But their crucial importance in determining how successfully CSMP is implemented in the classroom lies in the extent to which the program demands and relies on them. The "pedagogy of situations" is in some ways a problem solving approach, and the list of question-asking techniques given above contains many that are necessary for any good problem solver. One reason problem solving is not taught often or well is that these are not easy techniques to learn. For example, in developing lessons, some teachers shortened the lesson to what was virtually, "Here is the rule. Now apply it." Although the lessons in the Teacher's Guides are full of suggested sequences of questions and possible responses, they can never be more than guides. Following the guide slavishly created as many problems for teachers as straying too far from it did.

The vast majority of teachers handled some of these question-asking techniques well, others not so well. Perhaps the hardest to achieve was responding effectively to an incorrect answer when that answer should have provided a tip-off about an important misunderstanding of a concept. For many teachers it was clear that CSMP was their first experience in a curriculum which explicitly required these techniques and they were making a genuine effort to use them according to the Teacher's Guide.

It is this fact which prompts many coordinators to think that the real strength of CSMP is in the teacher training it provides through the Teacher's Guide. Visible improvement could be seen in some teachers after a year's experience; they became better question-askers. It is unfortunate that most did not receive the kind of intensive in-classroom support from coordinators that would build these skills faster.

A related issue of critical importance was the way teachers incorporated CSMP's guided discovery approach. Decisions had to be made throughout the lesson about how long to wait for an answer (or try for the correct answer), how much to explain, how many questions to ask, etc. Though there is general agreement on what the good question-asking techniques are (observers know them when they see them), the effectiveness of the best kind of discovery approach has always been a source of disagreement among educators. When observing CSMP lessons it was most often the pace of the lesson that had the greatest impact on the observer.

There was wide variation in how quickly the lesson moved along. For a given lesson which might have an intended development time of, say, 25 minutes, about 20% of the teachers might do it in 15 minutes while about 35% would require at least 40 minutes. Some of the variation in pace was related to the overall ability level of the class, but most was due to teacher differences. Probably more teachers erred on the side of too slow a pace than too fast. Some teachers slowed down when computation was required and then speeded up during the problem solving part of the lesson. Certainly the most effective lessons were those with a crisp pace controlled by clever questioning and supported by thorough preparation and understanding of the lesson. The most painful to watch were the ones which dragged interminably as teachers belabored unimportant points or repeated unnecessary examples.

This difficulty in judging pace is understandable given the nature of most CSMP lessons. Because many different mathematical ideas are touched on in most lessons, there is often no single focal point for the teacher to concentrate on by skipping parts or adding other parts. In most cases of substantial deviation from the lesson plan, the resulting lesson was less effective than the original. Compounding the problem was the natural, and perhaps justified, reluctance to zoom on to the next part of the lesson, while students were still having difficulties. In some cases it would have done no harm because of the nature of the lesson since the developer may have expected some students to get more out of it than others, or the concept was to be developed more fully later. But in other cases, that part of the lesson was truly a prerequisite for understanding what would come next. Only a thorough understanding of the lesson, and other lessons in the sequence, could enable the teacher to make an accurate decision about when to stop and regroup and when to move on.

Overall, lessons took longer than intended by the developers. A single long lesson might be split into two lessons by the teacher. An additional lesson might be prepared by the teacher for consolidation or as a worksheet assignment because the whole previous math period was needed for the teacher-led part of the lesson. This lengthening of lessons, in an already full yearly schedule (with occasional time taken for supplementation), caused many teachers not to complete the schedule or to drop segments of the schedule that they considered to be too hard or too much off the main track, such as geometry and probability. Again, this happened more often in lower ability classes.

On the whole, most teachers did a fairly good job of pacing their lessons and learned to improve with experience. The teachers who had the most difficulty in maintaining pace were teachers who were naturally inclined towards mastery approach, but who nevertheless attempted to teach the lessons according to the guides. At the other extreme were a few teachers who preferred a directed teaching approach, changed the spirit of the lesson to fit this preference, and thereby did most of the thinking for the students.

Teachers' Understanding of CSMP. There were three ways in which teachers' understanding of the program played an important role in the quality of the lessons observed. At the lowest level was simply being prepared for the lesson: knowing in advance what the sequence of activities was, preparing needed blackboard demonstrations, having other student or teacher materials available, having some idea of the way questions would be asked, and knowing how long to devote to various portions of the lessons. This is a fairly onerous job for first year CSMP teachers since many of the lessons run eight pages or more in the Teacher's Guide. It was not uncommon for teachers to have the Guide firmly in hand throughout the lesson. Some teachers had obviously done little preparation and this contributed to sense of floundering, long pauses and eventual loss of interest by students, a generally vicious circle that made lessons very long. Other teachers were superbly prepared and in full control. Most fell somewhere in between. Gradually, dependence on the Guide decreased with time but even for experienced teachers it was rare not to see the Guide opened at the right page and handy for occasional reference.

The next level of teacher understanding was the content: how to solve the problems, know the good strategies for playing the games, know why some answers are good and others poor, and know all this well enough to respond rapidly to classroom situations. Long pauses while the teacher figured out an answer almost always disrupted the smooth flow of the lesson. It was at this level that the more mathematically able teachers were at an advantage, but even for less able or interested in mathematics such problems could often be tied to inadequate preparation, i.e., not actually going through the various problems and situations and thinking about them as they did so. Wrong answers were given by teachers on occasion, or they accepted an incorrect answer from the student. Because of the potential damage of such errors, this possibility became a source of tension for some teachers and they became flustered.

In other classes, students were obviously used to this happening occasionally and corrected the teacher who made a matter-of-fact adjustment and continued with the lesson. In many ways this response fostered a very healthy and cooperative atmosphere for learning. In defense of the teachers, it must be said that because the CSMP materials are so rich and layered with many levels of mathematical thinking, the curriculum is replete with situations amenable to teacher blunders or long pauses. Such errors have been observed in classes taught by CSMP development staff. Most teachers were somewhat apprehensive about the CSMP content when they first began teaching the curriculum, and this was especially true of teachers at the upper grade levels. But with experience and conscientious preparation, they were observed (and reported themselves) to have improved dramatically.

The highest level of CSMP understanding, and the most difficult to attain, was an understanding of why things were done the way they were, i.e., the purpose behind the various lessons and exercises. There are many general statements in the Teacher's Guides about the various mathematical aspects of the lessons, and about the problem solving and higher order skills being emphasized. But these are not described anywhere in detail, or in behavioral terms, nor are they categorized or referenced. It was often difficult for the teacher to know where a lesson was going or why a particular sequence of lessons appeared in the curriculum. The lack of understanding about, and in some cases disagreement with, the philosophy and goals of the program occasionally affected teachers' attitudes towards the program and their subsequent performance in the classroom.

This attitudinal problem was likely to get worse rather than better with experience. Some teachers came to see the program as having an excessive commitment to nebulous kinds of unmeasurable thinking skills resulting in a weak development of the familiar skills and concepts that teachers approve of and know how to teach. Among the ways in which this attitude manifested itself in the classroom were the following: an impatience in getting to the point of lesson, a fixation on getting the correct answer, a need to see observable progress in students' performance, subtle to drastic changes in lessons and sequences of lessons, an increased emphasis on student written work, limited expectations of what students are capable of doing, and sharply defined expectation of mastery of certain skills at certain times.

Summary of Teacher Observations. In summary, teachers who had good generalized teaching skills, who were willing to prepare adequately in order to learn the content and lessons of the program, and who understood and agreed with the philosophy of the program, were able to do an outstanding job in the classroom. Many memorable lessons were observed which cried for a wider audience to see the power of CSMP in the right hands. But this combination was hardly the norm; more commonly observed were lessons presented in a fairly competent way by teachers doing the best they could with a difficult curriculum. They usually got better with experience and the highs generally outnumbered the lows. For a significant minority of teachers, several pieces of the combination of factors listed above were absent and the teaching of CSMP moved inexorably towards the more traditional approach.

Summary

CSMP has been successfully implemented in many different kinds of school districts with many different kinds of students. Through 1982, 134 school districts had used the program and as of 1984, approximately 55,000 students were using CSMP. The program tends to be used less often in grades 4-6 than in grades K-3. There is also a trend toward usage by higher ability classes in the upper grades.

In any given year recently, over 90% of the districts using CSMP one year continued to use it the following year. The curriculum is still healthy in spite of virtually non-existent support for dissemination from NIE since late 1982.

The role of the local coordinator has been vital to the success of CSMP; without a skilled and influential person at the helm, a solid implementation was not likely. Coordinators from outside the district (such as a local University professor), or with single-school responsibilities (such as a principal or teacher), were much less successful than coordinators with district wide responsibilities (such as a mathematics supervisor).

The coordinators' biggest concern, and most difficult job, was training teachers for CSMP. Teachers and/or financial support were not always available to the extent necessary to meet the CSMP recommendations for training (from two to five days depending on grade level). Consequently, at least half the teachers received much less than the recommended amount of training. This job got harder as more classes used CSMP, at higher grade levels, and as new teachers entered the system.

Another constraint on the use of CSMP was the cost of the program, which tended to be competitive with traditional programs in start-up costs but more expensive to maintain, particularly in grades 4-6 where consumables needed to be purchased each year.

Teachers who had good general teaching skills, who were willing to spend the time in training and daily preparation, and who agreed with CSMP's overall philosophy, were able to do an outstanding job of teaching the program. The absence of any one of these three attributes - skills, commitment and philosophical agreement - reduced the program's impact in the classroom, and it came to look more like the traditional mathematics curriculum. But in any case, most teachers supplemented the program with computation practice and dropped portions of the curriculum, especially lessons in geometry and probability.

Questionnaire data from a large number of CSMP teachers, showed that teachers rated CSMP higher than the previous curriculum they had used, and higher than Non-CSMP teachers rated their curriculum, in:

overall quality,

student interest and involvement,

students' logical reasoning ability,

appropriateness for high ability students, and

student achievement in mathematical concepts.

On the other hand, teachers rated CSMP less appropriate for low ability students, and less effective in teaching computation skills, than the previous curriculum they had used.

In summary, although CSMP is a difficult program to implement, but it can and has been implemented successfully for several years in many different settings.

The previous chapter concluded with a list of CSMP features that make it a distinctive curriculum, and suggested why such features should make it a desirable program. The features will be reviewed here, and it will be shown that each of them is a double edged sword with equal potential for making it an undesirable curriculum.

1. CSMP contains recommended new content.

   The content is also new to teachers, most of whom have very little formal mathematics background and do not understand why such content is needed. They resist it and it is the first thing to be dropped in a time crunch.

2. CSMP resequences certain arithmetic skills and slows their rote development to ensure understanding.

   Traditional wisdom holds that students should master certain skills in certain grades: addition-algorithm in second grade, subtraction in third, basic multiplication and division facts in third, etc. There is pressure to continue this timetable because of test standards, student mobility, parent expectations and some teachers' belief that this is the way the world is and should remain.

3. CSMP promotes higher order thinking skills by presenting rich mathematical situations. Such situations do not usually culminate in a specific target for mastery, but instead emphasize the process of getting there. Each lesson may have several objectives but none has to be achieved for the lesson to be successful.

   This organization contradicts much current educational practice which emphasizes an instructional process of stating objectives, providing instruction to meet those objectives, measuring student outcomes, and basing next instruction on the results of this measurement. Teachers see games of strategy as frills, rather than as a way to learn thinking skills.

4. CSMP lessons extend the length of time teachers engage the whole class.

   This extension requires more preparation by teachers and is physically demanding. Teachers have less time to work individually or with groups of students.

5. CSMP has developed an extensive training program and training materials to help teachers use the curriculum successfully.

   Inservice training is difficult for most districts because of the cost and extent of training, the time required for teachers to participate and the need for skilled trainers.

6. CSMP's schedule of lessons incorporate the spiral approach.

   The lack of specific behavioral objectives flies in the face of current mastery teaching which generally prevents students from progressing to a new topic until they have learned the old one. Teachers feel uncomfortable when topics are left uncompleted and when students don't remember everything from the last time a topic was covered.

7. CSMP uses representational languages which are mathematically potent and reduces the verbal load on students.

   These languages take time for the teachers to learn, require catch-up time for new students and are difficult to explain to parents and administrators. Sending work home sometimes creates problems with parents.

8. CSMP reduces the time spent on rote development of computational skills.

   Most teachers have, over the years, developed good methods for teaching these skills. Since the skills are easily measured and hold a dominant position in standardized achievement tests, they have gained acceptance as the "real" mathematics content for students. There is increasing pressure on schools to be held accountable for student performance (for example, through state mandated criterion-referenced testing programs). Teachers believe these skills are the one outcome that all students must achieve.

9. CSMP provides extensive Teacher's Guides with detailed lesson plans.

   Teachers need to put in more preparation time. Some teachers think that the guides are overly prescriptive.

10. Student materials are attractive, high quality and easy for students to use.

    Because they are consumable, new materials need to be bought each year. This makes the program more expensive in the upper grades than traditional textbook programs where the text can be reused for several years. Moreover, since student materials are not in textbook form, schools sometimes can't use regular textbook funds to buy them and it is difficult to get the program on state-approved textbook tests.

The traditional mathematics curriculum, used virtually nationwide, is relatively robust. It can simultaneously withstand many different kinds of criticism because of its low cost, its easy-to-measure goals, its familiarity to all teachers and its established position. CSMP, on the other hand, is relatively fragile; any single one of the many problems described above can scuttle an implementation.

Sweeping changes on so many fronts at the same time, as CSMP attempted, are bound to be resisted. One need only look at the discrepancy described in the NCTM Prism survey between math supervisors, teacher trainers, and researchers on the one hand and principals, school board members, and the public on the other, to know that the first group - the mathematical experts - has limited power to change the views and practices of the second group.

It may also be the case that CSMP is viewed, even by some educators who agree with the reasoning behind its approach, as a somewhat eccentric program. A single, consistent philosophy and way of doing things are omnipresent; one could not call CSMP eclectic. Perhaps the point of view that sparked development, also prevented a practical accommodation to the exigencies of marketing and implementation. Or perhaps the creative single-mindedness necessary to produce a program of this scope and consistency is incompatible with such an accommodation.

IV. STUDENT ACHIEVEMENT: TOTAL MANS SCORES

Overview

The ultimate question to be answered in the evaluation of any curriculum is "How are students' knowledge and skills different, as a result of their participation?" Answering this question with respect to CSMP presents some interesting problems for assessment. Goals are given only at the most general level, such as "dynamic creativity." In the spiral approach, content is interwoven at successively more complicated levels, but expectations of mastery levels at any point in the curriculum are absent. Topics in which certain mathematical ideas or processes are used may disappear after brief usage. There is a continual interchange between content and process. And most difficult of all, the special CSMP languages are the vehicles in which almost everything takes place: concept development, applications and problem solving.

The main vehicle for the evaluation of student learning was the MANS Tests, Mathematics Applied to Novel Situations, a series of short tests, different at each grade level, developed by the evaluation staff. The tests probed important mathematical processes, such as relational thinking and estimation, by presenting students with generally unfamiliar mathematical situations that did <u>not</u> use any of the special CSMP terminology. The tests were administered to large numbers of CSMP and Non-CSMP classes in grades 2-6. This chapter will describe the MANS tests and present student data. The next chapter will describe student performance on each of the MANS categories.

On the other hand, CSMP is an elementary school curriculum which is intended to be <u>the</u> mathematics program for schools which adopt it. Thus, users have an expectation that the program will provide students with the knowledge and skills that are generally expected at these grade levels, regardless of the intentions of the program developers. In order to investigate this concern, a wide variety of standardized tests was used over the course of the evaluation. Because of the concerns expressed by many teachers about inadequate computational skills of CSMP students, this part of the evaluation came to focus on the computation sections of standardized tests. The results of these test administrations will be described in Chapter VI.

Testing was carried out in two ways. The main source of data for this report was from tests administered during the Extended Pilot Test for each grade level of the CSMP materials. These Pilot Tests were initiated by CEMREL, with school districts cooperating as part of their participation with CSMP. A secondary source of data for this report was a series of Joint Research Studies, initiated by local districts and carried out cooperatively between CEMREL and a local district on an individual basis. These Joint Research Studies took place after the Extended Pilot Test and involved revised versions of both the curriculum and the MANS.

In both kinds of studies, the designs were comparative in nature, with the performance of CSMP classes compared with that of Non-CSMP classes. The method of analysis was an Analysis of Covariance on class means, with class score on a reading or vocabulary test used as a covariate.

The MANS Tests

Description

The MANS Tests were the principal measures of student outcome used in this evaluation. They are a collection of short tests, designed to assess how well students can use mathematical thinking and skills in situations that are new or unfamiliar to them. The tests are in plain English and do not use terminology that is specific to any particular curriculum, including CSMP.

The MANS Tests are normally contained in two student booklets at each grade level, each of which requires a period of 30-60 minutes (depending on grade level) for administration. Each booklet contains several tests. Every test has its own directions which a specially trained tester follows in explaining the task and describing sample items after which students then complete the items in that test on their own. A flexible time limit, typically about 5 or 6 minutes, allows almost all students to finish. Most tests contain 5-9 items.

Each MANS test takes up one or two pages in a booklet so that diagrams and illustrations are large, words are easy to read and there is ample space for students to do scratch work. For most tests, students produce their own answers instead of selecting one of several given alternatives. Answers are to be written in the booklet and can be erased or crossed out; no special pencil is required.

At each grade level, one of the tests is a standardized vocabulary test, whose purpose is to derive an estimate of the ability level of each class which can then be taken into account in subsequent analyses of covariance.

A simple version of item sampling is used for most tests by having two versions of each test booklet. Each version looks the same at first glance; pagination, sample items and format are identical but the actually test items are different. The two sets of test items are similar in general difficulty but are not necessarily statistically parallel. The class mean is the main level of analysis for the MANS Tests. Therefore, having a random half of the class take each version of the booklet allows class means for a test to be based on twice as many items without extending the testing time.

The MANS tests are different in each grade level (grades 2-6). Although some kinds of tests may be repeated from one grade to the next, with some overlapping of items, the tests are always somewhat different at each grade.

The tests are classified into categories based on mathematical process or content. There are seven process categories, each of which is represented by at least one test at each grade level. In addition, there are five special topic categories which are introduced at the upper grade levels. Appendix G describes each of the 57 MANS tests, grouped according to category. Each description includes an abstract of the test, how it is administered, and some sample items.

A brief description of each process category is given below, together with items from some of the MANS second, fourth and sixth grade tests. Many of the items shown have been much abbreviated from the versions seen by students, but the set of items for each category will give the reader a better operational understanding of what the categories mean.

Computation. Straightforward calculation with basic fact and algorithms. Standardized achievement tests of computation were sometimes used to assess this category. A description of the tests, and the subsequent results, will be delayed until the next chapter.

Estimation. Rapid calculation of approximate answers under short time limits. Most tests were made up of multiple choice items. A typical test contained eight items to be answered in 1 1/2 minutes, with suitable warnings to students not to calculate exact answers and with frequent announcements of how much time was left.

Sample items

Second Grade	Fourth Grade	Sixth Grade
90 - 12 is in which interval 0 - 10 - 50 - 100 - 500?	602 is about ? as large as 298? 2, 5, or 10 times	8 1/2 - 8 is: <1 or =1 or >1? Which interval contains 1,002.5 ÷ 21.5? 0 - 1 - 10 - 20 - 50 - 100?

Mental Arithmetic. Exact computation of problems amenable to non-algorithmic solution. The computation aspect of the problems was downplayed; numbers were either small or easy to work with (such as multiples of 25, 50 or 100). Scratch work was not usually allowed.

Sample items

300 - ? = 250

Hit = gain 5, miss = lose 1
start with : 3 below zero
end with : 5 above zero
# of misses : 2
# of hits : ?

12 x 75 = 900
13 x 75 = ?

1/2 x ? = 40
0.75 - 0.5 = ?

scratch work not allowed

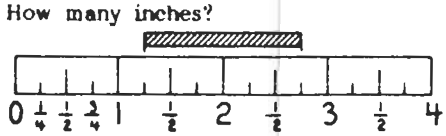
Number Representations. Recognition or production of ways of representing numbers. In the primary grades, the tests were concerned with whole numbers and place value; in the upper grades, fractions and decimals were emphasized.

Sample items

Write "two thousand, eleven"

100 more than 901 is ?

How many inches?

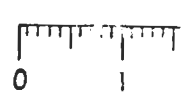


Name the 2nd largest 4-digit number using only 2, 5, 7, 8?

Which are equivalent to 1/3?
2/6 11/31 3/15 4/12 50/150?

Which are equivalent to 3/4?
0.750 0.075 0.75 7.5 75.0?

Put an arrow at 1.35 in.



Relationships and Number Patterns. Solution and application of patterns and number relationships. Tests involved various kinds of relationships including sequences, ordering, number rules and interpolation.

Sample items

2 - 10
8 - 4
10 - 5
6 - ?

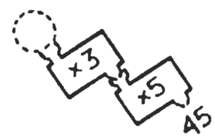
Which is Larger?
585 + 250 or 580+290

What is the missing number?
28, 25, ? , 19, 16, 13

Will 700 be in any of the boxes?

••	60	67	74	81	88	••

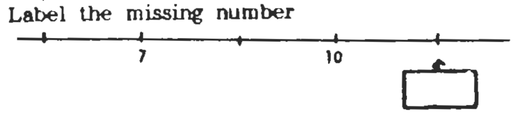
What are the missing numbers?
?, 50, ?, 200, 400, 800



Which is larger? 5/2 or 5/4
0.9 or 0.11111

Name a fraction (decimal) that is:
larger than 1/3 but smaller than 1/8
larger than 0.2 but smaller than 0.3

Label the missing number



Word Problems. Solutions of word problems requiring low levels of computation and reading comprehension, and classified according to types of problem, such as one-, two-and three-stage, extraneous data, fractions, decimals, and approximations.

Elucidation of Multiple Responses. Fluency in producing as many answers as possible that fit a given situation. There might be an infinite number of possible answers (as in the second grade sample) or a finite number of correct solutions (as in the sixth grade sample).

Sample items

Write # Sentences about 8
8 = 9 - 1
8 = 3 + 4 + 1
8 = 2 x 4

Take out 3 balls together
Add to get total score
Give all possible scores ____

Special Topic Categories. Special topic categories appeared only in the upper grades and were given less emphasis than the process categories. The basic premises of the MANS tests were retained. Problems were new to the students and did not contain any special CSMP terminology. Furthermore, tests in these categories did not require the knowledge of any particular content. They were rather general and process oriented. The special topic categories are listed below and will be described in the next chapter with the category results.

Pre Algebra (grade 6 only).
Geometry (grades 4-6)
Logic (grade 6 only)
Organization and Integration of Data (grades 5-6)
Probability (grades 4-6).

The classification of the MANS tests is somewhat arbitrary in that some tests could reasonably be placed in one of two categories. The categories themselves were based partly on the ten basic skill areas recommended by the National Council of Teachers of Mathematics and the National Council of Supervisors of Mathematics and partly on processes which are thought to be particularly important for mathematical thinking for elementary school students. Each of the following categories correspond to one of NCTM's and NCSM's ten basic skill areas: Word Problems, Estimation, Computation, Geometry, Organizing Data, and Probability. In addition, Problem Solving, the most important of the ten areas occurs throughout the MANS tests.

Development of the MANS Tests

The description of the MANS Tests given above is really a description of the tests after they had evolved into their present form. The first use of tests of this kind occurred in the first year of the Extended Pilot Test of second grade, when a total of 14 tests, some group and some individually administered, were given to classes in the local St. Louis area. In succeeding years, as Pilot Tests of higher grades were undertaken, the tests were gradually refined. Directions were simplified so that testers at distant sites could, with some training, administer the tests. A reading test was included in each booklet, thus providing a common measure across sites that could be used as a covariate. Item sampling by test halves was introduced, thereby increasing the number of items that could be administered to a class in the limited available testing time. Standardized computation tests were included, on a sampling basis, as part of the MANS Tests, eliminating the need for a separate testing period. A classification scheme for the tests was developed.

During test development for sixth grade, the entire set of MANS tests for grades 2-5 was revised to incorporate these changes at all grades, to integrate the tests from grade to grade, and to simplify administration, scoring and reporting so that school districts might undertake, in cooperation with CEMREL, their own evaluations of CSMP student learning.

For each grade level, the MANS Tests were developed using the process described below.

1. Development of Prototype Tests. Based on analyses of the CSMP curricular materials on the one hand and of available test materials (from all sources) on the other, a set of prototype tests was developed.

   The curricular review was usually rather informal, focusing on general processes that were repeated in different CSMP contexts. Occasionally more formal reviews were conducted and resulted in frequency counts of various types of items, operations, language usage, etc. Reviews of test materials included standardized achievement tests, tests of intelligence or academic ability, tests used in mathematics education research, and tests used in previous curriculum evaluations. The whole process was more inductive than deductive owing to the integrated nature of the curriculum and its lack of behavioral objectives.

   The prototype tests developed from this process consisted of a sketch of the directions, samples and diagrams for the student page, a summary of tester directions, and few test items.

2. Review and Revision. These prototype tests were reviewed by the evaluation staff, the external CSMP Evaluation Panel, and the CSMP development staff. Sometimes coordinators and teachers also reviewed the tests. Reviewers were asked to respond to the importance of the idea of being tested, the fairness of the testing situation for both CSMP and Non-CSMP students, and the likely technical quality of the prototype test as a test instrument. Reviewers made numerous suggestions for improving the task, the presentation, the directions and the items.

   Based on this review, a number of tests, perhaps 30% overall, were rejected out of hand for various reasons. The remainder were revised for pilot trials according to reviewer comments and new tests were created as a result of reviewers' suggestions.

3. Local Pilot. The revised tests, each with a full set of items and carefully written directions, were administered to classes of average ability in the St. Louis metropolitan area. Usually five or six classes were used in two stages because the first pilot inevitably revealed weakness necessitating revisions and further testing. At least half of the classes tested were Non-CSMP classes. Throughout the pilot testing, observers kept notes of what happened, especially concerning student questions and difficulties and time required (at this stage students were given as much time as needed).

   This pilot served two purposes. The first was to determine whether the test was, or could be made, practical. The major question in this regard was whether or not directions and samples could be prepared which would enable all students to at least understand the task. Many promising scales had to be rejected at this stage because of this difficulty, particularly in the lower grades. The second purpose served by the tests was to investigate the statistical properties of the proposed tests. At the test level the most important of these considerations were mean percent correct, reliability, percent reaching the last items, and distribution of scores, (i.e., not large percentages of students getting all or none of the items correct). At the item level the most important properties were percent correct, r-biserial correlation, distribution of wrong answers, and percent omitted.

   The pilot culminated in the selection of a set of tests for use in the First Year EPT. In addition to considerations of mathematical merit, practicality, and statistical properties, one other consideration was important in this selection. In the testing session in which these tests were to be used, new and difficult sets of problems would follow one after another. Thus, students' attitudes toward the tests, and their motivation for doing them, were crucial. After each pilot testing session, the tester asked students to indicate by a show of hands how much they liked the test and these student "votes" were one more consideration of test merit.

4. First Year Extended Pilot Tests. The selected tests were carefully formatted into two or three student booklets, each requiring one testing session ranging from 30 minutes for second graders to 60 minutes for sixth graders. The tests were then administered in the First Year EPT by one or two trained testers to about ten CSMP and ten Non-CSMP classes in the St. Louis area.

An extensive statistical analysis of the results of this administration was reported to the Evaluation Panel during the panel meeting held in St. Louis each fall. This review served two purposes. First, it provided a preliminary evaluation of CSMP students' achievement in comparison to Non-CSMP students. Second, it allowed the panel to make recommendations for test revisions. Virtually all tests were revised; some were revised substantially with new directions and format while others required only a revision of a few items. Other tests were eliminated entirely and new ones developed and pilot tested to increase coverage of certain topics.

5. Second Year of Extended Pilot Tests. The revised tests were formatted into 16-page student booklets and printed in quantity on inexpensive newsprint paper. Revised tester manuals were prepared and distributed to testers who were hired at each site and trained by CEMREL staff. Altogether about 60 classes were tested in the Second Year EPT, and the results of this testing form the main data presented in this chapter.

From the beginning of the development of the MANS Tests through final revisions at sixth grade, the evaluation staff remained fairly stable (two members of the usual three-person complement were on staff throughout) and the five-person Evaluation Panel changed not at all. A common understanding of what the MANS Tests were intended to accomplish and how to go about developing them, led to an efficient, informal and productive working relationship. At any time of the development process it was possible to sketch out two or three prototype tests, send them to the panel, receive comments, revise the tests, locate and schedule pilot sites, administer the tests, analyze them and make revisions, all within a very short period of time.

At each successive grade level, the selection of tests became more difficult because of the increased sophistication of students which allowed for more complicated mathematical situations, and the broader range of content and mathematical processes that needed to be measured. Furthermore, each array of suitable tests was larger than the previous year because it included not only the new tests especially developed for that grade level, but also all previously used tests in earlier grade levels, even some considered unsuitable because of their difficulty for those younger students.

MANS Technical Data

Content Coverage. The number and percent of items in each MANS category during the Extended Pilot Tests is shown by grade level in Table 8, next page.

Table 8
Percent of MANS Items by Category in Extended Pilot Tests

MANS Categories	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Average
Computation[1]	24	27	19	17	08	(19)
Estimation	11	18	14	17	09	(14)
Mental Arithmetic	21	14	14	14	11	(17)
Number Representations	08	04	11	15	09	(09)
Relationships and Number Patterns	15	21	20	08	15	(16)
Word Problems	07	06	11	24	05	(07)
Elucidation	14	07		08	14	(08)
Pre Algebra					11	
Geometry			03		01	
Logic					08	
Organization of Data		03		03	02	
Probability			08	10	05	
Total Number of items	115	180	249	309	424	

[1] Grades 2, 3. Computation was tested separately from MANS Tests, with different standardized tests used in each district. The numbers in the table are for the CTBS, which was used in four of the seven sites.
Grades 4, 5. The computation test of the Stanford Ach. Test and the CTBS were incorporated into the MANS Tests in grades 4 and 5 respectively.
Grade 6. A specially constructed computation test was part of the MANS.

The total number of items increased from grade to grade because of increased use of item sampling and because the older students worked faster. Computation, Estimation, Mental Arithmetic, and Relationships were the categories that generally received most emphasis, though they accounted for a lower than average proportion of the sixth grade MANS because of the inclusion of the five special top categories.

Standardized mathematics tests usually have three sections: computation, concepts, and word problems. There are separate MANS categories in computation and in word problems. There is no separate category for concepts since these kinds of items occur throughout the remaining categories. The average number of items in the mathematics sections of the seven leading standardized tests[1] is shown below, with the corresponding number of MANS items from the Extended Pilot Tests.

TABLE 9
Number of Test Items, MANS versus Standardized Tests

	Number of Computation Items		Number of Other Items	
	Standard Tests	MANS Tests	Standard Tests	MANS Tests
Grade 2	31	28	38	87
Grade 3	38	48	34	132
Grade 4	39	48	53	201
Grade 5	41	54	54	255
Grade 6	41	34	54	390

The MANS Tests have roughly the same number of computation items as standardized tests, but have three to five times as many non-computation items.

---

[1] CAT, CTBS, ITBS, MAT, SAT, STEP and SRA

Reliability. The KR 20 reliability was calculated for each scale and adjusted, using the Spearman Brown formula, to get the KR-20 for an equivalent 20-item test. The results are summarized below in Table 10.

TABLE 10
Summary of KR 20's Across Grades 2-6
Adjusted by Spearman-Brown to 20-item Test

Category	Total # of Tests	Average # of Items	Average KR 20	# KR 20's >.80	# KR 20's .75-.79	# KR 20's < .75
Computation[1]	6	7	.82	4	2	0
Estimation	18	8	.77	9	5	4
Mental Arithmetic	19	7	.88	19	0	0
Number Representations	12	9	.81	9	2	1
Relationships	22	8	.86	19	1	2
Word Problems	13	6	.86	11	1	1
Elucidation	6	11	.90	6	0	0
Special Topics	12	8	.83	8	2	2
Combined	108	7.5	.84	85	13	10

[1] Does not include standardized tests incorporated into MANS tests.

Most tests (79%) had a reliability of at least .80 and only a few (9%) had a reliability of less than .75. The category with the lowest KR 20's was Estimation, which included many multiple choice tests that had short time limits to promote rapid answering.

Correlations with Other Measures of Achievement. Table 11 shows correlations between total MANS scores and measures of reading ability that were used as covariates in the data analysis. Because of item sampling, different students took different sets of items; hence the median correlation coefficient across different forms is reported. In second and third grades, the median correlation across two or three achievement tests is reported.

TABLE 11
Median Correlations Between Total MANS Score
and Standardized Test Score

Grade	Standardized Reading Tests	Standardized Mathematics Tests	Kuhlmann Anderson Ability Test
2	.54	.76	.77
3	.57	.72	.70
4	.60	.64[1]	
5	.61	.60[1]	
6	.59		

[1] Computation test only.

The correlations with reading score are very consistent, between .54 and .61 regardless of grade level. Correlations with mathematics scores are higher and with Kuhlman Anderson higher still.

Correlations with Teacher Rating of Student Ability. In grades 4 and 5, teachers were asked to rate the mathematical problem solving ability of each of their students, using a 5-point scale. The median correlation with total MANS score was .66 in 4th grade and .57 in 5th grade.

Teacher Ratings. In fourth and fifth grades teachers were asked to rate the importance of each MANS Test on a 1 to 5 scale, where 1 = not important and 5 = very important. Average ratings for each test were calculated, then these average ratings were averaged for each category.

There was very little difference between CSMP and Non-CSMP teachers' ratings. For both groups, ratings fell into four groupings:

> two categories were always rated very highly, 4.4 or better (Computation and Word Problems),
>
> four categories had an average rating of around 4.0 (Organization of Data, Estimation, Number Representations and Mental Arithmetic),
>
> three categories had a rating in the upper 3's, i.e., 3.5-3.8 (Relations and Number Patterns, Elucidation and Geometry), and
>
> one category was rated below average in importance (Probability).

Setting

Extended Pilot Tests

As described earlier in Chapter II, part of the CSMP development cycle for each grade was a two-year Extended Pilot Test (EPT) of the materials. During the first year of the EPT, about 10 - 15 classes from school districts in the metropolitan St. Louis area used the CSMP curriculum. Teachers were trained in CEMREL-conducted summer workshops and materials were provided by CEMREL to participating classes. Extensive observations and teacher interviews were carried out by both evaluation and development staff, and student interviews were conducted. Evaluation instruments were developed and used at the end of the year to compare the performance of CSMP and Non-CSMP classes, with the Non-CSMP classes elected jointly by the evaluation staff and the local districts. Thus the first year resulted in preliminary evidence about CSMP's effects on students and potential implementation problems and, in addition, provided the evaluation staff with a chance to develop and test a variety of instruments for use the following year in the second year Extended Pilot Test.

In the second year of the Extended Pilot Tests, the program was available to districts nationally and about 40-60 classes per grade level participated. Districts trained their own teachers usually through the local coordinator who had been trained in a CEMREL workshop. All districts (including those whose classes had previously participated in the first-year EPT) had to purchase the materials. In order to participate, districts had to agree to name a local coordinator who would provide a CSMP-recommended amount of training to their teachers and would cooperate in any data gathering activities (testing, site visits, questionnaires, etc.). In practice, once districts adopted the program, they became fairly autonomous and adapted the program to fit local needs. They selected teachers and schools as they saw fit, trained teachers in ways that were different from what CEMREL recommended (and usually less exacting), and cooperated in data gathering activities in proportion to how useful the data was to them. This was both an advantage and a disadvantage for the evaluation enterprise.

The wide variation in treatment meant that no single "program" was being implemented uniformly. Furthermore, since sites were widely dispersed in distant locations, it was difficult to determine the exact nature of the adaptations; site visits could only be made occasionally and teacher logs and questionnaires were not always returned. On the other hand, this very freedom from restraint gave the sites greater ownership over the program and led to fairly natural implementations which would be far more informative in predicting CSMP's effects than would a detailed, rigidly-adhered-to plan of implementation. The nature of the curriculum, especially the spiral sequencing of content and the detailed lessons in the Teacher's Guide, made CSMP a difficult program to change drastically at the classroom level. Such changes inevitably led to the rapid demise of the program in the classroom. If the program continued to be taught, it could safely be assumed that it was being taught roughly as prescribed.

The design for assessing student achievement data was always comparative in nature. The performance of CSMP classes was compared with the performance of Non-CSMP classes. The selection of Non-CSMP classes to serve as control classes in this experimental comparison was always a source of concern since the random assignment of teachers and students to curriculum was not possible. Instead, coordinators were asked to select, from nearby schools, classes whose students and teachers were as similar as possible to the CSMP classes.

There are several reasons to believe that there were no systematic differences between the CSMP and Non-CSMP groups. Subsequent analysis of student test scores in reading usually corroborated the coordinators' judgments about student ability. A study of teachers whose classes were selected as control classes one year, showed that when these teachers started teaching CSMP the following year, their classes performed well in comparison to the previous year's classes. They did at least as well as the earlier CSMP classes had done and better than their own previous Non-CSMP classes. Interviews and observations by the evaluation staff confirmed district personnel's judgment regarding teacher comparability.

As the evaluation reached the higher grades, the concern for teacher comparability became less acute. Teachers were not individually selected nor did they not volunteer for the program as sometimes happened at lower grades. If a teacher was a fourth grade teacher in a school where all the third graders were studying CSMP, that teacher knew he or she would inherit both the program and a class of CSMP students next year as a matter of course. Hence in the later grades, the comparability issue focused on the school as the unit of adoption, rather than the teacher.

Table 12 lists the school districts who participated in one or more years of the Extended Pilot Test.

Table 12
Participating Districts, Second Year EPT
First Entry = # CSMP Classes, Second Entry = Non-CSMP Classes

District Number	Type of Community	Section of Country	Grade 2	3	4	5	6
1	Medium City	North Central				0-3	0-6
2	Suburb	East			2-2	2-2	2-2
3	Small City	Central	6-6				
4	Large City	North Central	6-6				
5	Large City	East					2-2
6	Suburb	Central	3-3	4-5	1-3	0-2	
7	Small City	West				0-5	
8	Suburb	North Central					1-6
9	Suburb	East					0-4
10	Suburb	East			3-3		
11	Small City	North Central					7-0
12	Suburb	Central	3-2	4-4	6-6	6-6	8-6
13	Large City	Central					0-6
14	Medium City	North Central					0-5
15	Suburb	Central			6-0	6-0	
16	Small City	East			6-0	6-0	6-0
17	Large City	South				5-3	
18	Suburb	Central	6-5		2-2	2-2	
19	Large City	East			3-4		
20	Small City	South		15-12			
21	Small City	East	6-6	6-12			
22	Large City	Central	3-3	3-3	1-1	1-2	
23	Suburb	Central				3-0	
Total Number of Class			33-31	32-36	30-21	31-25	26-37
Mean Percentile Rank on Reading Test			56-54	55-55	64-62	61-60	77-78

The CSMP and Non-CSMP classes were very similar in ability each year; covariate adjustments in MANS scores due to differences in ability between the two groups was always small, averaging less than 1%. There is an upward trend in overall ability levels so that by sixth grade, the median percentile ranks on the reading score were above 75. In sixth grade, there are several districts with no CSMP classes but some Non-CSMP classes. This is because at some other districts, CSMP was implemented district-wide requiring the use of comparison classes from other districts. In every case in which this was done, the other district was similar to the CSMP district, and was using CSMP at lower grades (i.e. started later) with the intention of continuing it on a year-by-year basis.

Joint Research Studies

Several districts who had begun the program a number of years after the pilot study began, and who were thus unable to participate in the Extended Pilot Tests, expressed an interest in conducting an evaluation of CSMP in their own districts. CEMREL cooperated in these efforts by supplying and scoring the tests. Local districts selected CSMP and Non-CSMP classes, trained testers and did the testing.

The MANS Tests used in these Joint Research Studies were the revised MANS, i.e., they incorporated the revisions that were made in after the completion of the Extended Pilot Tests in grades 2-5. The main changes were the following:

At each grade level only two test booklets, i.e., two testing sessions, were required.

The Gates McGinitie Vocabulary Test was incorporated at each grade level.

Whole number computation tests were developed for each grade based on analysis of the major standardized achievement tests.

Through item sampling, the total number of items was increased at all grades. Excluding the Vocabulary tests, the number of items ranged from 160 (second grade) to 266 (fifth grade), though an individual student would only do about half of these items.

Larger numbers of common items were included on tests which appeared in consecutive grades.

The directions were simplified, causing the elimination of some hard to administer tests. A Coordinator Training Manual was developed and the format for the Tester Manuals was standardized so that local district could carry out all phases of the testing.

The proportion of items in each category was changed somewhat. Each of the seven process categories was tested at each grade. Relationships and Number Patterns was the most heavily represented process category, containing an average across grades of 22% of the items. Word Problems, which require the most time per item to administer, was the least represented category, and average of 7%. The other five process categories each accounted for between 12% and 15% of the items.

Compared to the MANS Tests used in the Extended Pilot Tests, this was a decrease in emphasis in Computation and an increase in Relationships and Number Patterns and in Elucidation of Multiple Answers (the latter increase due to the fact that this category wasn't tested in fourth grade previously). There was one test in Geometry in fourth grade, and one test each in Geometry, Organization of Data, and Probability in fifth grade, none of which accounted for more than 5% of the items.

In addition to the changes in the MANS Tests, CSMP classes participating in these Joint Research Studies at second or third grades were using the final version of the curriculum, which incorporated the revisions made after the Extended Pilot Tests.

Table 13 lists the school districts who participated in one or more grades of Joint Research Studies. Those with ID's less than 24 also participating in the Extended Pilot Tests, usually providing comparison classes in fifth or sixth grade since their own implementation had not reached those grade levels.

Table 13
Participating Districts, Joint Research Studies
First Entry = # CSMP Classes, Second Entry = Non-CSMP Classes

District Number	Type of Community	Section of Country	Grade 2	Grade 3	Grade 4
1	Medium City	North Central	33-13	43-39	21-26
4	Large City	North Central	10-10	12-7	
7	Small City	West	5-5		
9	Suburb	North East			5-5
11	Small City	North Central	2-2	4-5	
13	Large City	Central			6-6
17	Large City	South	5-4	5-4	
24	Medium City	East		4-3	
25	Suburb	North Central	2-2		
26	Medium City	West	2-2		
27	Medium City	West	3-3		
28	Suburb	North Central		2-2	
29	Small City	North Central	7-4		

Altogether then, 29 districts participated in either the Extended Pilot Test or Joint Research Studies. Eleven of these were suburbs, seven were small cities, six were large cities and five were medium cities. There were eight districts from each of the Central, North Central and Eastern parts of the country, and there were three from the West and two from the South.

Analysis of Class Mean Scores on Total MANS

Summary

This chapter describes results of Total MANS scores. Except for the Extended Pilot Test at second and third grades, this Total score includes Computation. In the next chapter a category-by-category analysis will be presented (except for Computation, which will be deferred to the following chapter).

The primary method of analysis throughout the Extended Pilot Tests was an analysis of covariance on class means. For each class, a mean score was calculated for both the MANS and a Reading or Vocabulary Test. In cases of item sampling, averages were computed by adding together the average score on each half of the test. Then a one-way ANCOVA was carried out to compare CSMP and Non-CSMP classes, using Reading or Vocabulary score as a covariate.

In second and third grades, different covariate measures were used in different sites (Kuhlmann Anderson Test of Mental Ability, CTBS Reading Comprehension, ETS Cooperative Reading Test, and Stanford Achievement Total Reading). Thus, separate analyses were carried out for each district and will be reported in the grade-by-grade analysis in the following pages. In order to summarize further for each grade, second and third grade classes were converted to a common metric using equipercentile methods, and a single Analysis of covariance performed.

In fourth through sixth grades, a covariate test was built into the MANS Tests so that all districts used the same Reading or Vocabulary Test. The tests used were the Stanford Achievement Reading Comprehension, the CTBS Reading Comprehension and the Gates McGinitie Vocabulary, respectively, for grades 4-6.

Following the brief summary of Extended Pilot Test data given below, a grade-by-grade analysis of both Extended Pilot Test and Joint Research Study data will be presented, together with graphs of class means on the Total MANS score.

Table 14 summarizes total MANS scores in the Extended Pilot Tests at each grade level. The adjusted mean scores were calculated by computing mean score across all CSMP classes and across all Non-CSMP classes, and then adjusting these two means for differences in Reading or Vocabulary score. The adjustments were always small because the groups were well matched in Reading or Vocabulary score.

Table 14
Summary Data of Total MANS Scores
Extended Pilot Tests

Grade	Number of Classes CSMP	Number of Classes Non-CSMP	Adjusted Mean MANS CSMP	Adjusted Mean MANS Non-CSMP	Diff/SD[1]	P-Value[2]
2(Form A)	16	15	20.0	17.4	.41	.04
2(Form B)	17	16	25.5	22.5	.51	.02
3	32	36	70.1	63.2	.60	.01
4	30	21	145.3	127.5	.77	.01
5	31	25	193.9	177.2	.47	.01
6	26	37	290.0	263.7	.60	.01

[1] Difference in adjusted means/standard deviation of means.

[2] Covariance F-test with 1 and 28,30 d.f. (second grade) and 1 and 65, 48, 53 and 60 d.f. (grades 3-6 respectively).

Table 15 shows the average percent correct across all items, obtained simply by dividing the adjusted means from the previous table by the total number of items. Also shown is the percentile rank corresponding to the mean Reading or Vocabulary score.

Table 15
Average Percent Correct on MANS Items
By Grade in Extended Pilot Tests

Grade	Covariate Percentile Rank		Average Percent Correct on MANS	
	CSMP	Non-CSMP	CSMP	Non-CSMP
2	56	54	52	46
3	55	55	52	46
4	64	62	58	51
5	61	60	63	57
6	77	78	68	62

The difference is remarkably constant across years, always between 6 to 8 points in favor of CSMP classes. More items were answered correctly in the upper grades and this is consistent with the higher ability level of participating students in those grades.

Second Grade Total MANS.

Table 16 summarizes the second grade results on a district-by-district basis from both the Extended Pilot Test and the Joint Research Studies. The scores from the Joint Research Studies are higher because there were more items in these revised MANS Tests.

Table 16
Summary of Second Grade MANS Results,
Total MANS Score

District	Description	Number of Classes		Adjusted Means		Signif at .05[1]
		CSMP	Non-CSMP	CSMP	Non-CSMP	
Extended Pilot Tests						
3	Small City, Central	6	6	54.2	48.5	Y
21	Small City, East	6	6	46.0	33.6	Y
6	Suburb, Central					
12	Suburb, Central	15	13	41.6	35.3	Y
18	Suburb, Central					
22	Large City, Central					
4	Large City, North Central	6	6	41.5	42.2	N
Joint Research Studies						
27	Medium City, West	3	3	112.0	93.0	Y
11	[2]Small City, North Central	2	2	111.4	104.4	NA
1	Medium City, North Central	21	26	98.2	85.6	Y
7	Small City, West	5	5	95.7	83.3	Y
29	Small City, North Central	7	4	90.2	76.0	Y
25	Suburb, North Central	2	3	86.6	74.0	NA
26	Medium City, West	2	2	73.7	70.7	NA
4	Large City, North Central	10	10	62.2	57.6	N
17	Large City, South	5	4	52.7	44.9	N

[1] Y = Significant, N = Not significant, NA = too few classes for application of Analysis of Covariance on class means.

[2] Upper track students.

Altogether, there are 13 comparisons in the table, 10 of which had enough classes (n>5) to reasonably carry out an Analysis of covariance. Seven produced significant differences in favor of CSMP. The other three comparisons were from large urban districts in classes of below average ability (Districts 4 and 17). The Joint Research Studies from these two districts produced CSMP advantages of 8% and 17%, but were not significant because of the wide variation in scores. This will be illustrated after the presentation of third grade results.

Another statistic that can be used to compare performance is percentage of items answered correctly. If each study is weighted equally, the mean percent correct in the Extended Pilot Tests was 53 for CSMP classes versus 46 for Non-CSMP; in the Joint Research Studies the percentages were almost the same, 52 versus 46.

The graphs below show the performance of each participating class. Each class is represented by an entry on the graph, "x" for a CSMP class and "•" for a Non-CSMP class. Horizontal position on the graph is determined by Reading score; the farther to the right - the higher the average reading score of the class. The vertical position is determined by Total MANS score; the farther up - the higher the average MANS score for the class. The regression line which has been drawn on the graph is the best linear prediction of MANS score for a given Reading score.

Figure 6 shows class means from the second grade Extended Pilot Test. Two graphs are needed because half the classes took one set of tests, Booklet A, and the other half took a different set of tests, Booklet B.

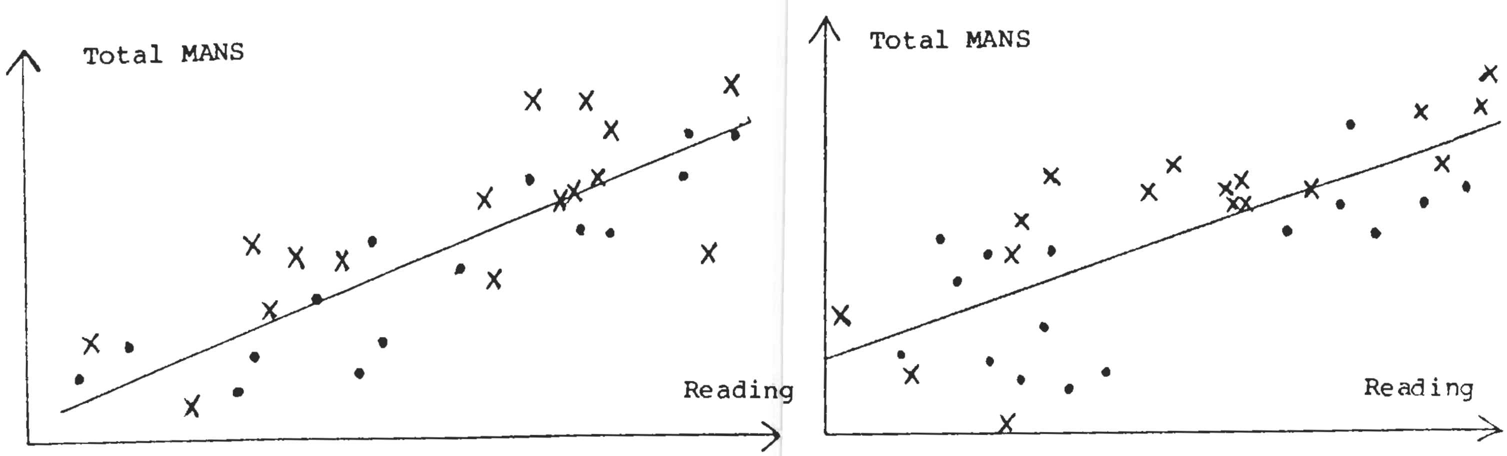


Fig. 6. Second Class Class Means,
Extended Pilot Test
Booklet A (left) and Booklet B (right)
(X = CSMP Class, • = Non-CSMP)

Figure 7 shows mean scores for all classes which have participated in Joint Research Studies.

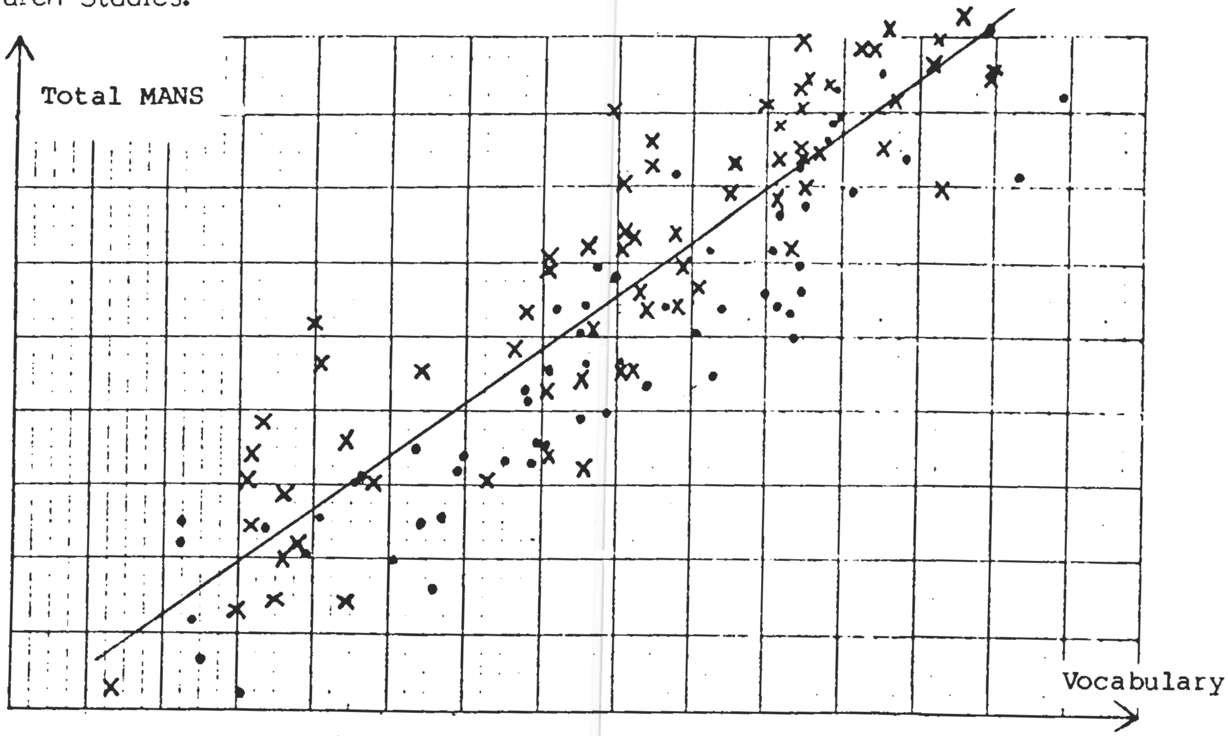


Fig. 7. Second Grade Class Means, Joint Research Studies (X = CSMP Class, ● = Non-CSMP Class)

Third Grade Total MANS.

Table 17 summarizes the third grade results of the Extended Pilot Test.

Table 17
Summary of Third Grade MANS Results,
Total Mans Score

District	Description	Number of Classes CSMP	Number of Classes Non-CSMP	Adjusted Mean CSMP	Adjusted Mean Non-CSMP	Signif at .05[1]
Extended Pilot Tests						
21	Small City, East	6	12	77.1	67.4	Y
6	Suburb, Central	4	5			
12	Suburb, Central	4	4	72.7	65.0	Y
22	Large City, Central	3	3			
20	Small City, South	15	12	58.2	54.5	N
Joint Research Studies						
24	[2] Medium City, East	4	3	151.6	125.8	Y
11	[2] Small City, North Central	4	5	145.2	116.9	Y
1(82)	Medium City, North Central	20	26	122.8	106.2	Y
1(83)	Medium City, North Central	33	13	119.7	110.3	Y
28	Suburb, North Central	2	2	114.7	93.1	NA
4	Large City, North Central	13	6	90.3	82.5	N
17	Large City, South	5	4	71.5	68.0	N

[1] Y = Significant, N = Not Significant, NA =too few classes for application of Analysis of Covariance on class means

[2] Upper track or gifted students.

Altogether there are 10 comparisons in the table. Six produced significant differences all in favor of CSMP and one had too few classes to test for significance.

The average percentage of items answered correctly (obtained by average across studies) was 52 for CSMP versus 47 for Non-CSMP in the Extended Pilot Tests and 60 versus 52 in the Joint Research Studies. The higher percentages correct in the Joint Research Studies no doubt reflect the fact that in two of the eight districts, upper track or gifted students were tested.

Figures 8 shows third class means from the Extended Pilot Test. Most classes took both test booklets, but in one district, half the classes took Booklet A and the other half took Booklet B. Hence all classes cannot be represented on a single graph.

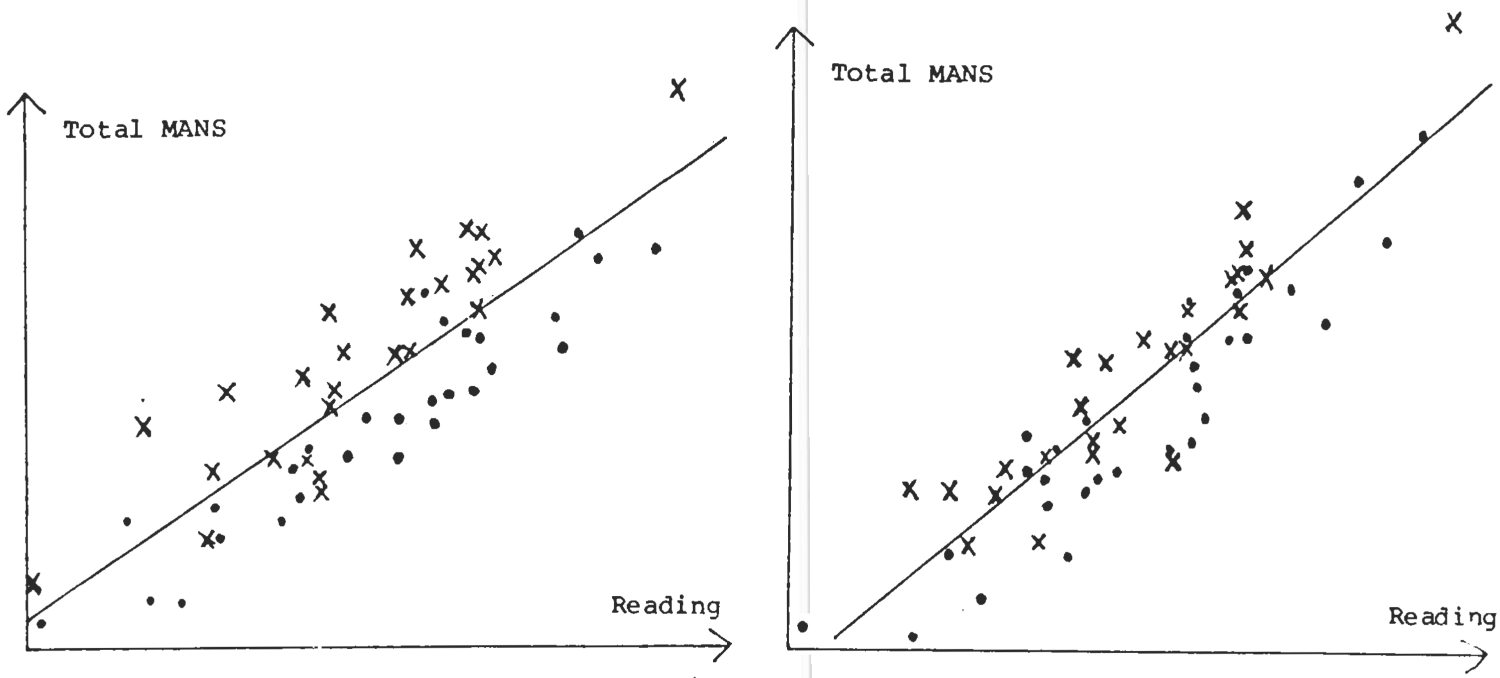


Fig. 8. Third Grade Class Means
Extended Pilot Test
Booklet A (left) and Booklet B (right)
(x = CSMP Class, • = Non-CSMP)

Figure 9 shows the mean scores for all third grade classes which participated in Joint Research Studies.

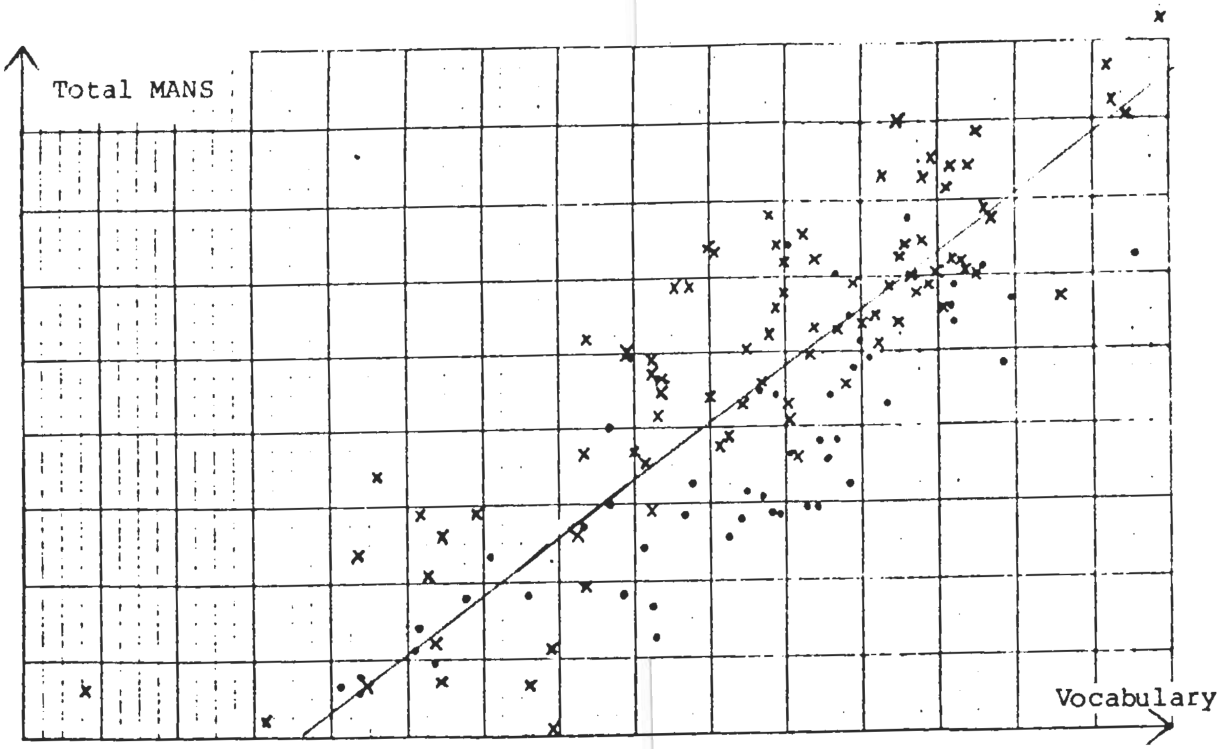


Fig. 9. Third Grade Class Means
Joint Research Studies
(x = CSMP Class, ● = Non-CSMP)

For all of the data reported thus far for either second or third grade, including both Extended Pilot Tests and Joint Research Studies, there were only three districts in which significant differences in favor of CSMP were not found (excluding districts with too few classes to properly perform the analysis). Two of these districts were Districts 4 and 17, both large city school districts. It is instructive to look at graphs of class means for these districts; two such graphs are shown below.

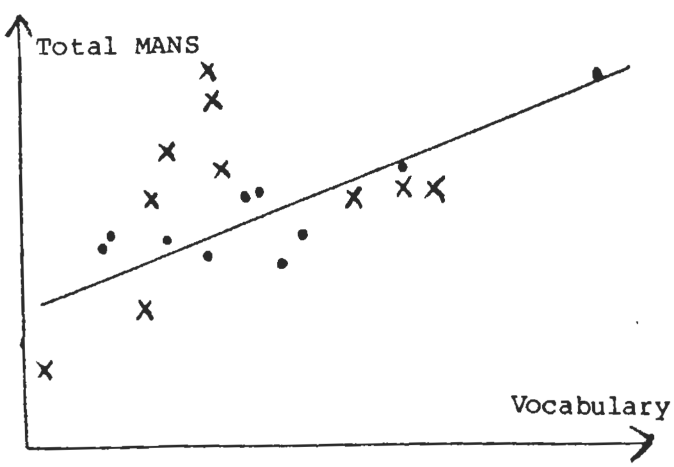


Fig. 10. Second Grade Class Means, District 4
(X = CSMP Class, ● = Non-CSMP Class)

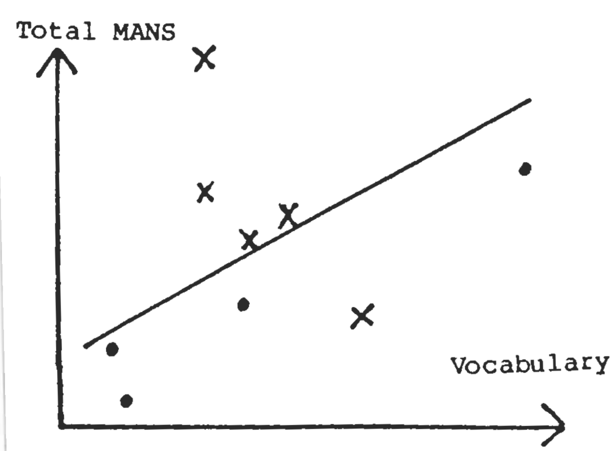


Fig. 11. Third Grade Class Means, District 17

These figures both show that CSMP classes performed better than Non-CSMP classes overall. But both figures also show one or two extremely high scoring CSMP classes (much higher than all other CSMP classes) and one or two low scoring CSMP classes (lower than all Non-CSMP classes). For the Non-CSMP classes, MANS scores are predicted quite well from Vocabulary scores. None of the Non-CSMP classes, however, did particularly well or particularly poorly in relationship to reading score. If a regression line were drawn through Non-CSMP classes (which actually is the case in the first figure), most Non-CSMP classes would fall close to that line.

This inconsistency of CSMP performance, with wide dispersion from the regression line (i.e., unpredictability), is very different from what is usually observed. Ordinarily, there are occasional outliers, but most CSMP classes fall fairly close to their regression line. Not enough is known about the implementation of CSMP in the aberrant classes of these two districts. However, in both districts, coordinators were able to name the teachers of very high and low scoring classes before seeing the data. The reasons given had to do with teacher attitude and extent and quality of implementation, though how much these were related to general teaching ability remains unknown.

Fourth Grade Total MANS

Table 18 summarizes the fourth grade results. The Extended Pilot Test results are given in a single row. There were several districts with only a few participating classes,thus a single Analysis of Covariance wascomputed for the entire group of 51 classes. Classes from nine districts altogether were represented and these districts were listed earlier in Table 12.

Table 18
Summary of Fourth Grade MANS Results,
Total MANS Score

District	Description	Number of Classes CSMP	Non-CSMP	Adjusted Means CSMP	Non-CSMP	Signif at .05
Extended Pilot Test		30	21	145.3	127.5	Y
Joint Research Studies						
13	Large City, Central	6	6	168.1	145.3	N
1	Medium City, North Central	23	23	158.4	141.2	Y
9	Suburb, North East	5	5	158.2	137.7	Y

Figures 12 and 13 show fourth grade class means from the Extended Pilot Test and Joint Research Studies respectively.

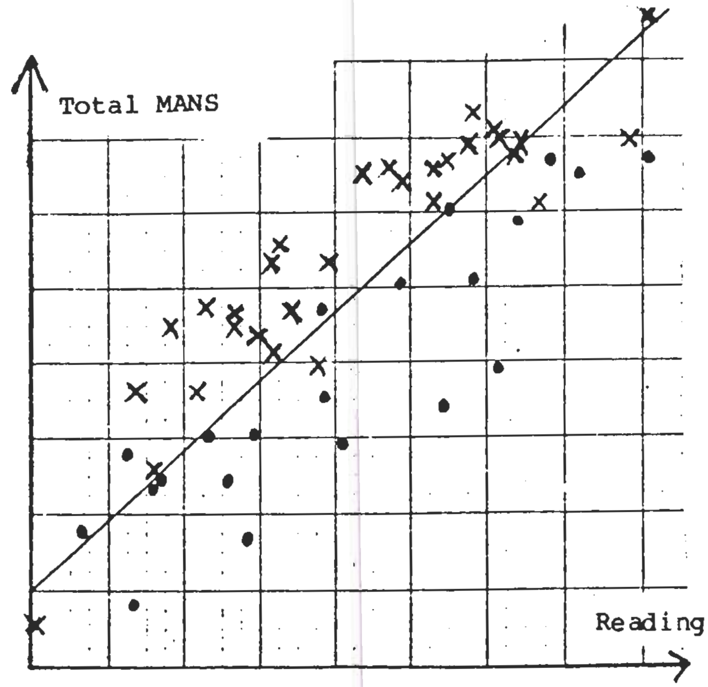


Figure 12. Fourth Grade Class Means
(Extended Pilot Test)

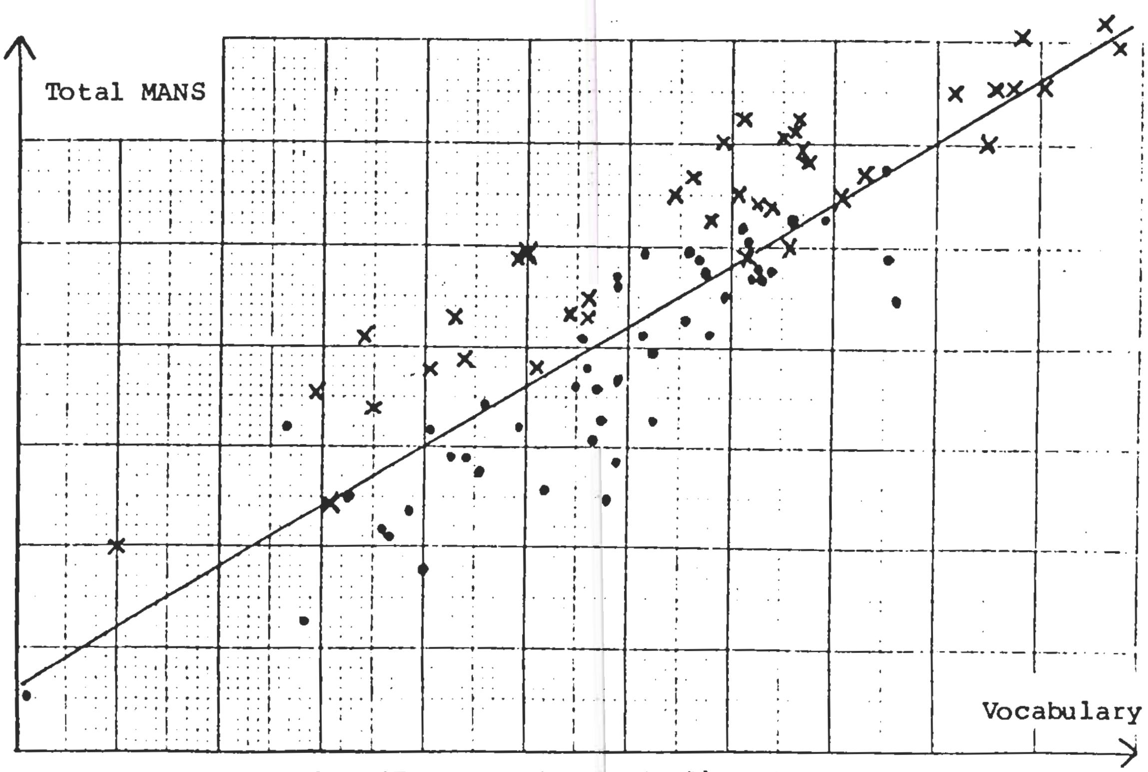


Fig. 13. Fourth Grade Class Means
(Joint Research Studies)

Fifth and Sixth Grade Total MANS

Only a very few classes in fifth or sixth grades have participated in Joint Research Studies. (the Extended Pilot Test of sixth grade was only completed in 1982, and these isolated classes are not reported here.) Table 19 gives summary data from the Extended Pilot Tests at these grade levels.

Table 19
Summary of Fifth and Sixth Grade MANS Results
Total MANS Score

Grade	Number of Classes		Adjusted Means		Signif at .05
	CSMP	Non-CSMP	CSMP	Non-CSMP	
5	193.9	177.2	31	25	Y
6	290.0	263.7	26	37	Y

Figures 14 and 15 show graphs of class means on the fifth and sixth grade Total MANS.

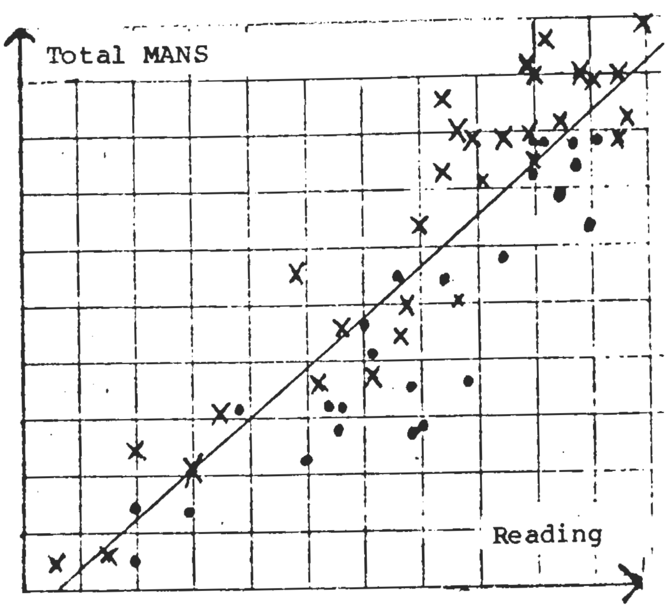


Fig. 14. Fifth Grade Class Means
Extended Pilot Test
(x = CSMP Class, ● = Non-CSMP)

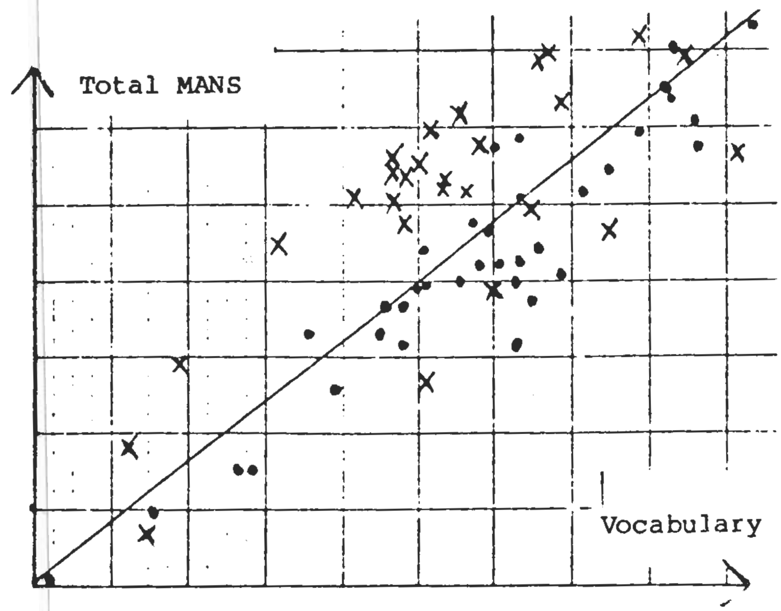


Fig. 15. Sixth Grade Class Means
Extended Pilot Test
(x = CSMP Class, ● = Non-CSMP)

Other Levels of Analysis

Analysis at the School and District Level

At fifth and sixth grades, analyses were carried out using school and district as units of analysis instead of class. There was very little difference in these results compared to the class level analyses reported earlier. In fifth grade, for example, the t-statistics for differences in means were 4.5, 4.8 and 4.1 for class, school and district level analyses respectively, all significant at .01. The school data were based on 12 CSMP and 12 Non-CSMP schools and the district data on 6 CSMP and 6 Non-CSMP districts. (Most of the districts contained both CSMP and Non-CSMP schools, but none of the schools contained both CSMP and Non-CSMP classes.)

The sixth grade data are shown in Figures 16 and 17 below. Figure 16 shows school means; each entry on the graph represents a school. In Figure 17, each entry represents a district.

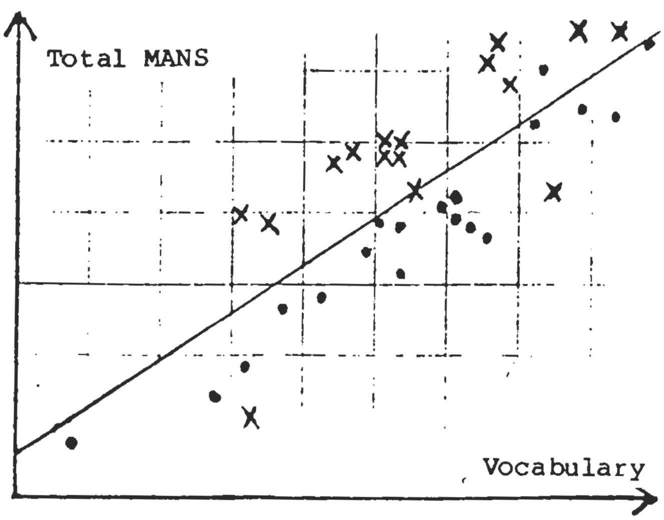


Fig. 16. 6th Grade School Means
(x = CSMP school, ● = Non-CSMP)

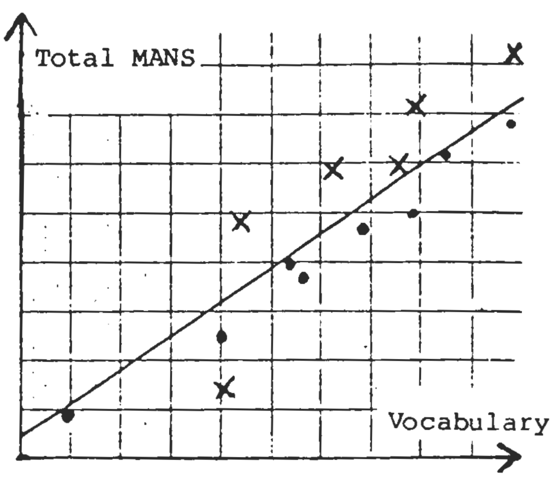


Fig. 17. 6th Grade District Means
(x = CSMP district, ● = Non-CSMP)

These higher levels of accumulation tend to stabilize MANS scores relative to Vocabulary score. In Figure 17 for example, a regression line through only the Non-CSMP districts would predict MANS scores very accurately; all districts would be very close to the line. The CSMP districts were also fairly predictable except that one CSMP district did very poorly, and was farther below the regression line than any other district. In that district, one school, containing two CSMP classes, participated in the testing. Not enough is known about the circumstances of the implementation in that district to explain this finding. The coordinator was not greatly surprised by the results and thought that CSMP classes then in the lower grades would do much better when they get to sixth grade than the present group.

District level graphs from the fourth and fifth grade Extended Pilot Tests are shown below in Figures 18 and 19. Once again, the data are more stable when district means are used. Also, in fourth grade, it was again true that one CSMP district (a different one than in sixth grade, but also one with relatively lower ability students) did not perform as well as other CSMP districts.

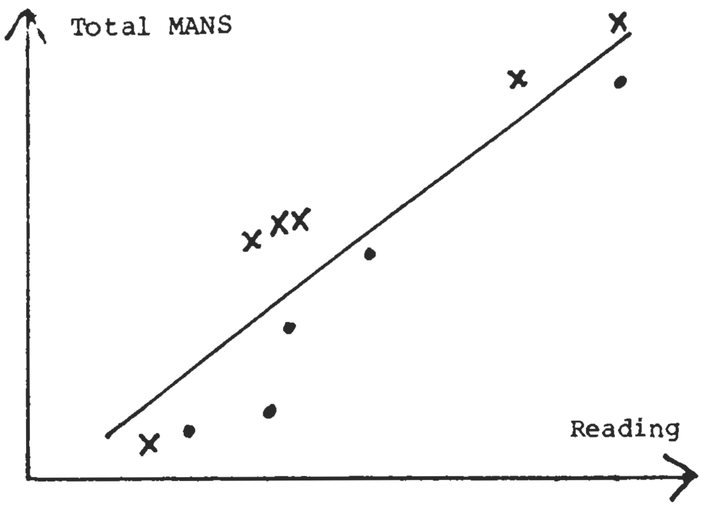


Fig. 18. Fourth Grade District Means
(x = CSMP District, ● = NON-CSMP)

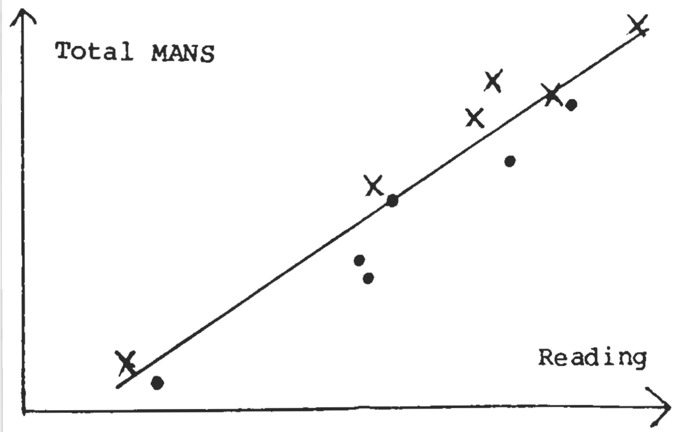


Fig. 19. Fifth Grade District Means
(x = CSMP District, ● = Non-CSMP)

Student Level Analysis

At each grade level, students were divided into four or five groups according to their reading or vocabulary score. National norms were used to determine these groupings. For each such group, an average total MANS score was calculated. These means scores are plotted, separately for CSMP and Non-CSMP students, in Figures 20 to 21 below. The first two figures are slightly less accurate because a separate reading score was not calculated for each group. Also, in Figure 20, the graph points were determined by adding together the separate totals from Booklets A and B, each containing different tests, with classes randomly assigned either A or B.

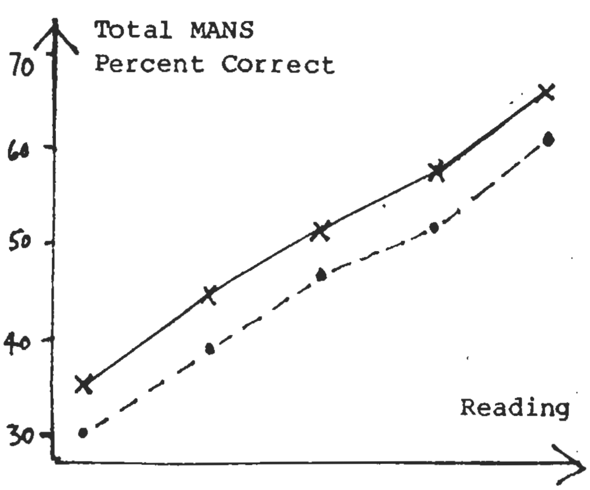


Fig. 20. Second Grade Student Means
Students grouped by reading score
(x = CSMP Students, ● = Non-CSMP)

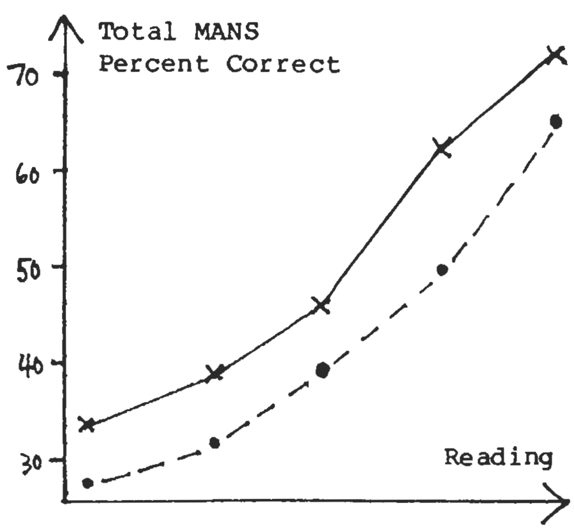


Fig. 21. Third Grade Student Means
Students grouped by reading score
(x = CSMP Students, ● = Non-CSMP)

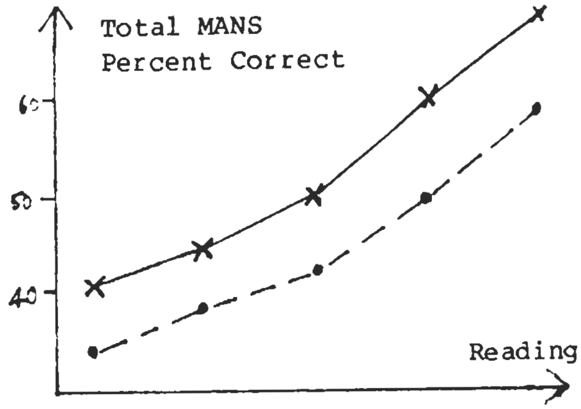


Fig. 22. Fourth Grade Student Means
Students grouped by reading score
(x = CSMP Students, ● = Non-CSMP)

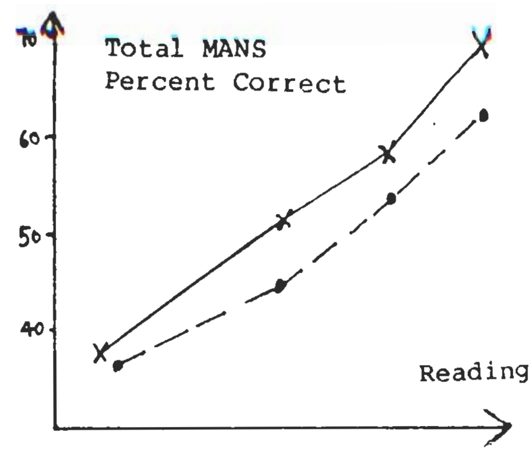


Fig. 23. Fifth Grade Student Means
Students grouped by reading score
(x = CSMP Students, ● = Non-CSMP)

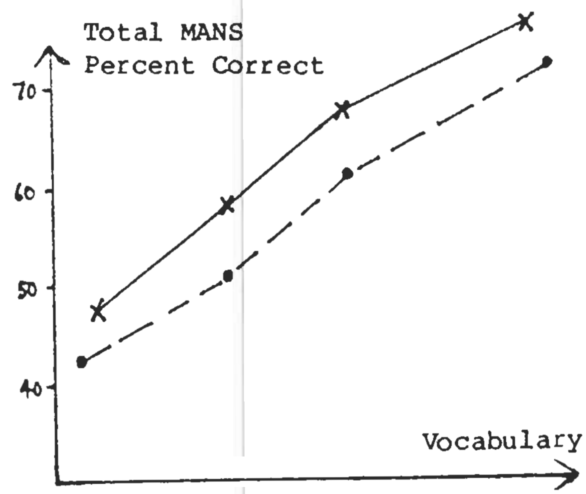


Fig. 24. Sixth Grade Student Means
Students grouped by reading score
(x = CSMP Students, ● = Non-CSMP)

The results are very consistent; CSMP students outperformed Non-CSMP students at every ability level in all grades. The largest differences occurred in fourth grade. It was shown previously that from analysis of class means, the largest difference in standard deviation units was also at fourth grade. In fifth and sixth grades, the difference in performance at the lowest ability level is smaller than at other ability levels.

Summary

The MANS Tests were an attempt to assess some of the underlying thinking skills of CSMP without overtly using CSMP representational languages or terminology. The usual emphasis on computation and word problems was drastically reduced so that tudents could be presented with a wide variety of often unfamiliar situations requiring some mathematical application. The tests contained many "problems", though most of the MANS items were not "problem solving" in the strictest sense, nor could any paper and pencil, group-administered test qualify in that sense. But the MANS tests were closer to true problem solving than most standardized achievement tests in mathematics, and they turned out to be a rather valuable, frequently used, product with potential use independent of CSMP.

The original MANS tests were administered to at least five districts and 50 classes in each of grades 2-5 during the formal CSMP Extended Pilot Tests. The revised MANS were administered in 13 districts to over 300 classes in subsequent Joint Research Studies, which were cooperative ventures between CEMREL and the local district. CSMP and Non-CSMP classes were comparable in ability, as measured by standardized reading and vocabulary tests. Similarly, schools were comparable, usually from the same area with similar teaching staffs. Class mean scores were analyzed using Analysis of Covariance on the class means, with reading or vocabulary scores as covariate.

The results leave no room for doubt. CSMP students, classes, schools and districts performed better than their Non-CSMP counterparts. This happened at all grade levels, for all ability levels, and in every kind of school. Looking at graphs of class means becomes a repetitive exercise. It is this consistency of results which leaves no doubt that something happened and that CSMP caused it.

The importance of this overall finding, the educational significance, depends on how big the difference is and how important the abilities being tested are.

Consider the student level effect size, i.e., difference in scores divided by standard deviation. At sixth grade, this 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 all 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 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 of Educational Progress, in mathematics, 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:

> "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 practiced 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."

The CSMP curriculum is a demonstration that such gains are possible.

V. STUDENT ACHIEVEMENT: INDIVIDUAL MANS CATEGORIES

Summary

Results from each category will be presented separately. The tests making up a category will be reviewed with sample results. The results will be based on data from the Extended Pilot Tests, except for a few graphs of class means from Joint Research Studies which will be used to illustrate certain findings. In addition, results from individually administered tests in third and fourth grades will be described briefly at appropriate places. Between 100 and 150 students in the St. Louis area were tested on an individual basis, during the Extended Pilot Test, using more extensive and open ended formats than were possible in a group setting; results are described more fully in Appendix G.

Before presenting category-by-category results, a brief overview of the results for all categories will be given. Table 20 shows adjusted class means for CSMP and Non-CSMP classes at each grade. The means were derived in the usual way, i.e., computing a mean category score and a mean reading or vocabulary score for each class and then performing a one-way Analysis of Covariance of the class means.

Table 20
Adjusted Means, MANS Process Categories
(First entry = CSMP mean, Second entry = Non-CSMP)

	Grade 2		Grade 3		Grade 4		Grade 5		Grade 6	
Estimation	4.4	4.1	16.6	15.4	22.9	21.2	28.5	27.4	24.4	22.5
Mental Arithmetic	13.9	11.0	11.9	9.4	22.2	19.5	26.1	21.8	31.5	28.3
Numb. Reps.	5.6	4.5	3.3	3.0	13.3	12.9	30.0	27.7	28.8	26.3
Relats. & Numb. Patts.	9.9	9.2	24.2	21.2	32.2	23.7	15.6	13.4	46.1	40.3
Word Problems	4.3	4.0[N]	5.3	4.8	15.2	13.1	14.1	12.2	15.1	13.6
Elucidation	7.4	6.8[N]	5.9	5.8[N]			16.3	13.3	38.8	31.9
Total	45.5	39.6	67.2	58.6	105.8	90.4	130.6	115.8	184.7	162.9

N = Not significant; all others significant in favor of CSMP at .05 on F-Test

A total of 29 out of the 35 comparisons in the above table produced significant differences in favor of CSMP. In the other six comparisons, CSMP classes had higher mean scores, but the differences were not significant.

In Table 21, below, the numbers from the previous table are translated into percent correct to allow a common basis for comparison.

Table 21
Adjusted Mean Percent Correct,
MANS Process Categories,
(First entry = CSMP, Second entry = Non-CSMP)

	Gr. 2	Gr. 3	Gr. 4	Gr. 5	Gr. 6	Average
Estimation	34-32	50-47	55-52	62-60	61-56	52-49
Mental Arithmetic	58-46	48-38	62-54	62-52	68-62	60-50
Numb. Reps.	62-50	41-38	49-48	64-59	72-66	58-52
Relats. & Numb. Patts.	52-48	64-56	66-48	68-58	74-65	65-55
Word Problems	54-50	53-48	58-50	60-51	66-59	58-52
Elucidation	53-49	49-48		65-53	65-53	58-51

Mental Arithmetic and Relationships and Number Patterns were the categories with largest CSMP advantage, an average of 10 percentage points in each case. This difference translates into about 20% more correct answers for CSMP classes than for Non-CSMP classes. Number Representations, Word Problems, and Elucidation produced average differences of 6 or 7 percentage points, i.e., about 13% more correct answers. Estimation was the category with the smallest difference, with an average difference of only about 3 percentage points.

Ordinarily, these average percentages would be somewhat deceptive since they are unweighted. Categories with a disproportionately small number of items in the lower grades may have undue influence. But in the previous table, the average percentages reflect the findings at each grade level fairly well. In only one category, Elucidation of Multiple Answers, were the findings very different across grade levels; the differences were quite small in second and third grades, but quite large in fifth and sixth grades. An explanation for this discrepancy will be given when that category is discussed.

A third way of looking at the data is to compare the difference in adjusted class mean scores with the standard deviation of the class means. The results are shown in Table 22, below. For second and third grades, data is from 107 and 75 classes, respectively, which participated in Joint Research Studies. This was necessary since relevant data from earlier Extended Pilot Tests are not available.

Table 22
Differences in Adjusted Means in Standard Deviation Units
MANS Process Categories

	Gr. 2	Gr. 3	Gr. 4	Gr. 5	Gr. 6	Average
Estimation	.35	.60	.54	.26	.42	.43
Mental Arithmetic	.45	.74	.57	.84	.49	.62
Numb. Reps.	.31	.36	.18	.43	.52	.36
Relats. & Numb. Patts.	.49	.67	1.64	.62	.74	.83
Word Problems	.06	.20	.67	.56	.45	.39
Elucidation	.21	.24		.86	.96	.57

The largest effects again were in Mental Arithmetic and in Relationships and Number Patterns, where the difference was usually 1/2 to 3/4 of a standard deviation. The averages given for Elucidation and Word Problems are somewhat deceptive; the effects were relatively small in second and third grades and relatively large in fifth and sixth grades.

A fourth method of comparing results across categories is to look at individual tests within a category and simply count whether or not the test produced a significant difference. Again the two categories containing the highest proportion of significant tests were Mental Arithmetic (17 significant results out of 20 tests) and Relationships and Number Patterns (16/22). Tests in each of the other categories were significant about half the time: Estimation (9/18), Number Representations (5/12), Word Problems (8/14), and Elucidation (3/6).

The rest of this section will describe findings in each category except computation, which will be described in the next chapter.

Mental Arithmetic

CSMP classes scored significantly higher than Non-CSMP classes in the Mental Arithmetic category at each grade level, and the differences were fairly consistent. Across all grades and test items, CSMP classes had an average of 60% correct versus 50% for Non-CSMP, and 17 out of 20 tests in this category the difference was significant. Figures 25 and 26 illustrate the findings. Figure 25, which shows third grade classes on the revised MANS, is a fairly typical result; the differences are large and obvious between CSMP and Non-CSMP classes. Figure 26, sixth grade class means, shows the least impressive results of any grade level, mainly because of the very poor showing of a few CSMP classes; nevertheless the overall results still clearly favor CSMP by a large margin.

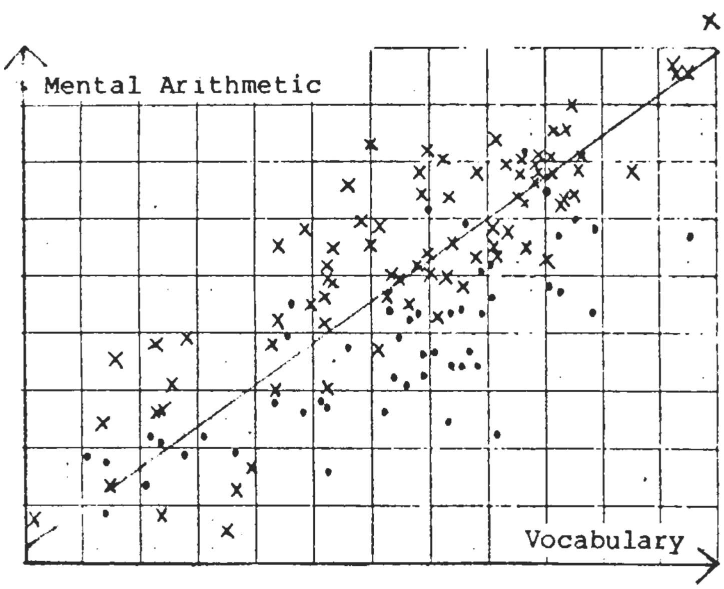


Fig. 25. Third Grade Class Means
Mental Arithmetic
Joint Research Studies
(x = CSMP Classes, • = Non-CSMP)

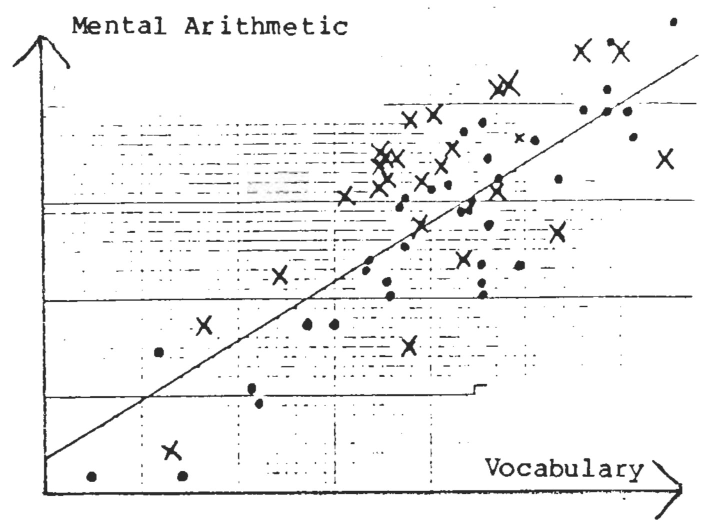


Fig. 26. Sixth Grade Class Means
Mental Arithmetic
Extended Pilot Tests
(x = CSMP Classes, • = Non-CSMP)

Whole Number Open Sentences. Eg. 9,001 + ? = 9,100. Doing scratch work was either discouraged or prohibited; students had to figure out these problems in their head. The box to be filled in could be on either side of the equal sign. The computational requirements were not heavy. For example, a problem like 7 X 63 would be inappropriate because the emphasis is on computation and memory, and partial results must be retained mentally for later processing. Most problems contained numbers which were multiples of 25, 50 and 100, so the arithmetic itself wasn't hard. But determining what operation to use and how to use it is not easy for elementary students. For example, only 71% of the fifth grade CSMP students and 58% of the Non-CSMP students gave the correct answer, 99, to the apparently easy example given above.

At every grade level, CSMP students had significantly higher scores on this test, and the differences were large, an average of about 60% versus 51% correct in favor of CSMP. In second and third grades, CSMP students' superiority came mostly from items involving multiplication or containing larger numbers (for example, in the hundreds). This result is unsurprising given CSMP's early emphasis on these concepts.

In grades 4-6, however, CSMP students continued to do better with problems involving multiplication; they also did better when division was required. This is very interesting considering that CSMP students do not do particularly well with straightforward multiplication and division problems, especially those involving algorithms. Consider the two results shown below.

$6\overline{)12,000}$	CSMP: 73% correct Non-CSMP: 79% correct	800 divided by ? = 200	CSMP: 60% Non-CSMP: 45%

Below is a list of a few typical items on which CSMP students did particularly well (grade levels are shown in parenthesis).

$\boxed{} \times 2 = 8$ (2)	$50 + 150 = \boxed{}$ (2)
$2 \times 400 = \boxed{}$ (2)	$35 + \boxed{} = 65$ (2)
$3 \times 125 = \boxed{}$ (3,4)	$525 - \boxed{} = 225$ (3,4)
$525 - \boxed{} = 225$ (3,4)	$\boxed{} + 125 = 250$ (3,4,5)
$\boxed{} \times 250 = 500$ (4,5)	$500 \div 2 = \boxed{}$ (4,5)
$4,999 + \boxed{} = 10,000$ (5)	$11 \times 273 = 3,003$ so $22 \times 273 = \boxed{}$ (5)
$\boxed{} - 250 = 150$ (5,6)	$\boxed{} \div 20 = 41$ (6)
$(8 \times 29)+(2 \times 29) = \boxed{}$ (6)	$12 \times 75 = 900$ so $13 \times 75 = \boxed{}$ (6)

<u>Fraction and Decimal Open Sentences.</u> These tests were similar to the whole number tests, except that they involved fractions or decimals and appeared only in third, fifth and sixth grades. On fraction open sentences, CSMP students had much higher scores than Non-CSMP students in third grade (50% correct versus 35%), and significantly higher scores in fifth grade (53% correct versus 43%). By sixth grade there was virtually no difference in scores. Items on which CSMP students did best were:

$1/2 \times \boxed{} = 10$

$1/2 \div 2 = \boxed{}$

In third grade, a similar scale was used, consisting of items requiring the calculation of 1/2 or 1/3 of number, for example, 1/3 of 15 = ? or 1/2 of ? = 16. CSMP classes did much better than Non-CSMP classes (an average of 50% correct versus 35%); this finding reflects CSMP's early emphasis on the partitioning aspect of fractions.

Decimal open sentences appeared only in sixth grade, with CSMP students scoring much higher (78% versus 60%) than Non-CSMP students. Typical items were:

0.75 - ? = 0.5 and 25 x ? = 2.5

Negative Hits and Misses. Students had to determine the missing pieces of information in the situation illustrated below. (During testing, the game was carefully explained, with samples problems, and a thermometer-like scale was available for use on the student page.)

	Each Hit Gain 5 points	Each Miss Lose 1 point	
Started with score of	Number of Hits	Number of misses	Ended with score of
0	1	1	?

Some form of this test was administered in each of grades 4-6. The sixth grade test was more difficult because each miss cost two points instead of one. At each grade there was a significant difference in favor of CSMP classes, who had an average of 65% correct answers versus 54% for Non-CSMP classes. The scale involves the concept of negative numbers, but this alone does not account for the difference. For example in both fifth and sixth grades, one of the items started with a non-negative score, with no hits and a number of misses sufficient to make the ending score negative; in both grades identical percentages of CSMP and Non-CSMP students got the item correct.

The CSMP advantage was greatest on items in which the missing information was something other than the ending score, as shown in some typical examples below.

Started With	Hits	Misses	Ended With
?	2	0	15 below zero
3 below zero	?	2	5 above zero
10 below zero	1	?	12 below zero

Hints and Problems. In this moderately speeded third grade test, students were given pairs of related addition or subtraction problems. The answer to one of the problems was given and students had to use that answer to figure out the answer to the other problem, as in the example given below. (Students were discouraged from trying to use an algorithm to calculate the answer and were not given much time to do this set of problems.)

$$\begin{array}{cccc} 538 & + & 198 & = \underline{736} \\ 539 & + & 199 & = \boxed{} \end{array}$$

CSMP classes had significantly higher scores on this test, 40% correct versus 32%.

Above and Below Zero. In the revised MANS in second and third grades, there was a short test required students to use negative numbers in the simple context shown below.

Score at the start: 5 below zero
then: won 2
Score at the end: 7 below zero   3 below zero   3 above zero   7 above zero

CSMP students did slightly better in second grade (average percent correct = 38 versus 35) but the difference was not significant. CSMP students did quite a bit better in third grade (mean percent correct = 53 versus 42) and the differences were significant in most Joint Research Studies.

Individually Administered Problems. In one of the problems administered in third grade, each student was shown a partial calendar with "69 cents" written under each day of the week and told that "Bill gets 69 cents every day this week". They were then asked to describe the fastest way, on a calculator, to figure out "how much Bill would earn by the end of the week".

CSMP students were more likely to suggest a multiplication process (88% versus 53%) and less likely to suggest an addition process.

Relationships and Number Patterns

CSMP classes had significantly higher scores on the Relationships and Number Patterns category at every grade level except second grade (where the difference approached significance). Sixteen out of 22 tests produced significant differences; except at second grade, almost all tests produced significance. Across all grades and test items, 65% of the CSMP responses were correct versus 55% for Non-CSMP.

Figures 27 and 28 illustrate the findings. Figure 27 shows second grade class means using the revised MANS in Joint Research Studies. In these more recent studies, CSMP performance improved to the extent that the difference was significant, as can be seen from the figure. Figure 28 shows fourth grade class means and needs no comment.

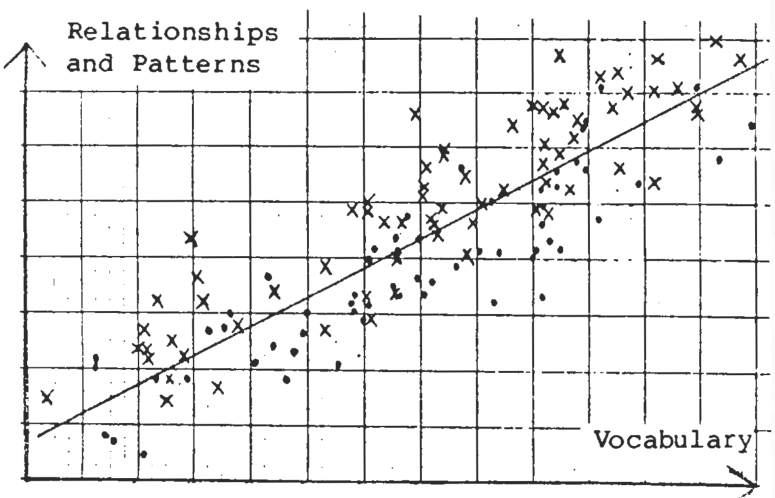


Fig. 27. Second Grade Class Means
Relationships & Patterns
Joint Research Studies
(X = CSMP Classes, ● = Non-CSMP)

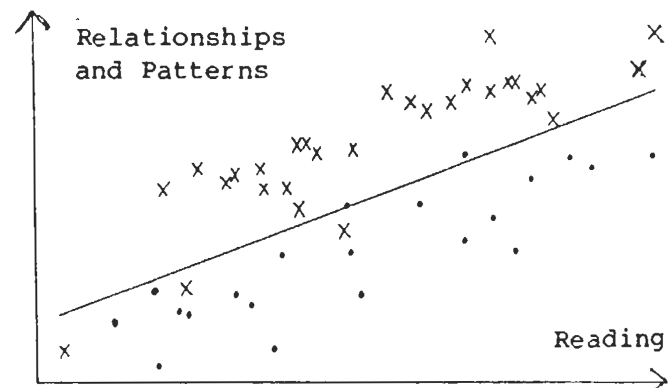


Fig. 28. Fourth Grade Class Means:
Relationships & Patterns
Extended Pilot Test
(X = CSMP Classes, ● = Non-CSMP)

Of all the categories, Relationships and Number Patterns produced the most consistent differences across ability levels. Figure 29 shows sixth grade student means when students were grouped into quartiles according to Vocabulary scores. Notice that the line segments joining the points are virtually parallel.

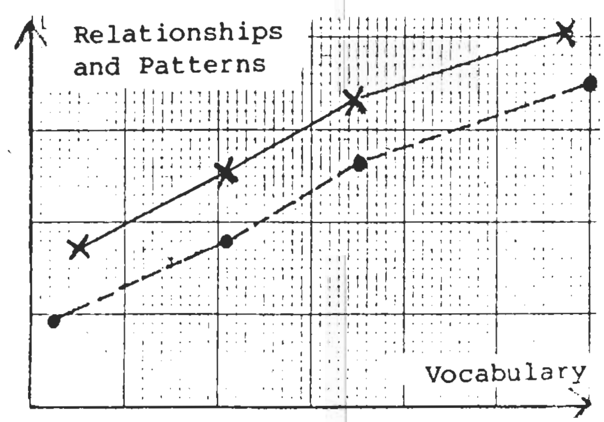


Fig. 29. Sixth Grade Student Means,
Relations and Number Patterns
Students grouped by Vocabulary Score
(X = CSMP Students, ● = Non-CSMP)

Solving Number Rules. This test has been used at all grade levels, though in different formats. The simplest to understand, pictorially, is shown below from fourth grade. Students were told that Machine A always did the same thing to any number that went into it; the first three rows gave examples of how the particular machine worked.

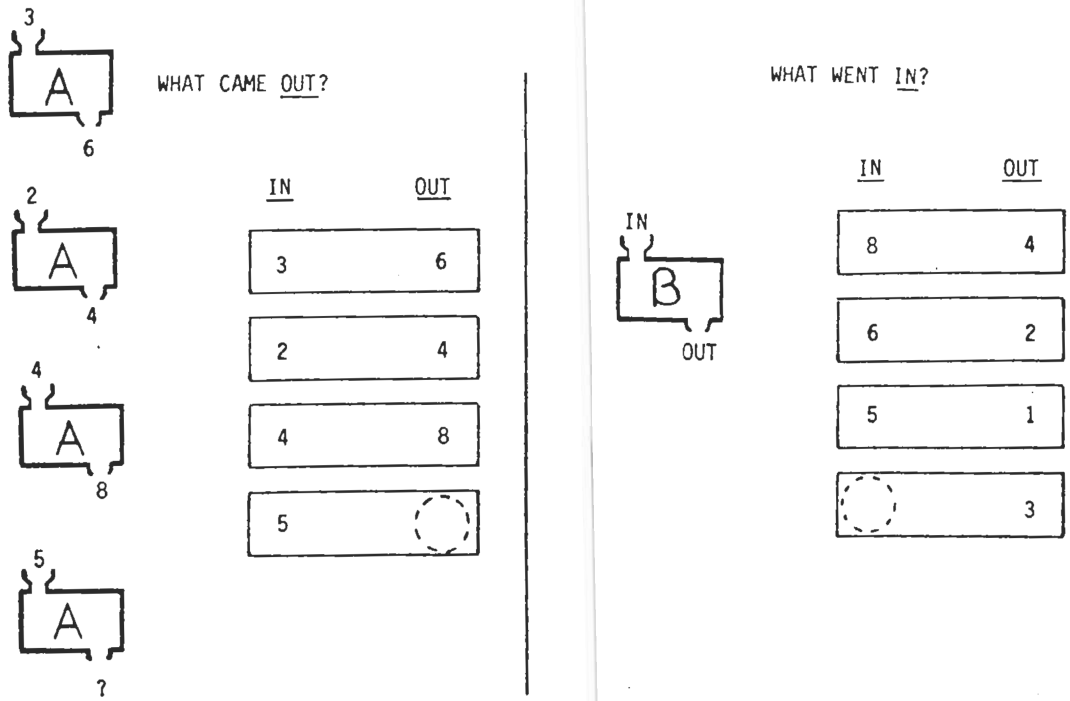


Figure 30. Items from Fourth Grade Test: Solving Number Rules
Left item = example (part of explanation),
Right item = test item

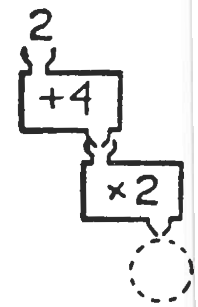
Students first had to figure out the common relationship between the given ordered pairs and then use that knowledge to figure out the missing entry output. The test got progressively more difficult by grade level. In third grade, the missing entry was sometimes an "input" rather than an output. In fifth grade, more complicated relationships were sometimes used; for example, the output number was one less than 10 times the input number. In sixth grade, some items used decimal numbers.

CSMP students always did much better than Non-CSMP students, their scores being 14%-25% higher (even in the non-significant second grade results).

CSMP students did better on all types of items; on every one of the 41 items in the various grades, CSMP students had a higher percentage correct. Their advantage was a little larger on multiplicative (versus additive) relationships and on to what might be called "two-stage" relationships. Examples are shown below, in abbreviated format.

2nd Grade	3rd Grade	4th Grade	5th Grade	6th Grade
2 - 1	6 - 12	6 - 3	36 - 6	100 - 304
4 - 2	2 - 4	16 - 13	100 - 10	0 - 4
8 - 4	7 - 14	8 - 5	81 - 9	10 - 34
10 - ?	? - 8	? - 9	? - 2	1 - ?

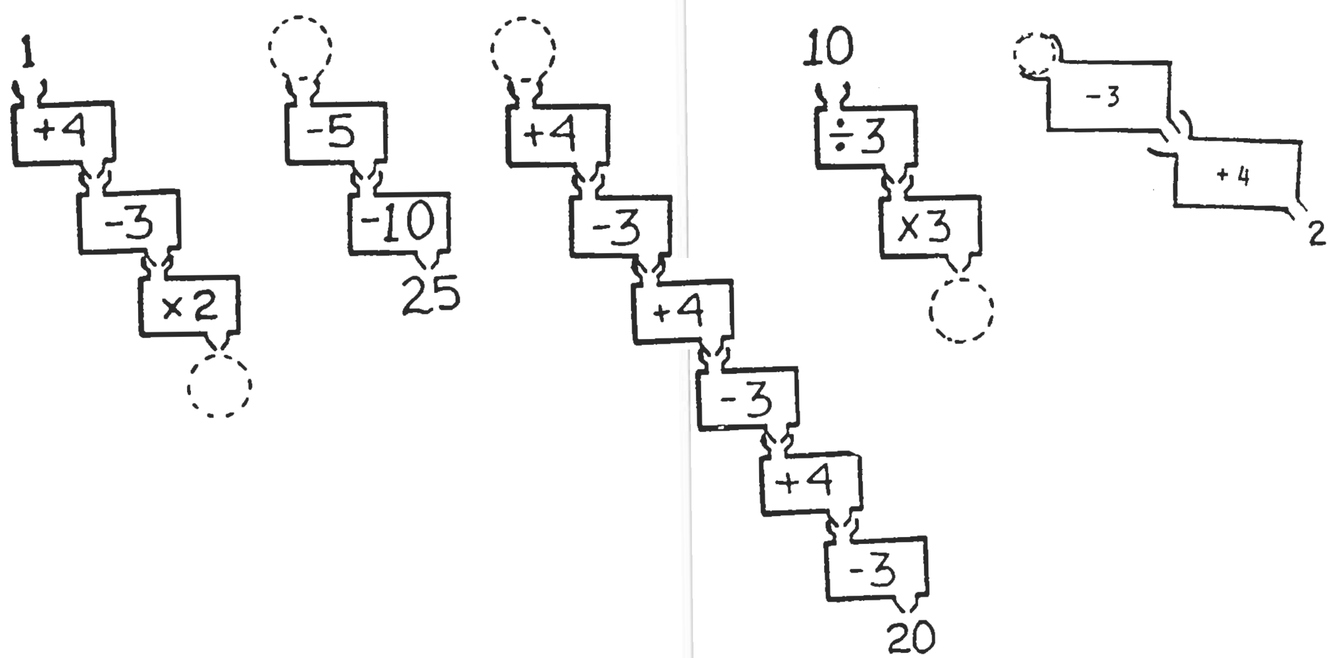
Using Number Machines. This test, also administered in one form or another in all grades except second, was positioned in the test booklet after the test described above and used the same concept of number machine. On this test, number machines could be hooked together as in this sample from the fourth grade test.



In the upper grades, most problems involved a missing input. In sixth grade, some items used decimal numbers.

CSMP classes scored significantly higher at every grade level, with an average of 62% versus 52% correct. They also had a higher percentage correct on every item at every grade level.

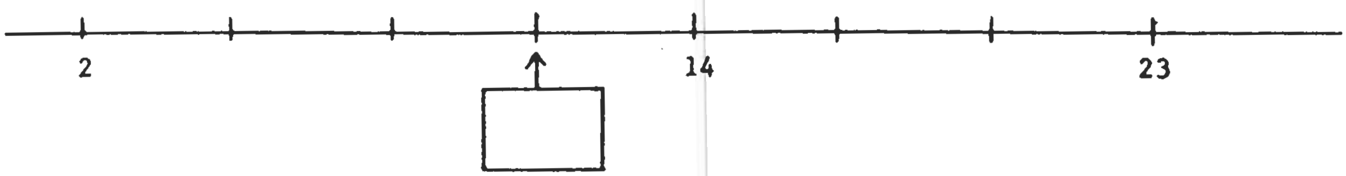
CSMP students did especially well, relative to Non-CSMP classes, on problems where combining machines (composing functions) was a better strategy than working one step at a time. The last three items below show examples of these problems. The first two are more general and straightforward problems.



Labelling Number Lines. Students had to label the empty box in a partially labelled number line, as in the example below from third grade.

This test was administered in all grades except fifth grade. More frequent use of interpolation and extrapolation were required in the upper grades, and some sixth grade items used decimal numbers.

CSMP classes classes got better at this with each passing year. Their scores were almost identical to Non-CSMP students in second grade, higher but not quite significantly so in third grade, and substantially and significantly higher in fourth and sixth grades (an average of 68% versus 49% correct). They did best, relative to Non-CSMP students, on items with large "gaps" to work with, requiring either interpolating or extrapolation, as in the examples shown below.



Which (whole number, fraction, decimal) is larger? In third grade, this test concerned whole numbers, for example 3 x 162 versus 4 x 160. Students were given only a few minutes to do many items so that computing exact answers was unwise (and strongly discouraged). CSMP classes did significantly better on this test and their advantage was greatest on items with fractions, and on subtraction items such as

500 - 201 versus 500 - 189.

In fifth and sixth grades fractions were used instead of whole numbers. CSMP classes had slightly higher scores each year (an average of 75% versus 70% correct); this difference was significant at sixth grade, but not at fifth. The largest difference was on an item with improper fractions: 5/2 versus 5/4.

In sixth grade there was also a test that used decimal numbers. CSMP classes had significantly higher scores than Non-CSMP classes, 82% versus 72% correct. The largest difference occurred on the following items:

6.1 versus 6.01 and 0.9 versus 0.111.

Additive and Multiplicative Series. In fourth grade, there were two scales dealing with series of numbers. The first one is illustrated below:

Counting by 98's

...						...
...	0	98	196	294	392	...

Will 492 be in any of the boxes? ____

Will 980 be in any of the boxes? ____

Will 690 be in any of the boxes? ____

CSMP classes had significantly higher scores, though the difference in scores was small (average percent correct = 63% versus 60%).

The other test in fourth grade concerned multiplicative series, as illustrated below:

...			50		200	400	800		...

The difference in scores was dramatic; average percent correct = 60% versus 32%. This test produced the largest CSMP - Non-CSMP differences out of all the tests administered in the Extended Pilot Tests. Figure 31 shows the graphs of class means for this test. The X's and dots are widely separated except for three low scoring CSMP classes; these classes were all from the same school and were the only classes from their district (a large urban district) which participated in this Extended Pilot Test.

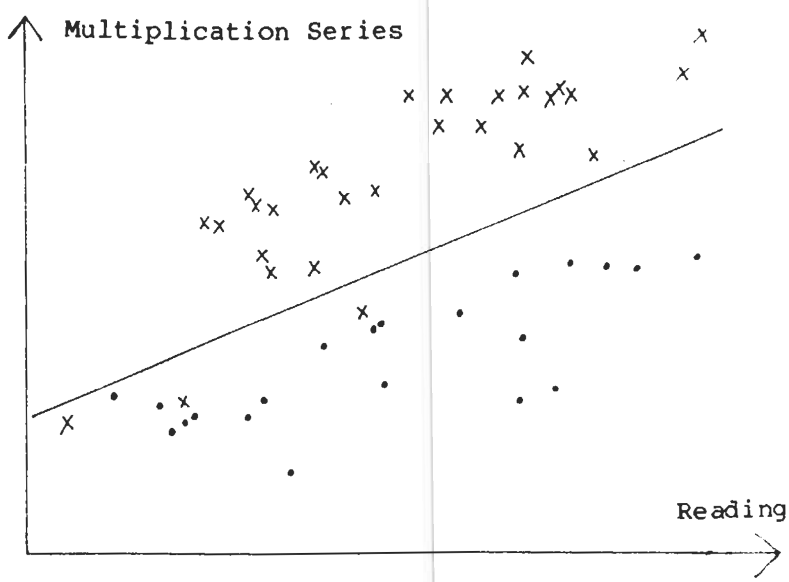


Fig. 31. Fourth Grade Class Means, Multiplication Series Extended Pilot Trials
(X = CSMP Classes, • = Non-CSMP Classes)

Fractions and Decimals Between Two Others. For these two tests, students had to produce a fraction (decimal) number, as shown by these examples.

_____ is larger than 1/10 but smaller than 1/3.

_____ is larger than 1.25 but smaller than 2.00.

On the test with fractions, there was virtually no difference between CSMP and Non-CSMP classes. On the test with decimals CSMP had significantly higher scores; mean percent correct = 89% versus 81%. These percentages are high because only two out of seven items required the student to extend the number of decimal places in the answer, and on these two items, shown below, CSMP students did much better than Non-CSMP students:

_____ is larger than 0.2 but smaller than 0.3.

_____ is larger than 0.42 but smaller than 0.43.

Sequences. In a second grade test, students had to determine the missing number in a sequence, such a the following:

7, 11, 15, ?, 23, 27

In the Extended Pilot Test, CSMP classes had higher scores (48% correct versus 42%) but the difference was not quite significant. In subsequent Joint Research Studies, the differences were larger and more likely to be significant.

Individual Administered Problems. In a fourth grade problem, students secretly drew a number out of a hat (but the interviewer knew that the number was 24) and answered a series of questions about their secret number. The questions dealt with concepts of order, whole numbers, negative numbers, multiples and divisors. The students were also asked whether the question itself was a good one. (For example, after finding out that the number was less than 100, a question about whether it was less than 200 was not a good question.)

CSMP classes had significantly higher scores, 82% correct versus 67%.

Elucidation of Multiple Answers

The second and third grade tests in this category differed considerably from the tests used in fifth and sixth grade with rather different results. In second and third grades students had to give as many answers as possible out of a potentially infinite number of correct answers, for example:

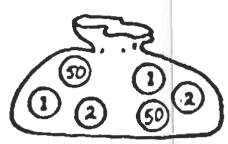
Sentences about 8, $8 = 5 + 3$
$8 = 2 \times 4$
etc.

Equations, using only these symbols: = + - x 1 2 3

eg. $2 + 2 + 2 = 2 \times 3$
$2 \times 2 = 1 + 3$
etc.

For each of the three tests in grades 2 and 3, CSMP classes produced about 6% more correct responses than Non-CSMP classes, a small but non-significant difference. Most of the difference occured at the higher ability levels.

In fifth and sixth grades, this category contained several problems, each of which had a number of correct answers (6-12) that would satisfy the given constraints. Altogether in the two years, a total of seven different kinds of problems were used and are described briefly below; in all cases students were to give as many possible correct answers as they could. The first problem is shown as it appeared on the student page; the others are shown in abbreviated form.



Rules: Take out three balls.

Add to get a total score. ~~50~~, ~~100~~

Give all the possible scores. 53, ________

start at zero, count by X's and reach exactly 24

pick out 3 balls, add the numbers, what total scores are possible?
(using a container with 6 balls numbered 1, 1, 2, 2, 50 and 50)

what whole numbers use only the digits 1, 2 and 3?
(no digit to be used more than once in a number)

what numbers are multiples of 2, and multiples of 3 and smaller than 50?

if P + P + Q = 7, what could P and Q stand for?

what whole numbers are even numbers, divisible by 5, and < 80?

CSMP classes did much better than Non-CSMP students on these tests; the average percents correct were the same at each level, 65% for CSMP versus 53% for Non-CSMP.

It was also true both years that the greatest difference in scores occurred at the lowest ability level, as illustrated in Figure 32, which shows average Elucidation scores for sixth grade students grouped according to reading score.

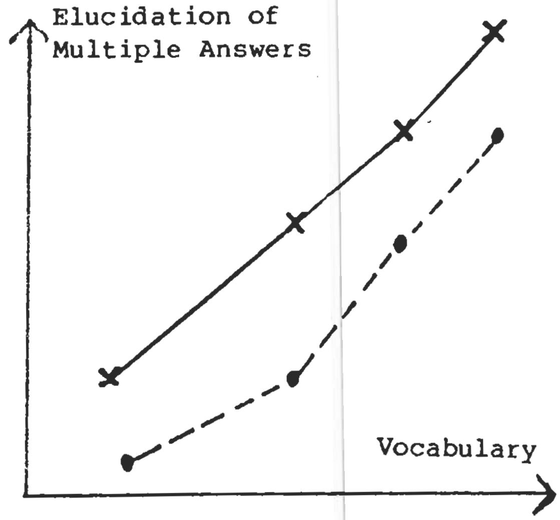


Fig. 32. Sixth Grade Student Student Means, Elucidation Students grouped by vocabulary score (X = CSMP students, ● = Non-CSMP)

In the revised version of the MANS, item formats like the ones used in fifth grade were extended down to third and fourth grades. Subsequent Joint Research Studies in those grades resulted in higher scores on these tests for CSMP classes, significant about half the time. The fluency format was retained in second grade with results similar to those found in the Extended Pilot Tests, i.e., slightly higher scores for CSMP classes, but not significantly so.

In sixth grade there was a test similar to the fluency-like tests in the lower grades. Starting from zero, and using any of the four operations with numbers 2, 3, 5 an 7, students were to construct sequences of calculations which would produce an end result of 12. For example:

$0 \ \underline{+\ 5} = 5 \qquad 5 \ \underline{+\ 3} = 8 \qquad 8 \ \underline{\div\ 2} = 4 \qquad 4 \ \underline{\times\ 3} = 12$

CSMP students gave about 35% more correct solutions, an easily significant difference, though it should be noted that this format is very similar to arrow diagrams (although students rarely chose to draw such diagrams in their booklets for this test).

Word Problems

Word problems of the kind found in textbooks and standardized tests (mainly one step, computationally oriented problems posed in sentence form) do not appear in the CSMP curriculum and teachers have commented on their absence. Nevertheless, the curriculum is saturated with mathematical problems (albeit in different formats), and CSMP students have usually done as well as or better than Non-CSMP students on the word problem sections of standardized tests.

The Word Problems category of the MANS Tests contained tests which were constructed on the basis of the kind of problem being posed. Thus, rather than a single long test containing different kinds of items, there were several short tests, each containing several items of single kind. The computation and reading skills needed to solve the problems were kept abnormally low.

Altogether there were a total of 14 tests administered in grades 2-5. With two exceptions, to be discussed later, the results were remarkably uniform, regardless of type of test. CSMP students always did a little better than Non-CSMP students; typical percentages correct were 55% versus 50%. These results were either barely significant or not quite significant. Out of the 12 tests, five were significant at .01, two were significant at .05, and for the other five tests the p-value was between .06 and .14.

Word Problem scores were fairly well predicted by Vocabulary scores. This meant that relatively small differences in mean scores could still be significant. Figure 33, below, shows fifth grade class means on the total of two Word Problem tests, dealing with two-stage and three-stage word problems. The adjusted mean scores favored CSMP by 6.6 versus 6.1, a fairly small difference which was nevertheless significant at the .03 level. The graph shows that most classes were represented fairly close to the regression line and there were few outliers.

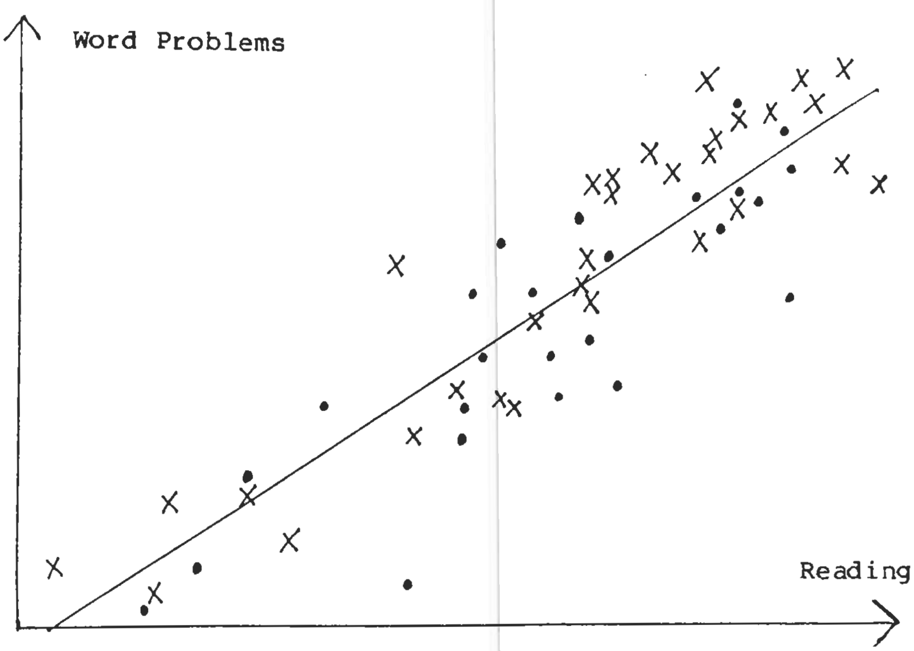


Fig. 33. Fifth Grade Class Means, Word Problems
Extended Pilot Tests
(X = CSMP Classes, ● = Non-CSMP)

Some of the tests were administered in more than one grade, though at least some of the items were different from grade-to-grade. Altogether, eight different kinds of tests were used and these are listed below with sample items.

One stage word problems.

> Mr. Rich lost $100 from his wallet.
> Afterwards he still had $200.
> How much did he have to begin with?
> (Said aloud by tester while second graders looked at cartoons.)

Two stage word problems.

> There are 40 apples in our barrel now.
> We will eat 2 apples every day.
> How many apples will be left in our barrel after 5 days?

Three stage word problems.

> Joe puts boxes into piles.
> Each box is 1/2 foot high.
> Each pile is 5 feet high.
> How many boxes does he need to make 3 piles?

Word problems with rounding.

> It takes 4 men to lift a piano.
> We have 14 men ready to work.
> How many pianos can they lift at the same time?

Word problem approximations.

> Martha can walk 2 blocks in 5 minutes.
> About how many blocks can she walk in 13 minutes?
> 5 blocks   10 blocks   15 blocks   18 blocks

Extraneous data.

> Sue has 12 bottles.
> It takes 36 bottles to fill a case.
> It takes 6 bottles to fill a carton.
> How many cartons can Sue fill?

Fractional word problems.

> 1/3 of a dozen eggs is ________ eggs.

Novel word problems.

> (This sixth grade test contained 12 miscellaneous problems including decimals, fractions, estimation, rounding, and 2-variable problems like the one below.)
>
> Steve has 7 bills.
> Some of them are $1 bills and some are $2 bills.
> Altogether he has $10 in bills.
> How many $2 bills does he have?

Two exceptions to the general overall pattern of small CSMP advantages were alluded at the beginning of this section. These exceptions concerned a test used in fourth and fifth grade involving decimals. The test consisted of a series of questions, all of which began "X has 6.5 gallons of gas". The item below happened to appear in both fourth and fifth grades.

> Joe has 6.5 gallons of gas.
> He uses up four gallons.
> How much gas will he have left?

CSMP students did much better than Non-CSMP students; an average of 46% versus 30% correct in fourth grade and 64% versus 50% in fifth grade. No doubt these differences reflect CSMP's earlier introduction to decimals.

In the revised MANS Tests administered in Joint Research Studies these two tasks were replaced by different tests which were simpler to administer. In second grade the test contained two types of items. One required the students simply to write a three-or four-digit number read aloud by the tester. The other required students to write the number that is 1 (or 10 or 100) greater than (or less than) a given number, for example, "What number is 10 more than 495?". CSMP did better (average percent correct = 59 versus 54) but the difference was not significant in most districts.

The revised third grade test required students to determine whether one number was 1 or 10 or 100 or 1000 more than another number. None of the answers was exactly correct. Students had several questions to do in a short time and were discouraged from calculating the exact answer. An example is given below.

$$4{,}265 \text{ is } \begin{matrix} 1 \\ 10 \\ 100 \\ 1000 \end{matrix} \text{ more than } 4{,}254$$

CSMP classes had higher scores (average percent correct = 50 versus 46) and the difference was significant in about half the studies.

In fifth grade, there were four tests dealing with fractions and one with decimals. The fraction tests contained the following tasks: marking fractions on a ruler, shading fractional parts of geometric figures, selecting equivalent fractions, and showing fractions on a number line. None of these tests produced significant differences. The decimal test required students to show metric distances and compare the size of decimal numbers. CSMP students did much better than Non-CSMP students on this test, average percent correct = 66% versus 50%.

In sixth grade, there were two tests, both of which produced significant differences in favor of CSMP. One was an omnibus test of fraction and decimal representations, and CSMP's largest advantage was on the decimal portion of that test. The other test required students to determine which fractions or decimals which were equivalent to a given fraction. This was similar to a test in fifth grade where the difference did not quite reach significance; in sixth grade it was barely significant ($p < .03$).

Estimation

Tests on estimation produced significant differences in favor of CSMP classes on 9 of 18 occasions. However, the relative differences were usually quite small; average percent correct across all tests and grades = 52% for CSMP classes versus 49% for Non-CSMP classes. Most of this difference was due to the strong performances by above average CSMP students. Figure 34, fifth grade student means grouped by reading score, shows an extreme example of this result. The line segments actually cross and at the lowest ability level, Non-CSMP students have higher scores. This crossing effect did not occur at any other grade level but the CSMP advantage was almost always smallest at the lowest ability level.

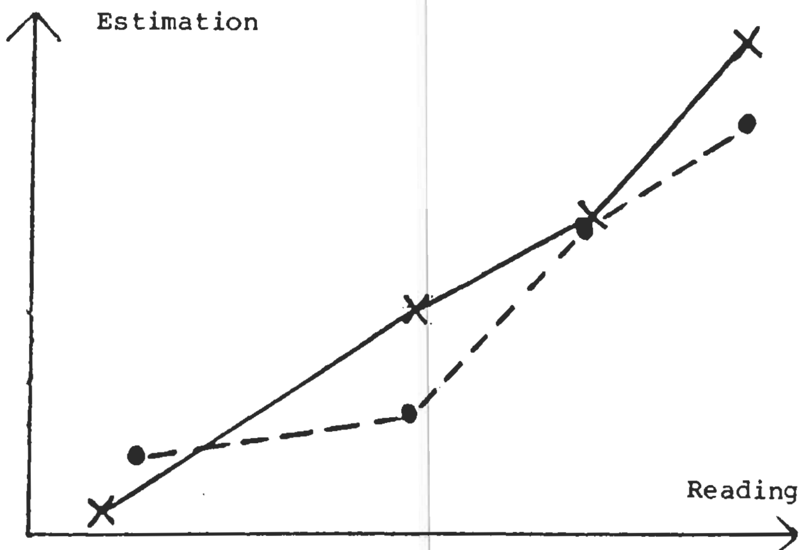


Fig. 34. Fifth Grade Student Means, Estimation, Students grouped by reading score (X = CSMP Students, ● = Non-CSMP)

Estimating Intervals. In what was by far the most common test used in this category, students had to respond to several computation items in a short period of time (an average of less than 15 seconds per item). For each item in the test there was a fixed set of intervals and students merely had to indicate which interval contained the answer. For example:

11 X 50 = 0 10 50 100 500 X 1,000

Only one arithmetic operation was used on any page, except in second grade where addition, subtraction and a couple of multiplication items were thoroughly mixed. Table 23 summarizes the results for each grade according to type of operation used.

Table 23
Mean Percent Correct, Interval Estimation Tests
First entry = CSMP, second entry = Non-CSMP
(* = sig. at .05)

Grade:	Addition	Subtraction	Multiplication	Division	Combined
2					34 - 32
3	54 - 53	30 - 31	50 - 43*		
4	77 - 73*		69 - 61*	45 - 40*	
5	83 - 79		74 - 67*	49 - 43*	
6			68 - 61*	51 - 48	

The table shows clearly that multiplication is the operation which produced consistent significant differences. The other operations produced modest differences, which were sometimes significant and sometimes not. The sixth grade CSMP advantages come almost entirely from items with decimal numbers.

Most of the multiplication and division items in fourth and fifth grade are the kind that would require an algorithm to find the exact answer. It is interesting to note that CSMP students were better than Non-CSMP students if the task was to estimate the answer to these items, but not as good if the task was to calculate the exact answer.

Other Estimation Tests. There were four other kinds of estimation tests used. In third and fourth grade, CSMP students were significantly better than Non-CSMP students (though the difference was relatively small) on a test with items like the following:

100 is about 2 or 5 or 10 times as large as 19?

In fifth grade, there were four tests, one for each operation, in which students had to select the best of three wrong answers. CSMP and Non-CSMP scores were virtually identical on all tests. A sample item is shown below.

$$15 \times 2{,}111 \quad = \quad \begin{matrix} 3{,}173 \\ 20{,}173 \\ 31{,}173 \end{matrix}$$

There was also a test in fifth grade, Measurement Estimation, in which students had to estimate quantities, volumes or areas from pictorial presentations. There was no difference in scores. It should be noted that, with respect to technical considerations (reliability, correlations, etc.) these fifth grade tests were among the worst ever produced by the evaluation staff.

In sixth grade, students had to estimate whether fraction computations, such as 1/2 + 4/7, were less than, equal to, or greater than one. Several items were given with a short time limit. CSMP scores were slightly higher than Non-CSMP scores, but the difference was not quite significant ($p < .06$).

Individually Administered Problems. Two kinds individually administered problems in third grade produced significant differences in favor of CSMP students. In one kind, students were shown a set of completed calculations which "a student at another school" had done (e.g. 6 X 13 = 53). They were then asked to rapidly indicate which answers "could be right" and which ones were "probably wrong". Finally students were asked to go back to each probably-wrong answer and tell why they thought the given answer was wrong.

CSMP students made a higher average number of correct decisions (70% versus 64%) and their explanations of wrong answers were more likely to be acceptable (89% versus 77%). The largest differences between CSMP and Non-CSMP students occurred for students of about average ability.

In the other kind of individually administered problems, students were asked to quickly estimate the number of dollar bills that would be needed to purchase seven items whose costs were as shown below, "but we don't want to take any more (money) then we'll need":

$$\begin{array}{r} \$1.22 \\ 1.81 \\ 1.51 \\ 1.53 \\ 1.33 \\ 1.33 \\ \underline{1.39} \end{array}$$

A higher proportion of CSMP students (50% versus 34%) gave good answers, defined as 10, 11 or 12, and a lower proportion (12% versus 25%) gave poor answers, i.e., <8 or >14.

Special Topic Categories

A total of fifteen tests were administered in the five special topic categories. Most were administered in sixth grade. The results are summarized below, in Table 24.

Table 24
Summary of Special Topic Category Results

Category	Grade	Number of Items	Adjusted Means CSMP	Adjusted Means Non-CSMP	Significant at .05
Algebra	6	12	7.4	6.2	*
	6	10	9.3	8.9	*
	6	10	4.4	3.8	*
	6	14	9.2	8.8	
Geometry	4	8	4.7	4.5	
	6	6	3.0	3.3	* (in favor of Non-CSMP)
Logic	6	18	13.1	12.9	
	6	18	10.2	10.4	
Organization of data	3	6	3.3	3.4	
	5	10	6.1	6.1	
	6	10	5.8	5.7	
Probability	4	19	6.9	5.7	*
	5	6	3.7	3.5	
	6	12	9.3	8.6	*
	6	10	5.9	5.4	*

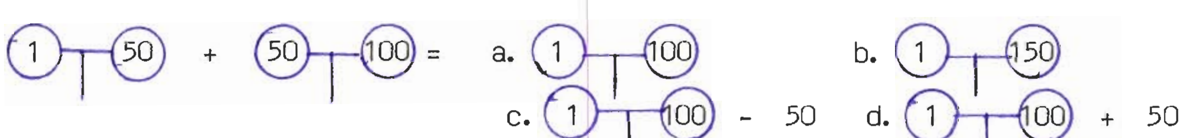
CSMP classes did better in the Algebra and Probability categories, with significantly higher scores on three of the four tests in each category. Scores on the other seven tests were virtually identical except for one Geometry test, where Non-CSMP classes had significantly higher scores.

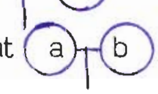
Algebra. Typical items from the three Algebra tests which produced significant differences are shown below.

If $g = 4$ and $h = 3$, then $5gh$ =? (students read 2 examples, including one showing that 3bc means $3 \times b \times c$)

If $q = 5$, then $2 \times q^2$ = ? (students read 3 examples explaining exponentiation)

If $k + 2 + k + 1 = 7$, then k = ? (students read 3 examples)



(The tester gave an explanation, through examples, to show that 

is the sum of a, a + 1, ... b, i.e., it is the summation operator.)

The following examples are from the fourth algebra test, which produced a non-significant difference in favor of CSMP.

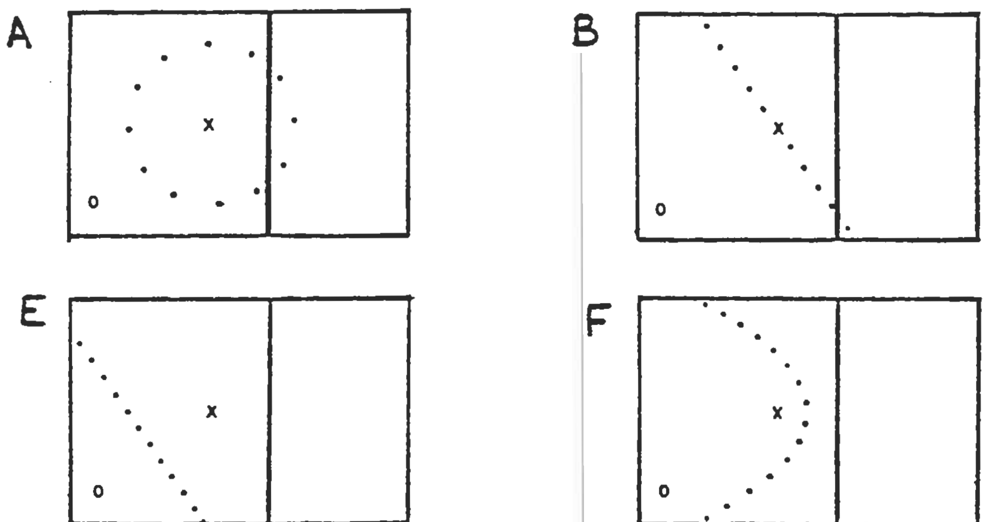
(xx o) = ?

do 9 times to o o X

(where has been shown, by examples, to be a 90° rotation and reverses the number of elements in the top and bottom row.)

<u>Geometry</u>. On a fourth grade test, in which students had to divide various geometry figures into congruent parts (e.g., an equilateral triangle into four congruent parts), there was no significant difference between CSMP and Non-CSMP classes.

In the revised MANS, this test was placed in fifth grade and a new test was used in fourth grade. Students were required to select the one picture (out of several pictures) which satisfied certain conditions, as in the example shown below.



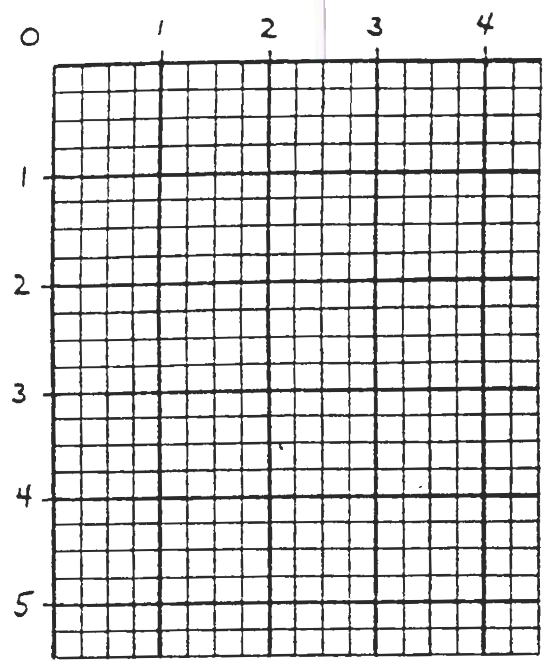
4. In which picture is each dot closer to x than to o? A B E F

5. In which picture is each dot just as close to x as to o? A B E F

These problems are about finding the locus of a point. In two districts which administered the revised fourth grade MANS there was virtually no difference in scores between CSMP and Non-CSMP classes. In the third district, Non-CSMP classes had higher scores (average percent correct = 73 versus 63) but the difference was not significant.

On a sixth grade test, students were given a page showing nine geometric shapes: two triangles, a square, rectangle, rhombus, hexagon, parallelogram and an "open" triangle and a rectangle. They were asked to study the figures, mark figures which were alike in some way, and explain why they were alike. Non-CSMP classes produced significantly more acceptable categories than CSMP classes - about 10% more. This difference was significant at .05 and was the only MANS Test at any grade level in which Non-CSMP students had significantly higher scores.

One of the individually administered problems in fourth grade produced a significant difference in favor of CSMP students. Students were given sheets of graph paper, with different ways of labelling the lines and some lines heavier than others. An example is shown on next page.



CSMP students were better able than Non-CSMP students to figure out how many little squares were shown, were more likely to use a length-times-width method, and were more likely to use the guide numbers in the margins versus a one-at-a-time counting process. They were also better able to do related problems of figuring out the area when pieces were combined or when one of the figures had a "hole" in it.

On a slightly different problem, CSMP students were better able to figure out how many squares were on a partly hidden role of paper marked off at every second square.

Logic. Two tests in Logic were administered in sixth grade and both produced almost identical scores for CSMP and Non-CSMP classes. In a typical problem from the first test, students were told there were six boys, each of whom played one of six sports. Students then had to use the given clues to figure out which boy played which sport. In the other test, students were to select or construct a situation which would make a given statement false. (The given statement concerned the placement of various geometric shapes above or below a line, as shown in a picture on the student page.)

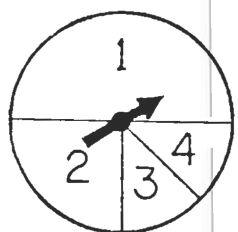
In two individually administered problems in third grade, CSMP students performed significantly better than Non-CSMP students. In one, students were shown an undifferentiated set of "people pieces" (simplified figures that were either tall or short, fat or thin, boy or girl, and red or blue). They were then asked to put them in piles so that all the pieces in a pile were similar in some way and so that the piles were all different from one other. They performed this classification in as many different ways as they could. CSMP students were able to make more complex sorts than Non-CSMP students, the average "best effort" being 3.0 dimensions simultaneously (versus 2.2 dimensions for Non-CSMP students).

In the other individually administered test, students were asked to figure out the interviewer's "secret" rule for the people pieces, by offering individual pieces to which the interviewer would respond with a "yes" or "no" according to whether the offered piece fit the secret rule. Examples of the secret rule were "blue" and "fat and tall". CSMP students needed to offer fewer pieces to figure out the rule. In four trials, the average total number of pieces needed was 14.8 for CSMP students versus 19.7 for Non-CSMP students.

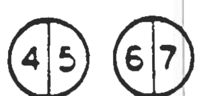
Organization of Data. The three tests in this category, administered at different grade levels (grades 3, 5 and 6) each produced almost identical scores for CSMP and Non-CSMP classes. Each test involved the reading and interpreting of data from a table (grades 3 and 6) or graph (grade 5). In fifth grade, some interpolation was required and in sixth grade some of the items required extrapolation.

Probability. Typical items from the three Probability are described below. CSMP classes did significantly better on each of the tests.

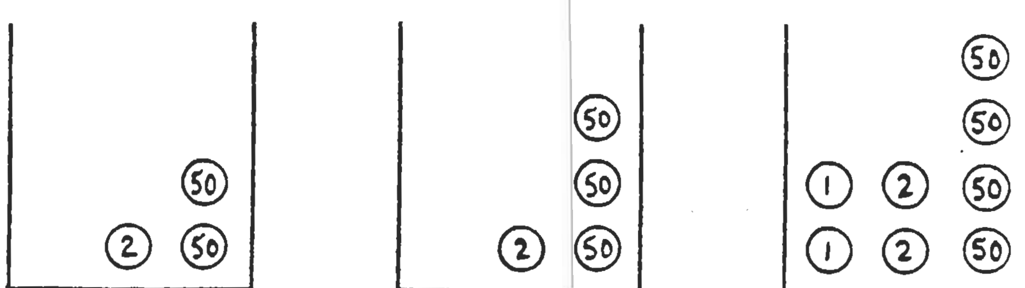
Students had to estimate how many times out of 100 spins they would get a particular result on a spinner. Spinners were divided into unequal, but easily calculable regions such as the one shown below. A range of answers was accepted for each question. (Fourth grade)



Students had to determine how often (never, less than half the time, half the time, etc.) a pair of spinners would land on numbers whose sum was at least 10. (Sixth grade) Pairs of spinners divided in various ways were used, for example:



Students had to select the best of three given boxes from which to make a blind draw. The boxes contained differing numbers of 1-cent, 2-cent and 50-cent balls. This test was administered in both fifth and sixth grades, producing a significant difference in sixth grade only.



Discussion

CSMP students performed best, relative to Non-CSMP students, on tests in two categories; Mental Arithmetic, and Relationships and Patterns, and worst in Estimation (that is, they were only a little better in Estimation). The categorization scheme used with the $\overline{\text{MANS}}$ Tests is one of several possible ways of organizing the testing and reporting of student learning. It has turned out to be a useful scheme and seems to convey the process orientation of the tests. But it may not be the most useful scheme for discussing the strengths of CSMP.

There are a number of fundamental processes and concepts at which CSMP students exell and which cut across categories.

1. Inverse operations. All of these problems share an aspect of having to think backwards or find an initial condition which will produce a given final result.

   $\boxed{} - 250 = 150$

   $1/2 \times \boxed{} = 20$

   Negative Hits and Misses, where beginning score instead of ending score is required

   Word Problems, such as (paraphrased), starting with \$10 and saving \$5 a week, how many weeks before one can buy a radio for \$30.

2. Recognizing numerical patterns. Examples are additive sequences, multiplicative series, partially labelled number lines requiring extrapolation, and multiples and divisors in Elucidation problems.

3. Relations. CSMP students seemed to understand the concept of relation better, that is, the independent existence of, say, +3 as concept, a thing in itself, that doesn't need some particularization (for example, $2 \underline{+3} = 5$) to give it meaning.

   This understanding was demonstrated most clearly in tests on solving number rules and using number machines. There is a sense in which these two tests are biased towards CSMP students, but any test dealing with relations would probably be biased in that sense. The concept of relations is such a fundamental one in mathematics that such criticism is not worth worrying about.

4. Relative sizes of numbers. Examples are selecting the larger of two whole numbers or decimals or fractions. For example, without actually calculating, which is larger:

   500 - 201 or 500 - 199 ?

5. Early presentation of concepts. CSMP students are introduced to the concepts of multiplication, negative numbers, fractions, and decimals earlier than most students and they are better able to apply this knowledge in a variety of situations.

6. Using intermediate answers. Examples are:

$$538 + 198 = 736$$
$$\text{so } 539 + 199 = \ ?$$

$$11 \times 273 = 3{,}003$$
$$22 \times 273 = \ ?$$

These are all very important processes in mathematics and the CSMP curriculum contains many instances of each of them. They are never formally presented or named, just used over and over in different ways in both teacher presentations and student materials. Together they make for what one might call "street number sense", and CSMP students seem to have it. What is surprising is that these processes have a heavy computation component, thus making the CSMP advantage on them particularly noteworthy since CSMP students are not particularly strong on straightforward computation. This may explain why the CSMP advantage on Estimation, very much a street sense attribute, is rather modest; although CSMP does emphasize some aspects of estimation, that skill is so computationally dependent (or possibly part of a very deep-seated quantitative trait) that large gains should not be expected.

In the special topics categories, the CSMP advantage in Algebra, which incorporates concepts of variables and transformations, is not surprising since these concepts arise in several ways in CSMP. Similarly, CSMP students should do better in Probability and they do. The two sixth grade tests in Logic produced no CSMP - Non-CSMP differences, meaning that CSMP's informal logical thinking, as in the string game for example, do not transfer to the more formal paper and pencil MANS items.

In Geometry, CSMP students did no better, and on one test, significantly less well than Non-CSMP students. The three MANS geometry tests were very general kinds of problems dealing with locus, congruency, and creating geometric categories, none of which were particularly stressed by CSMP. No doubt a test more oriented to the specifics of the CSMP curriculum in geometry would have produced rather different results.

The same could be said for tests oriented to other specific CSMP content such as negative numbers, modular arithmetic, binary numbers and other number theoretic work. Such tests would have been against the spirit of the MANS tests, which avoided terminology and content specific to CSMP. However, the absence of such tests, under whatever rubric, was a mistake. It leaves any CSMP reviewer in the position of suspecting there are many specific pieces of mathematical content that CSMP students know better than Non-CSMP students, but not knowing for sure. In this sense, the evaluation was conservative and underestimated CSMP student learning.

VI. COMPUTATION AND STANDARDIZED TEST RESULTS

Computation Results

A considerable amount of data has been collected on CSMP students' computation skills. Data will be presented from three sources: standardized tests administered as part of the Extended Pilot Tests, specially constructed computation tests incorporated into the revised MANS for administration in subsequent Joint Research Studies, and district-initiated standardized test comparisons.

Standardized Computation Test Results from Extended Pilot Tests. Table 25 summarizes the data from standardized computation tests administered as part of the Extended Pilot Tests. Unless otherwise indicated, the scores were from the Computation subtest of the standardized test. In second and third grade, separate studies were conducted in each district since the MANS Tests did not include a computation section. Thus each comparison shown for second and third grades represents one district. In fourth and fifth grades, districts were combined in the analysis since the MANS Tests contained a standardized computation test. In all cases, an Analysis of Covariance on class means was used, with class mean score on Reading or Vocabulary as covariate.

Table 25
Summary of Standardized Computation Scores
from Extended Pilot Tests

Grade	Test	Number of Classes		Adjusted Means		Sigif	In
		CSMP	Non-CSMP	CSMP	Non-CSMP	at .05	Favor of
2	CTBS	15	13	20.6	19.3	N	CSMP
	CTBS	6	6	20.5	18.6	N	CSMP
	ETS Coop Prim[1]	6	6	36.2	35.2	N	Non-CSMP
	Stanford Ach	6	6	47.0[2]	52.0[2]	N	Non-CSMP
3	CTBS	12	12	36.1	35.4	N	CSMP
	CTBS	15	13	35.8	36.8	N	Non-CSMP
	ETS Coop Prim [1]	6	12	42.3	41.8	N	CSMP
4	SAT	30	21	23.3	25.3	Y	Non-CSMP
5	CTBS	31	25	34.9	34.3	N	CSMP

[1] Total Math: This test does not have a separate computation section.
[2] Percentile Ranks

There a total of nine comparisons given above, and three different tests were used. Five of the comparisons favored CSMP and four favored Non-CSMP. Only one of the nine produced a significant difference: in fourth grade, Non-CSMP classes had significantly higher scores than CSMP classes on the Computation Test of the Stanford Achievement Test.

Figure 35 shows the graphs of class means for fourth grade classes from the Extended Pilot Test. It can be seen that the Non-CSMP advantage is due to the relatively poor performance of CSMP high ability classes.

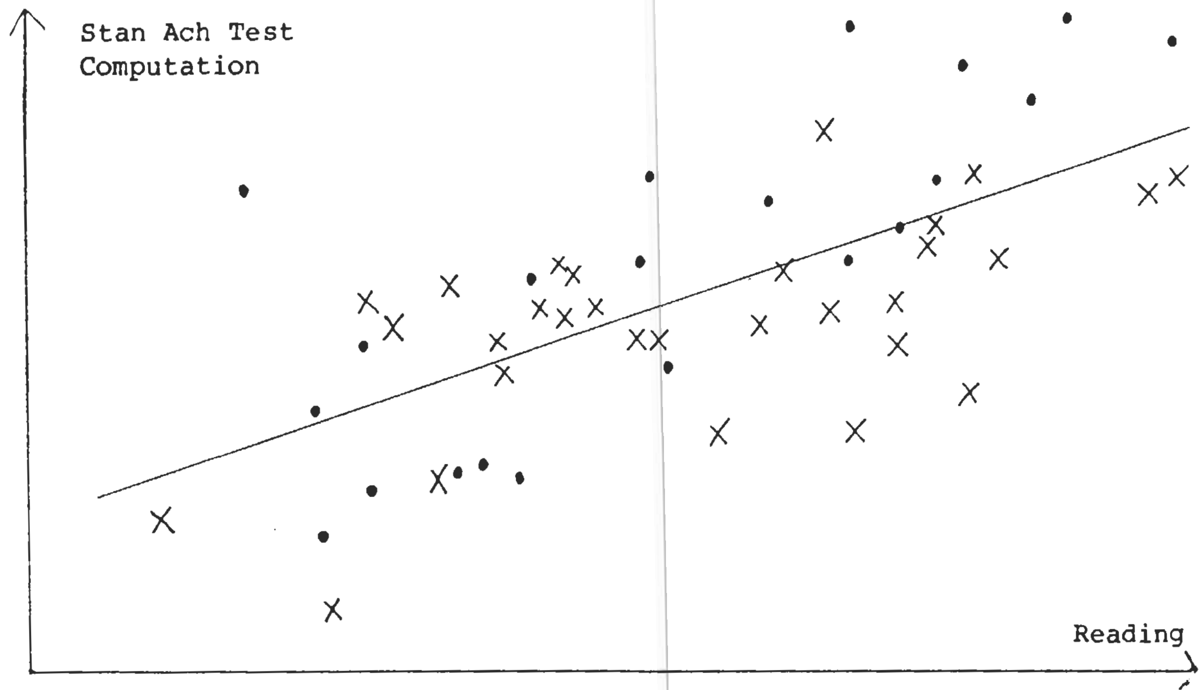


Fig. 35. Fourth Grade Class Means, Computation
Extended Pilot Test
(X = CSMP class, ● = Non-CSMP class)

Figure 36, below, shows the graphs of class means for fifth grade classes from the Extended Pilot Test. There is no discernable pattern between CSMP and Non-CSMP scores; in fact, computation score is not very well predicted by vocabulary score.

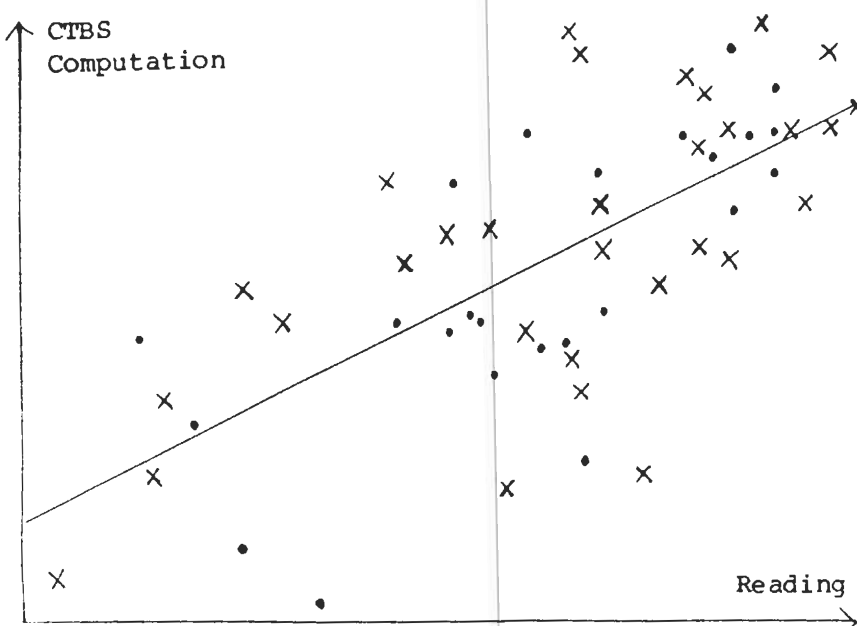


Fig. 36. Fifth Grade Class Means, Computation
Extended Pilot Test
(X = CSMP class, ● = Non-CSMP class)

Computation Results from Revised MANS Tests. In the revised MANS administered in Joint Research Studies and in the EPT sixth grade MANS Tests, a computation test was developed in order to reduce testing time and eliminate royalty costs. The items were restricted to whole numbers and selection was based on an analysis of the type and frequency of items found in the leading standardized tests at each grade level.

Table 26 summarizes the computation results from the revised MANS. Except for sixth grade, each row of the table shows results for a single district. Analysis of Covariance on class means was used each time with Gates McGinitie Vocabulary score as covariate.

Table 26
Summary of Revised MANS Computation Scores

Grade	Number of Classes CSMP	Non-CSMP	Adjusted Means CSMP	Non-CSMP	Sigif at .05	In Favor of
2	3	3	15.2	13.4	Y	CSMP
	10	10	10.0	10.5	N	CSMP
	4	5	9.1	10.2	N	Non-CSMP
	21	26	12.6	11.8	N	CSMP
	5	5	12.1	12.6	N	Non-CSMP
	7	4	11.3	10.1	N	CSMP
	2	2	15.2	14.0	-	CSMP
	2	3	12.1	9.3	-	CSMP
	2	2	10.0	11.4	-	Non-CSMP
3	20	26	17.2	16.7	N	CSMP
	5	4	15.7	18.0	N	Non-CSMP
	13	6	16.1	15.3	N	CSMP
	4	3	20.6	19.6	N	CSMP
	33	13	17.2	16.8	N	CSMP
	2	2	17.5	15.0	-	CSMP
4	23	23	21.3	21.5	N	Non-CSMP
	6	6	21.0	19.9	N	CSMP
	5	7	19.3	19.5	N	Non-CSMP
6	27[1]	36[1]	18.9	19.2	N	Non-CSMP

[1] Sixth grade EPT; classes from several districts combined into a single study

There are a total of 19 comparisons given above. Twelve of the comparisons favor CSMP and seven favor Non-CSMP. Only one of the nineteen produced a significant difference, a second grade comparison favoring CSMP.

Figure 37 shows class means in Computation for all second grade classes participating in Joint Research Studies.

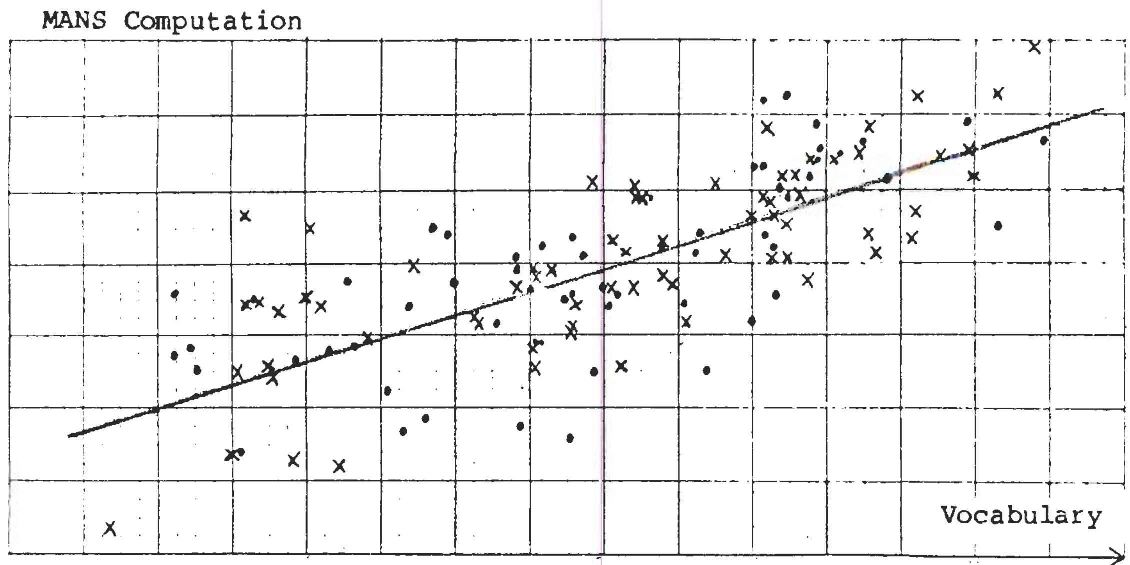


Fig. 37. Second Grade Class Means, Computation
Joint Research Studies
(X = CSMP class, ● = Non-CSMP class)

Figure 38 shows third grade computation means from Joint Research Studies. Although there is a slight CSMP advantage overall, it is hard to discern from the graph. Computation scores are poorly predicted by vocabulary scores; for lower ability classes especially, there is great variation in scores.

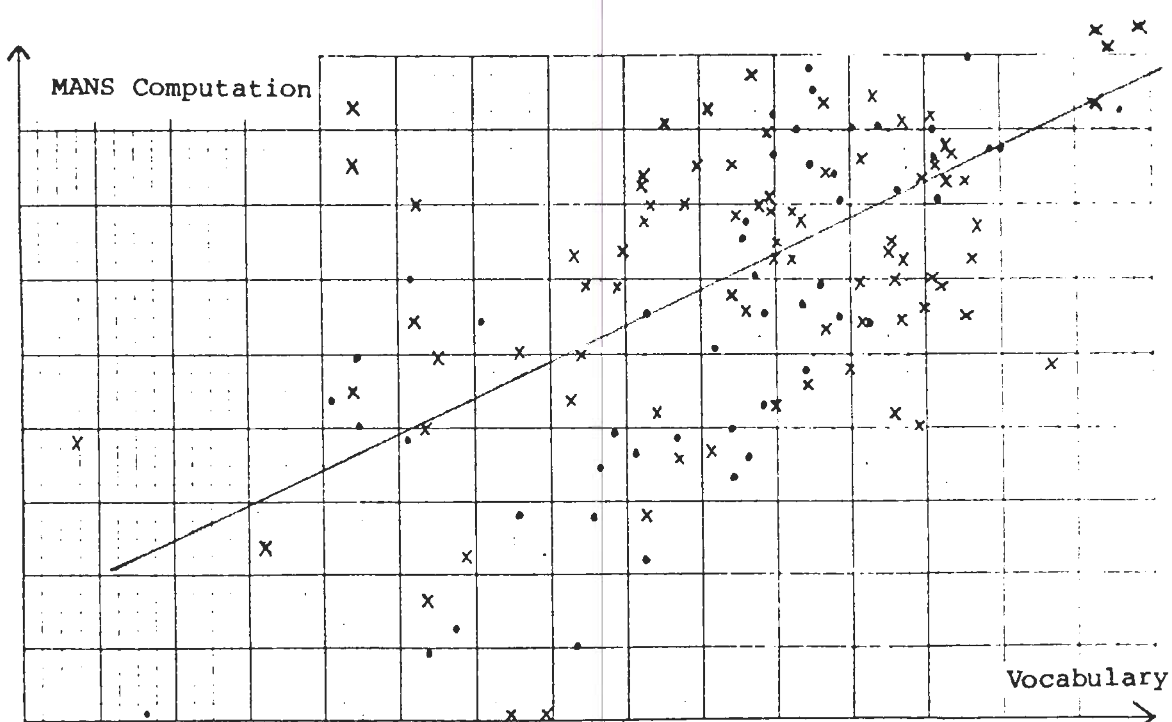


Fig. 38. Third Grade Class Means, Computation
Joint Research Studies
(X = CSMP class, ● = Non-CSMP class)

Computation Results from Locally Conducted Evaluations. In a number of sites, district personnel conducted a formal evaluation of CSMP by comparing the performance of CSMP and Non-CSMP classes (or students) on their district-administered standardized achievement test. No doubt many more evaluations were carried out than could be located for this report, and certainly many informal evaluations were completed and never officially reported.

Table 27 summarizes the data from those districts which reported separated computation scores. Different methods of aggregating and analyzing the data were used at each site, and significance tests were not generally reported.

Table 27
Summary of Computation Scores
from Locally Conducted Comparison Studies

District	Grade	Test	Comparison	Mean Score CSMP	Mean Score Non-CSMP
19	2	CAT-Comp	17 CSMP classes versus same teachers' previous Non-CSMP classes	277	280[1]
4	3	CAT-Comp	100 CSMP and Non-CSMP students sampled from 6 schools	60	51[2]
17	2	CTBS-Comp	6 CSMP classes versus same teachers' previous CSMP classes	307	312[3]
17	3	CTBS-Comp	Same as above	376	367[3]
17	4	CTBS-Comp	Same as above	394	388[3]
8	4	ITBS-Cr Ref SRA-Cr Ref	16,17 CSMP and Non-CSMP students sampled from 2 schools	69 81	68[4] 92[4]
8	5	ITBS-Cr Ref SRA-Cr Ref	20-24 CSMP & Non-CSMP students sampled from 2 schools	63 69	74[4] 91
12	6	ITBS-Comp	5 year longitudinal study of 70 CSMP versus 90 Non-CSMP students	6.4	7.0[5]
12	7	ITBS-Comp	same as above	7.8	7.9[5]

[1] Mean scaled score across classes;
an approximately equal decline occurred in Reading.
[2] Gain, from previous year, in mean student scale scores.
[3] Adjusted mean scaled score across classes.
[4] Average mean percent correct across items.
[5] Mean grade equivalent scores across students.

Altogether, five districts conducted nine studies, two in each of grades 2-4 and one in each of grades 5-7. Of these nine, three gave results favorable to CSMP and six favorable to Non-CSMP. Only three studies produced large differences: District 4 (in favor of CSMP), and District 8, grade 5 and District 12, grade 6 (in favor of Non-CSMP). Three studies using Total Math score (reported in the next section), in which computation was one component, produced virtually no differences.

Fraction and Decimal Computation. In each of the grades 4-6 Extended Pilot Tests, the MANS Tests included a short test of 6-10 items requiring straightforward computation with fractions. CSMP classes had higher scores at each grade level; much higher by about 23% in fourth and fifth grades, and slightly higher (though not quite a significant difference) in sixth grade.

CSMP students were much better at taking fractional parts of whole numbers (one-nth of a number) and anything involving commonly used fractions such as 1/2, 1/3, 1/4, and 1/1. On the other hand CSMP students were not as good at working with fractions of unlike denominators, i.e., the algorithmic part of fractional computational skills. These results are also consistent with CSMP's curricular emphasis.

In the sixth grade Extended Pilot Test, the MANS Tests included a short test of straightforward decimal computations on which CSMP classes had significantly higher scores than Non-CSMP classes.

Discussion. Among these three kinds of comparisons, a total of 37 studies were conducted. A total of 32 districts participated, either in separate studies (20) or as part of larger studies (12). Twenty results favored CSMP and seventeen favored Non-CSMP. Only 6 of the 37 studies produced significant or "large" differences, three in favor of CSMP and three in favor of Non-CSMP. Thus one can safely say that, overall, CSMP and Non-CSMP classes performed about equally on tests of computational skills.

However, if one analyzes the results separately by primary grades (2,3) and intermediate grades (4-6), the results are somewhat different. In the primary grades, 17 of the 24 studies favored CSMP, including all of the studies producing large differences in favor of CSMP. In the intermediate grades, nine of the twelve results favored Non-CSMP, including all three of the studies producing large differences in favor of Non-CSMP. It is still true that most studies, regardless of grade level, produce little or no CSMP-Non-CSMP difference, but there is some indication of better CSMP performance in the lower grades and poorer CSMP performance in the upper grades.

In addition, based on analysis of Extended Pilot Test data, there were certain computation skills which CSMP students were better at than Non-CSMP students and there are other skills in which they were worse. Furthermore, the pattern of these differences was consistent with the differences in curricular emphasis.

In second and third grades, CSMP students were a little better in addition and multiplication, and a little worse in subtraction, especially when it required borrowing.

In fourth grade sharper differences became apparent. There were no differences between the two groups on addition and subtraction questions, nor on one-digit multiplication and division questions (i.e., basic facts). But CSMP students did significantly worse than Non-CSMP students on multiplication and division questions containing multi-digit numbers and requiring an algorithm.

This difference persisted in fifth grade, though it was a smaller difference and counter balanced by better CSMP performance on items involving column addition and decimals.

By sixth grade the difference between the two groups' performance was very small, never more than 5 percentage points on any item. But CSMP students were still better in addition, worse in division, and once again worse in subtraction.

These findings are consistent with the differences between the CSMP curriculum and what is in most standard mathematics text books. The multi-digit algorithms for multiplication and division are introduced later in CSMP, are not taken to the "final" efficient form that most students finally learn, and are practiced less often.

Most teachers recognized CSMP's slower and later emphasis on these algorithmic skills and supplemented the program accordingly to remedy the problem. The amount of supplementation affected class performance on computation tests. In fifth grade, for example, high supplementation was one of a group of factors associated with higher scores in computationally oriented tests and with lower scores on content emphasized by CSMP. The other factors were:

more teacher experience,
more homework assigned,
less CSMP training, and
less playing math games.

This indicates that increasing supplementation and homework tended to produce more traditional student achievement, i.e., higher in computation but lower in other content.

Given the different patterns of achievement in computational skills, the results of any comparisons between CSMP and Non-CSMP classes are likely to depend somewhat on the composition of the test used; CSMP classes are at a disadvantage on tests which emphasize algorithmic skills and de-emphasize other kinds of computational skills.

The data with regard to differential computational skills at different ability levels were inconsistent. Through fifth grade the results vis-a-vis CSMP versus Non-CSMP were similar regardless of student ability level. If anything, lower ability CSMP students (those scoring in the lowest quartile on the covariate reading test) did better in this regard than did CSMP students at the higher ability levels. At fourth grade for example, scores at the lowest ability levels were the same, but at the highest ability level they favored Non-CSMP. This result is shown in Figure 39, next page.

At sixth grade, however, the results were reversed; the lowest ability CSMP students scored lower than corresponding Non-CSMP students, but there was no difference at any of the other ability levels. This result, the only instance of this phenomenon, is shown in Figure 40.

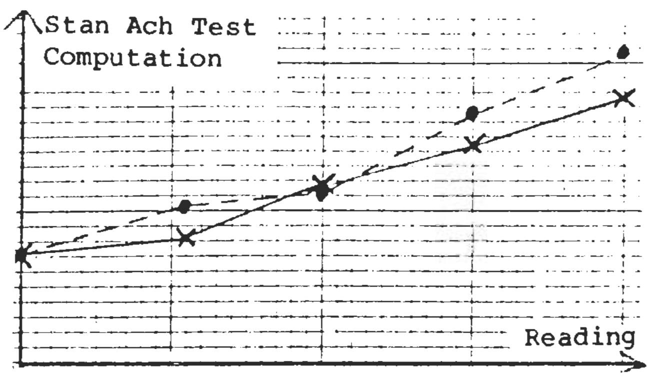


Fig. 39. Fourth Grade Computation Means,
Students grouped by Reading score
(X = CSMP students, ● = Non-CSMP)

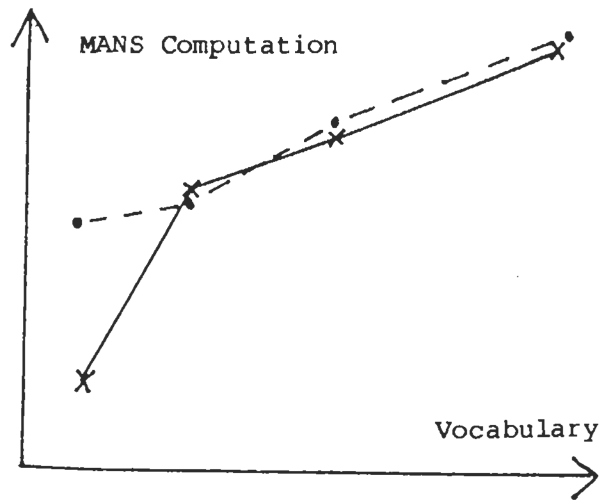


Fig. 40. Sixth Grade Computation Means,
Students grouped by Reading score
(X = CSMP students, ● = Non-CSMP)

Other Standardized Test Results

Extended Pilot Tests. In second and third grades of the Extended Pilot Tests, standardized tests were administered by individual participating districts. The computation portion of these tests were reported in the previous section. Table 28 summarizes the results from the other mathematics tests in these batteries.

Table 28

Summary of Standardized Math Scores (Other than computation) from Extended Pilot Tests

Grade	Test	Number of Classes CSMP	Non-CSMP	Adjusted Means CSMP	Non-CSMP	Sigif at .05	In Favor of
2	CTBS, Conc & App	15	13	18.3	18.2	N	CSMP
	CTBS, Conc & App	6	6	18.1	17.5	N	CSMP
	Stan Ach Test, C & A	6	6	43.5	47.7[2]	N	Non-CSMP
	Coop Prim Test[1]	6	6	36.2	35.7	N	CSMP
3	CTBS, Conc & App	15	13	31.2	33.8	Y	Non-CSMP
	Coop Prim Test[1]	6	12	42.3	41.8	N	CSMP

[1] Total Math Score since this test does not have separately scored tests.
[2] Percentile Ranks

Four of the six comparisons favored CSMP, though the only significant difference was a third grade comparison which favored Non-CSMP classes. It should be noted that this significant result was derived from scores on the CTBS, the regularly administered standardized test for Non-CSMP classes but unfamiliar to the CSMP classes.

Locally Initiated Studies. Several districts initiated their own comparison studies between CSMP and Non-CSMP classes. Some of these are reported in the previous section on computation scores. Those dealing with standardized mathematics tests other than computation are reported in Table 29.

Table 29
Summary of Locally Initiated Standardized
Test Scores, Other than Computation

District	Grade	Test	Comparison	Mean Score CSMP	Mean Score Non-CSMP
19	2	CAT-Conc & App	17 CSMP classes versus same teachers' previous Non-CSMP classes	323	332[1]
4	2	CAT-Total Math	10 CSMP versus 10 Non-CSMP classes	49	51[2]
4	3	CAT-Conc & App	100 CSMP and Non-CSMP students sampled from 6 schools	48	39[3]
17	2	CTBS-Conc & App	6 CSMP classes versus same teachers' previous Non-CSMP classes.	250	240[4]
17	3	CTBS-Conc	Same as above	383	366[4]
		CTBS-App		351	357
17	4	CTBS-Conc	Same as above	416	386[4]
		CTBS-App		386	363
2	3	CAT-Total Math	14 CSMP versus 13 Non-CSMP classes	389	385[5]
8	4	ITBS-Conc	16,17 CSMP and Non-CSMP students	76	72[6]
		ITBS-Prob	sampled from 2 schools	76	62
		SRA-Conc	same as above	84	83[6]
		SRA-Prob		88	96
8	5	ITBS-Conc	20-24 CSMP & Non-CSMP students	78	74[6]
		ITBS-Prob	sampled from 2 schools	80	79
		SRA-Conc	same as above	83	78[6]
		SRA-Prob		80	76
12	6	ITBS-Conc	5 year longitudinal study of	7.8	7.7[7]
		ITBS-Problems	70 CSMP versus 90 Non-CSMP students	7.5	7.4
12	7	ITBS-Conc	same as above	9.2	8.9[7]
		ITBS-Prob		8.7	8.5
16	6	CAT-Total Math	2 CSMP versus 3 Non-CSMP classes	57.4	57.4[8]

[1] Mean scaled score across classes;
an approximately equal decline occurred in Reading.
[2] Adjusted mean raw score across classes.
[3] Gain, from previous year, in mean student scale scores.

[4] Adjusted mean scaled score across classes.
[5] Mean scaled score across classes.
[6] Average mean percent correct across items.

[7] Mean grade equivalent scores across students.
[8] Mean raw score across students; mean IQ scores = 110.0, 110.5
mean Reading scores = 56.3, 53.8

Out of twelve comparisons (defining a comparison as one grade level at one district), eight favored CSMP and three favored Non-CSMP, though the difference in most comparisons was quite small and in only one comparison (District 17, grade 4, in favor of CSMP) was there a large difference. Furthermore, there were no grade level distinctions, nor were there much different findings with regard to Concepts versus Problems (or Applications).

Summary

There have been no less than 55 studies, involving 32 school districts, comparing the performance of CSMP and Non-CSMP students on standardized tests. The results could hardly be more even: 32 studies favored CSMP and 23 favored Non-CSMP. Large differences were found in only eight studies, four in favor of CSMP and four in favor of Non-CSMP. For the most part, findings were similar in each of the usual subdivisions: computation, concepts and applications, and at each level of student or class ability.

In spite of these findings, most CSMP teachers consider the program to be deficient in providing sufficient practice in computational skills, particularly rapid recall of basic facts (lower grades) and proficiency with multi-digit algorithms (upper grades). Inadequate rapid recall, if it does exist with CSMP students, does not affect their performance on standardized tests through third grade.

However, proficiency in multi-digit algorithms is lower for CSMP students and does effect standardized test performance in grades 4-6. CSMP students do not do as well on items requiring multi-digit algorithms but this is sometimes balanced by better performance on other kinds of computation items such as those using fractions and decimals.

Most teachers do supplement CSMP with computation practice and in this way may remediate the perceived deficiency. There is some evidence that increased supplementation improves computation scores. For many teachers this supplementation is done a few minutes each day, or sent home as homework, and is therefore fairly unobtrusive.

Just as the differences in computation, an area in which CSMP students might be expected to do poorly, were small and easily remediable, so too the differences on the standardized problem solving tests, where CSMP students might be expected to do better, were also small. But "problem solving" on standardized tests usually means solving one-step word problems with significant computation and reading requirements, so the results are unsurprising. On the MANS word problem tests, however, where computation and reading requirements were kept low, CSMP students had a small but consistent and significant advantage.

On the basis of standardized testing alone, CSMP doesn't seem to make much difference one way or the other. If that is the single criterion for making a curriculum decision, then CSMP must be rejected because of its cost and teacher training requirement. Of course in the case of CSMP there is a great deal of other evidence which demonstrates rather persuavely that teachers like the program and that students do better in many areas of problem solving. One wonders, however, how many other innovative curricula, national and local, did not have the resources to perform the kind of research and evaluation that was done on CSMP, and were scuttled because they didn't get the necessary gains or standardized tests.

VII. OTHER FINDINGS

Entering CSMP

Rapid Implementation Model

In two districts where CSMP was implemented in a single school at grades K-5 at the same time, the MANS tests were administered to all second, third and fourth graders on three occasions: the year before start up, at the end of the first implementation year, and at the end of the second implementation year. (Fifth grade tests were not available the year before start up.) One school was in District 17, a large urban district, and had six classes per grade level. The other school was in a relatively affluent neighborhood of District 23, a suburban district, with three classes per grade.

The results were similar in both districts. In second grade, there was a large gain in adjusted MANS scores after one year of CSMP and an additional small gain after the second year. This finding is illustrated in Figure 41, which shows these districts' scores in Year 0 (circled dots), Year 1 (squares) and Year 2 (circled x's). These data have been superimposed on the graph of district means from the Extended Pilot Test (which are represented by regular x's and dots).

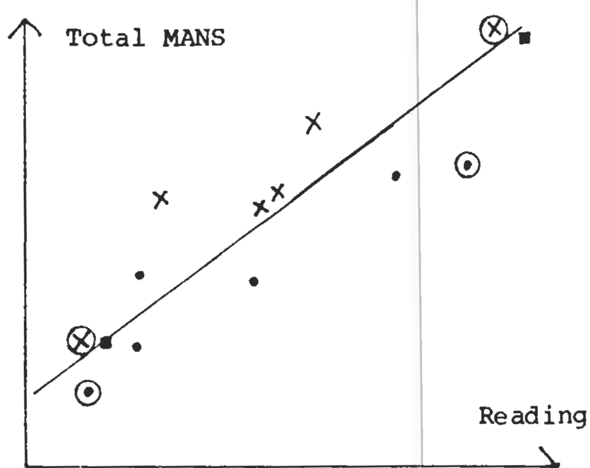


Fig. 41. District 17 and District 23 Second Grade Means
⊙ = year preceding CSMP, ■ = 1 year of CSMP, ⊗ = 2 years CSMP
(X, ● = CSMP, Non-CSMP district means from Extended Pilot Tests)

In third and fourth grades, sizeable gains were made after each of the first and second years though again most of the eventual gain occurred after one year's experience with CSMP. Table 30 summarizes the adjusted means.

Table 30
Adjusted Means, Grades 2-4
Rapid Implementation Model

		Year 0	Year 1	Year 2
Second Grade:	District 17	38.5	42.5	44.2
	District 23	57.1	65.9	68.4
Third Grade:	District 17	45.5	49.4	53.3
	District 23	86.1	95.5	NA
Fourth Grade	District 17[1]	48.8	60.9	64.9
	District 23[1]	123.4	141.7	150.4

[1] The test used in fourth grade differed from district to district.

The findings indicate that it is possible for schools to begin using CSMP right away in grades 2-4, rather than having to start at K,1 and advance one grade level each year, which is the normal implementation strategy. When successful implementations of this kind have taken place, they have been overseen by a strong coordinator with authority and commitment. In Districts 17 and 23, it was probably a strong teaching staff and able students, District 17, and training personnel with year-long, full time position in the school, District 23, that allowed the model to run successfully.

Entering Students

In the usual method of analysis of Extended Pilot Test data, students who entered class during the course of the year - both CSMP and Non-CSMP - were excluded from the analysis. In fifth grade, a separate analysis was made for these "late" students. A separate analysis was also done for students who transferred into class at the beginning of the school year ("new" students). These students had no CSMP experience in K-4, but then joined an experienced CSMP class in the fall of fifth grade.

There were 55 and 24 "new" CSMP and Non-CSMP students respectively and 31 and 25 "late" students (an average of exactly one late student per class). Mean scores on the covariate Vocabulary test and on the MANS tests were calculated for each group. MANS performance is plotted against Reading for each of these groups on the graph (next page). This data has been superimposed on the original graph showing performance of all other CSMP and Non-CSMP students when grouped into quartiles by reading score (shown earlier). Circled entries are for new students and boxed entries are for late students.

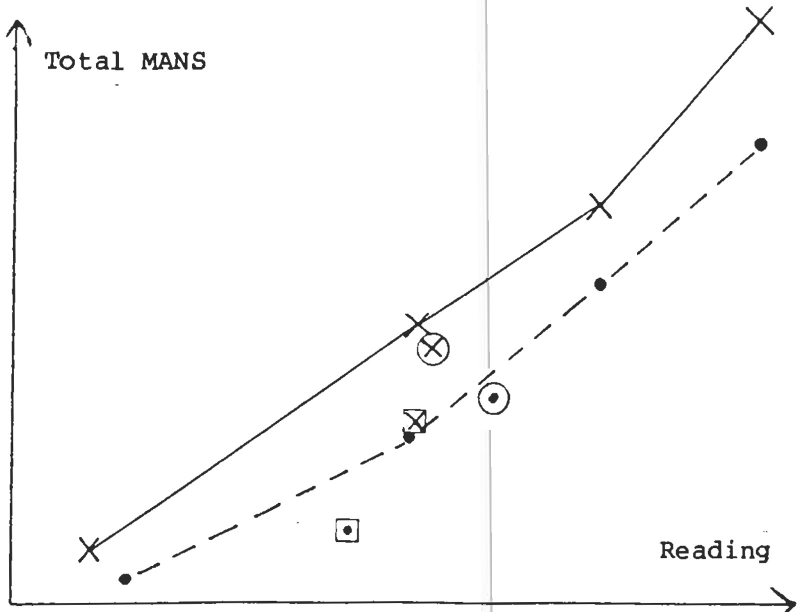


Fig. 42. "New" and "Late" fifth grade students' MANS score superimposed on graph of regular students' scores (Circled entries = New Students: X = CSMP, • = Non-CSMP) (Boxed entries = Late Students: $\boxed{X}$ = CSMP, $\boxed{\bullet}$ = Non-CSMP)

New students, both CSMP and Non-CSMP, scored slightly lower on the MANS scores than students generally, i.e., the circled entries were slightly below the corresponding line segment. Late students, however, scored quite a bit lower than students generally, i.e., the boxed entries are well below the line segments, and this finding also applies to both CSMP and Non-CSMP students. Interestingly, the CSMP advantage in the general population is preserved in each of these special groups.

In comparison to other CSMP students, new CSMP students had lower MANS scores, by about 10%, on items dealing with fractions and probability. Late students had lower scores in almost all areas, but especially in items dealing with mental arithmetic and the production of multiple answers where their scores were about 20% lower.

A situation half way between whole classes starting the program and individual students joining an intact CSMP class occurred in District 16. In one school, two second grade classes studied CSMP and the other two classes did not. The following year all these students studied CSMP as third graders. Classes were thoroughly mixed so that about half of each class had studied CSMP in first and second grade while the other half had no previous CSMP experience.

At the end of third grade the MANS Tests were administered to all students. Scores were adjusted for differences on the previous year's California Achievement Test. These adjusted means are shown in Table 31, together with adjusted means from classes of roughly comparable ability who had been tested previously during the Extended Pilot Test.

Table 31
Comparison of New and Experienced
Third Grade CSMP Students in District 16

	Adjusted Total MANS	
	CSMP	Non-CSMP
District 16	96.0	92.9 (CSMP only in third grade)
Other comparable classes during Extended Pilot Tests	93.1	85.6

CSMP students in the Extended Pilot Test outscored Non-CSMP students by about seven and one-half items. In District 16, experienced CSMP students outscored inexperienced CSMP students by about three items. One can infer from this data that, under these circumstances, third grade CSMP alone produces about half as large an incremental MANS effect as the grades 1-3 portion of the program does. This finding is in agreement with findings for Districts 17 and 23 who used the Rapid Implementation Model.

Leaving CSMP

CSMP Graduates

Not enough time has elapsed to gather much data on CSMP "graduates." However, in three districts (Districts 2, 12 and 18), seventh grade mathematics teachers were asked mid-way through the school year to rate each of their students using a 4-item, 5-point rating scale. Classes were always mixed, containing some ex-CSMP students and some ex-Non-CSMP students. In each district there were between 36 and 48 ex-CSMP students who had attended one elementary school, and between 74 and 210 ex-Non-CSMP students who had attended 2-4 elementary schools.

Teachers were asked to rate students on each of four characteristics:

Participation in class: high quality and frequent participation, listens well, attends well, volunteers responses.

Motivation: strong interest, works independently, interested in "why", likes new ideas.

Creativity and problem solving: reasoning and logic skills, tries new methods or several methods to solve a problem.

Practical applications: knows conventional terms and symbols, can organize and interpret, translates new problems into familiar forms.

Average ratings at each district were calculated for each item, and a total score was calculated from the sum of the four items. Total scores in the three districts, for CSMP and Non-CSMP respectively, were:

District 2:	adjusted means:	12.8 versus 12.8	
District 12:	adjusted means:	12.8 versus 11.3	($p < .20$)
District 18:	unadjusted means:	12.9 versus 10.8	($p < .05$)

There were virtually no differences among the four rating items. In District 2 CSMP and Non-CSMP scores were virtually identical on each item; in District 12 each item produced a slight difference in favor of CSMP; and in District 18 there was about a half-point difference in favor of CSMP on each item.

In District 12, math grades were compared for seventh grade ex-CSMP and ex-Non-CSMP students, using Analysis of Covariance, with Verbal section of the Cognitive Abilities Test as a covariate. The adjusted mean grades using (A = 5, B = 4, etc.) were always in favor of CSMP and are shown below for the first, second and third quarters respectively.

3.9 versus 3.6
3.8 versus 3.5
3.7 versus 3.5

These differences were significant at .05, .02 and .10 respectively on the Analysis of Covariance.

Students in District 12 were then divided into equal sized groups according to Cognitive Ability Test scores. The mean CAT scores for the four groups was about 90, 105, 115 and 125, illustrating the fact that this district was populated by students of fairly high ability. For each group, mean math grade and mean teacher rating were also calculated. Figure 43 shows the resulting means.

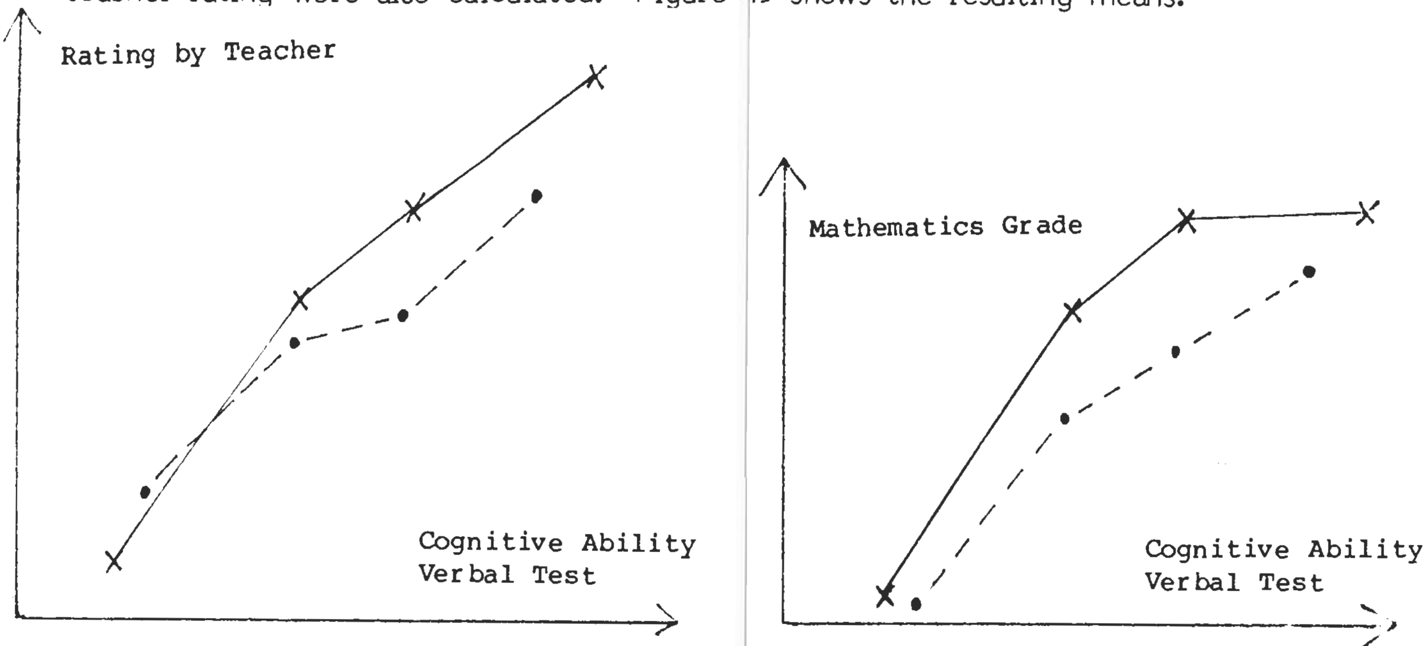


Fig. 43. District 12 Seventh Grade Student Means
when grouped by Cognitive Ability Verbal Score
(X = Ex-CSMP Students, ● = Ex-Non-CSMP Students)

The graphs show the relatively clear advantage for CSMP students in math grades and the small advantage in teacher ratings. The graphs also show that these advantages are to be found mostly at the upper ability levels; at the lowest ability level there are virtually no differences in teacher ratings and a small CSMP advantage in math grades.

It should be noted that in all three districts, seventh graders studied the regular district seventh grade mathematics curriculum; no special arrangements were made to take into account the special strengths of CSMP students. Thus, the results represent in a sense, the "worst case scenario". As districts start to use CSMP district wide, it will be to their advantage to alter their seventh grade curriculum accordingly, in which case the long range benefits of CSMP should be more strongly apparent.

Leaving CSMP After Third Grade

In District 21, a study of classes who stopped CSMP at third grade was carried out. The district decided not to begin implementation of the CSMP curriculum in grades 4-6, so in fourth grade these classes returned to a more traditional mathematics program from one of the standard textbooks.

At the end of fourth grade, the MANS Tests were administered to these ex-CSMP students, who constituted seven classes in two schools, and to seven classes of similar ability from two adjacent schools who had never studied CSMP.

Mean scores on the MANS Tests were calculated across ex-CSMP classes and across ex-Non-CSMP classes. These mean scores are shown in Figure 44, superimposed on the graph of district means generated from the earlier Extended Pilot Test of fourth grade. The circled entries represent the scores for District 21.

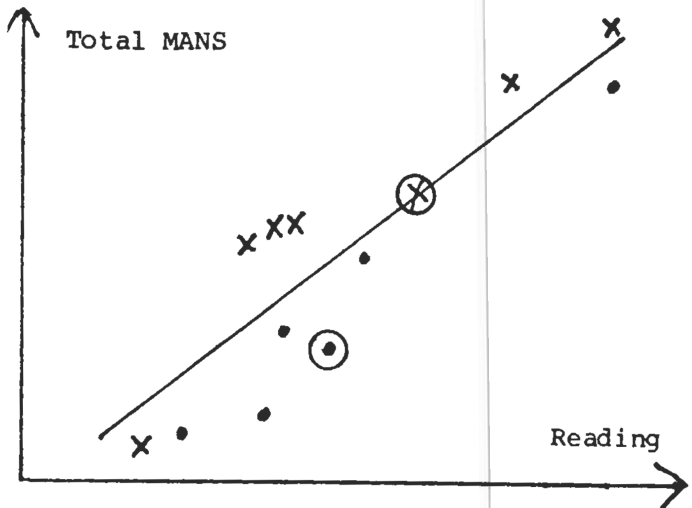


Fig. 44. Fourth Grade Class Means, District 21, Superimposed on EPT district means
(⊗ = District 21 - CSMP [K-3] + traditional [4th grade])
(⊙ = District 21 - traditional (K-4))

The graph shows that the ex-CSMP classes had higher MANS scores than the ex-Non-CSMP classes. This difference was significant at .05 on the ANCOVA of class means, though the differences in covariate scores between the two groups was larger than desirable for that kind of analysis.

When graphs of class means (seven ex-CSMP, seven ex-Non-CSMP) were examined, three ex-CSMP classes had very high MANS scores relative to ability; the other four ex-CSMP classes had MANS scores similar to ex-Non-CSMP classes. The three high scoring classes were not from the same school and the degree to which their teachers "followed-up" on students' CSMP background is not known.

The largest differences in favor of ex-CSMP students were found in two categories: Number Representations, and Relationships and Number Patterns. For this latter category, even though the difference was large and significant at .01, it was much smaller than the differences found during the Extended Pilot Test.

This study is another indication of some residual effects from CSMP after a year away from the curriculum. It also indicates that the MANS effects may not be long lasting if used only in the primary grades without specific follow-up.

Differences According to Sex of Student

Nine studies were conducted comparing boys' and girls' performance on the MANS Tests. For each study an effect size was calculated by dividing the boy-girl difference in mean scores by the pooled standard deviation. This was done separately for CSMP and Non-CSMP students.

There were two studies at each grade level except third grade (three studies) and sixth grade (one study). Usually, one study at each grade level was based on all data from the Extended Pilot Test at that grade and the other was based on one or two years of use of the revised MANS in Joint Research Studies. There were an average of about 1100 students per study; i.e., an average of about 275 CSMP boys, CSMP girls, Non-CSMP boys and Non-CSMP girls.

Table 32 summarizes the data by MANS category and grade level. The effect sizes given in the table are averages across studies, with different studies weighted according to the number of participating students.

Table 32
Average Effect Size, Boys versus Girls
(First entry = CSMP, second entry = Non-CSMP)

Category	2	3	4	5	6	Average
Computation	-14	-07	-07	-19	-09	-11
	-18	-18	-01	-12	-05	-11
Estimation	12	23	25	08	35	21
	18	15	36	27	35	26
Mental Arithmetic	18	21	27	10	41	23
	21	17	42	23	32	27
Number Represents	14	26	15	14	23	18
	20	10	17	11	03	12
Relationships & Patts	07	17	14	05	19	12
	14	07	13	14	26	15
Word Problems	05	23	20	15	29	18
	05	09	26	26	24	18
Elucidation of	-07	-11	-15	-16	-05	-11
Multiple Answers	-06	-19	-06	-06	-02	-08
Total MANS	06	16	17	09	15	13

Girls had higher scores than boys at every grade level, in both CSMP and Non-CSMP groups, in two categories: Computation and Elucidation of Multiple Responses. The difference averaged about 1/10 of a standard deviation in both categories (less at sixth grade) with CSMP girls having a larger advantage than Non-CSMP girls in grades 4-6.

Boys had higher scores than girls at every grade level, in both CSMP and Non-CSMP groups, in the other five categories. The difference averaged between 0.1 and 0.2 standard deviations in three categories: Number Representations, Relationships and Number Patterns, and Word Problems. The difference averaged about 1/4 of a standard deviation (slightly less for CSMP students) in two categories: Estimation and Mental Arithmetic. These differences favoring boys tended to be largest in fourth and sixth grades.

If one assumes a normal distribution of scores for both boys and girls, effect size can be illustrated in Figure 45 below.

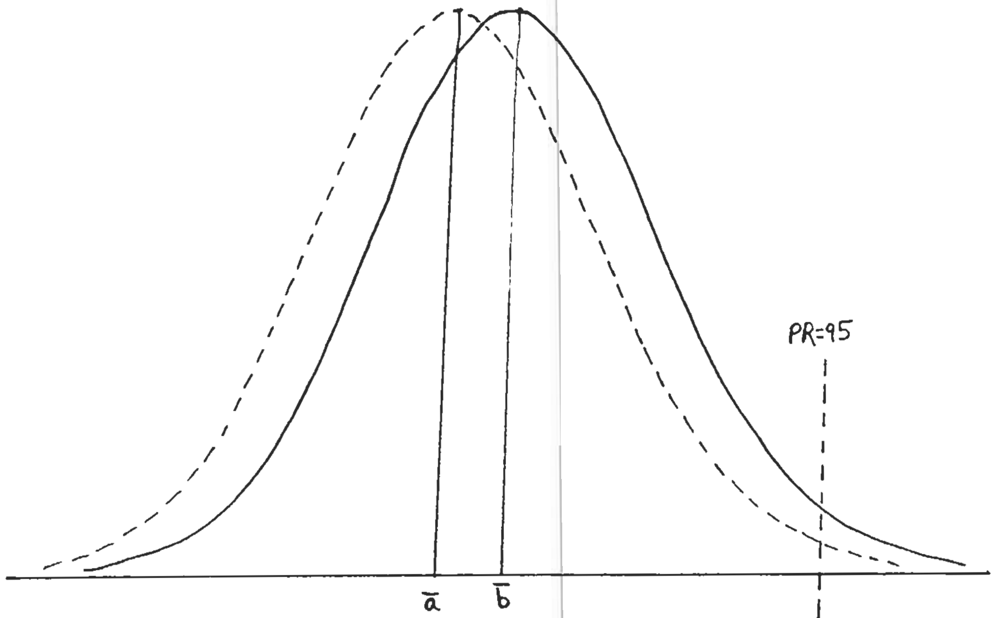


Fig. 45. Hypothetical normal distributions, MANS scores, (b, g = mean MANS score for boys and girls respectively)

The effect size determines the separation of b and g, the means for boys and girls respectively. In Mental Arithmetic and Estimation, the effect size was .25 (1/4 standard deviation), meaning that b corresponds to the 55th percentile rank on the combined distribution while g corresponds to the 45th percentile rank.

Furthermore, an effect size of .25 may result in a disproportionate number of boys in the tail of the distribution, i.e., above the 95th percentile (the portion to the right of the dotted line in the above figure). Under the assumption of normal distributions for boys and girls, boys would outnumber girls by nearly 2 to 1 in the top 5% of the combined distribution. This hypothesis was checked for the Estimation category in fourth grade. For CSMP students the effect size was .25 and about 3/5 of the students in the top 5% were boys. For Non-CSMP students, the effect size was .36 and about 4/5 of the students in the top 5% were boys. Thus, to the extent that the skills or abilities tested in Mental Arithmetic and Estimation are important components of mathematical thinking, girls as a group may be somewhat disadvantaged and under-represented in the top group of mathematical thinkers. Furthermore, this deficit is measurable as early as second grade.

Student Attitudes

In the fourth and fifth grade Extended Pilot Tests, students were asked to complete a series of attitudinal items borrowed from the National Assessment of Educational Progress (NAEP), as part of the MANS testing.

In fourth grade, there were three groups of items, none of which produced significant differences between CSMP and Non-CSMP students:

Attitude Toward School Subjects (6 items)

e.g. Science: Like In Between Do not like

Self Concept and Math (5 items)

e.g. I usually understand what we are talking about in mathematics

True about me Sometimes true about me Not true about me

Attitude Towards Math Activities (6 items)

e.g. Playing mathematical games:

I like it a lot I like it a little I don't like it

No single item produced a difference larger than 6 percentage points between the two groups, and responses were very close to those obtained from NAEP's national sample.

In fifth grade, there were seven groups of items, two of which produced significant differences between CSMP and Non-CSMP classes.

1. Math versus other subjects. This scale was scored by calculating the difference between the math "score" and the other subjects "score" (using items like the example given above under Attitude Toward School Subjects). Non-CSMP classes had significantly higher scores on this scale. Percentage of responses are shown in Table 33. For comparison purposes, the results from fourth grade are also shown.

Table 33
Percentage Responses for
Attitude Toward School Subjects
(First entry = CSMP, second entry = Non-CSMP)

		Like		In Between		Do Not Like	
Fourth Grade:	Math	58	57	29	30	12	12
	Other Subjects[1]	50	53	48	45	11	12
Fifth Grade:	Math	51	58	33	28	16	14
	Other Subjects[1]	51	44	35	40	13	16

[1] Average of responses for science, social studies, reading and spelling

Table 33 shows that there was very little difference in fourth grade between CSMP and Non-CSMP responses. From fourth to fifth grade, however, there were two changes. First, fifth grade Non-CSMP students liked math as well as fourth graders had liked it but liked other subjects less; this is a difficult result to explain. Second, CSMP fifth graders liked math less but stayed the same in other subjects. This finding is in contrast to teacher opinions about student involvement and enthusiasm in which CSMP, compared to previous math curriculum, was rated at over 4.0 on a 5 point scale in each of grades 4-6 (and at least 1/2 point higher than Non-CSMP teachers rated their curriculum).

2. Math is open. CSMP classes had significantly higher scores on this 3-item scale, an example of which is given below.

Being good at pretending helps people in math:

Always true   Usually true   Not usually true   Never true

Five other scales, containing from 2 to 5 items, produced no significant differences between the two groups: Self concept in mathematics, value of the spiral appraoch, value of estimation, math is closed, and math is mainly calculation.

Tests of Specific CSMP Content

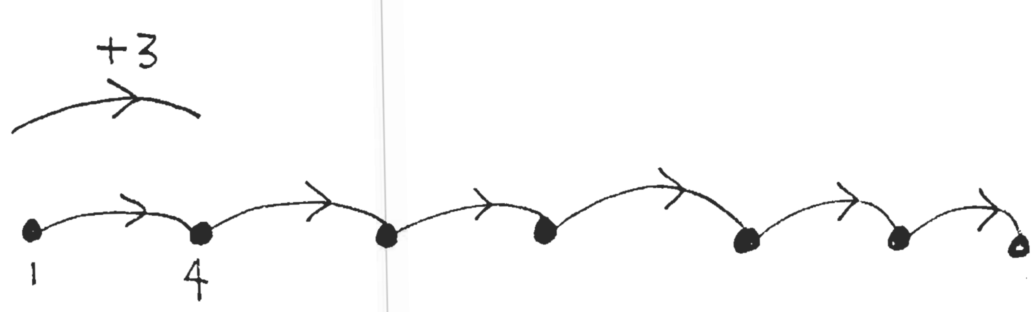
In each of grades 1-3, tests of specific CSMP content were administered to CSMP students. The tests were constructed to model the kinds of problems that were assigned to CSMP students in workbooks and worksheets. Primarily these tests were intended to assess how well students understood and could use the CSMP representational languages (minicomputer, arrow diagrams, and string pictures). This testing was discontinued after third grade because of the difficulty in interpreting the data, since there are no behavioral objectives or standards in CSMP's spiral curriculum, and since facility in these languages is not an end in itself but a vehicle for mathematical thinking.

The tests were administered as workbook problems to groups of 10-12 students. In format and in administration the task was always very similar to what students were used to doing in math class, i.e., a very non-test like situation. The total number of students tested was about 300, 600 and 100 students in grades 1, 2, and 3 respectively.

Sample items from each grade are given below for each of the CSMP languages, together with percentage of students getting the problem correct.

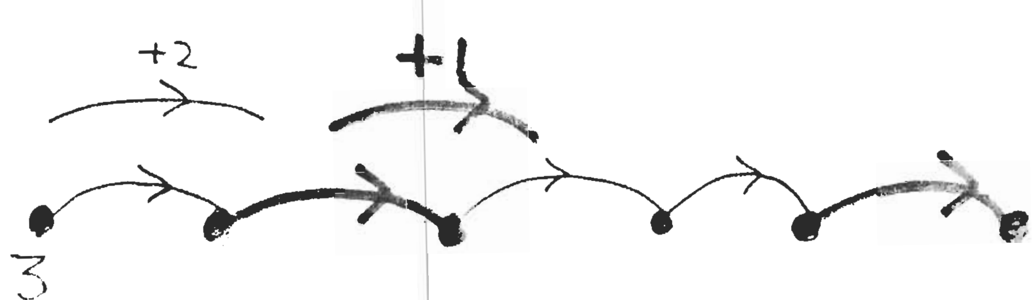
Arrow Diagrams

First Grade. (a)



Average percent of dots labelled correctly = 67.

(b)

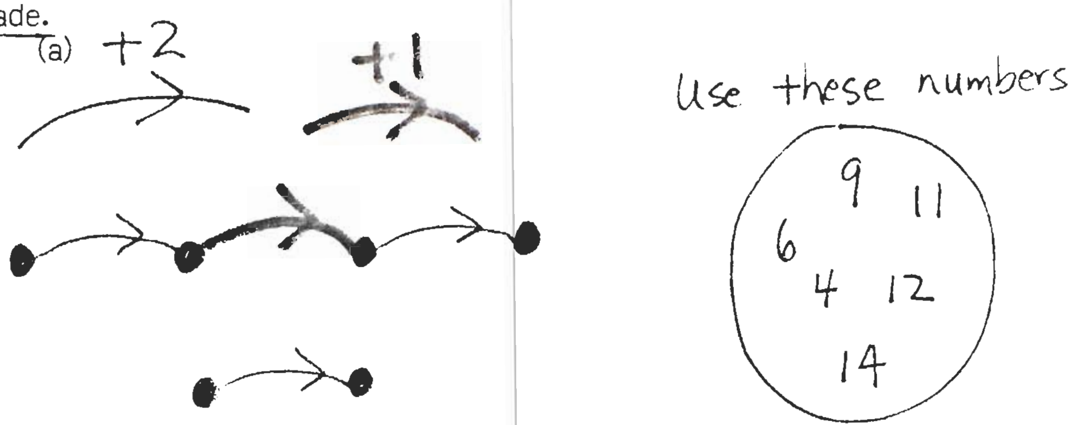


Average percent of dots labelled correctly = 55.

About 25% of the students did not know how to do these questions.

Second Grade.

(a)

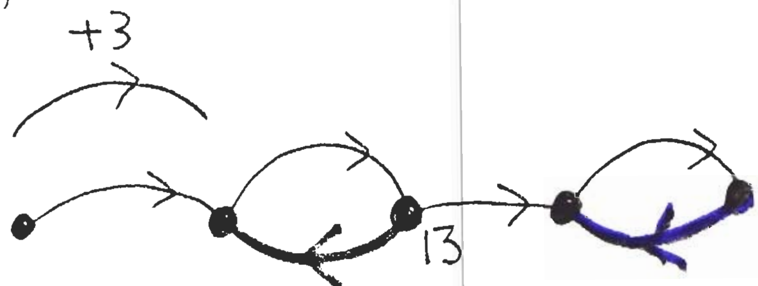


Average percent of dots labelled correctly = 74
Average percent of students with complete solution = 56

(b) Build a road from 1 to 8 with +3 and +2 arrows.

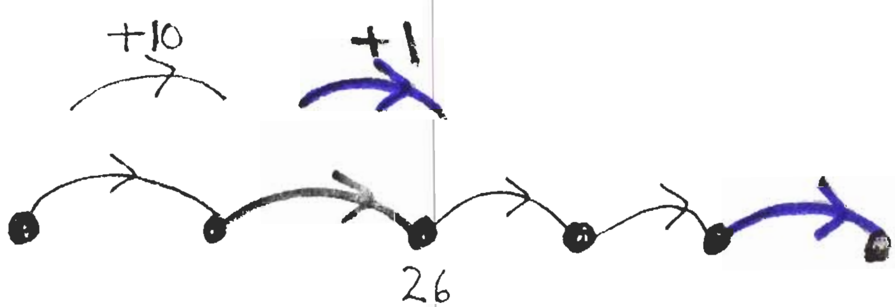
Average percent with a correct road = 58.

(c)



Average percent of dots labelled correctly = 77
Average percent of students with complete solution = 52

(d)



Average percent of dots labelled correctly = 74
Percent of students able to label return arrow correctly = 47

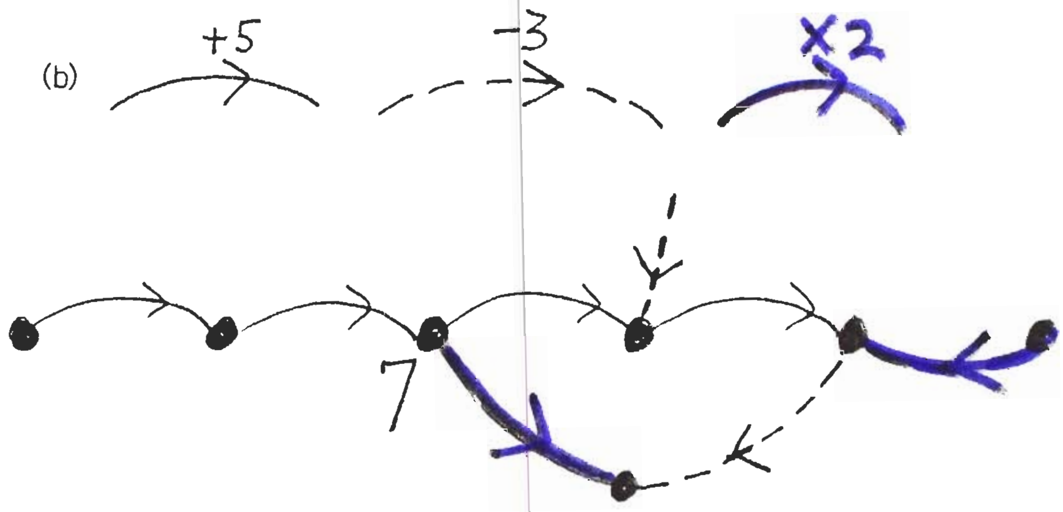
In addition to being able to do harder items in second grade, the percentages correct increased, especially among low ability students. For the lowest ability group (percentile rank <20 on Reading or IQ test), the average percent correct was 49.

Third Grade.

(a) Item (d) above

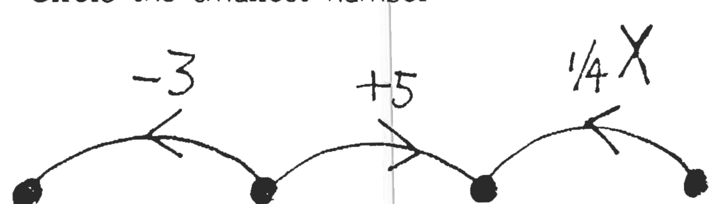
Average percent of dots labelled correctly = 86
Percent of students labelling return arrow correctly = 71

(b)



Average percent of dots labelled correctly = 61

(c) Circle the smallest number



Percent correct = 41

From second to third grade there was again substantial improvement, as in (a), and some genuinely hard problems are asked, as in (b) and (c).

Minicomputer

First Grade.

(a) Show 2 and 3-digit numbers on the minicomputer: approx 70% correct

(b) Read 2 and 3-digit numbers from the minicomputer: 65% correct

(c) Use minicomputer to add 2-digit numbers: 45%

(d) Use minicomputer to multiply 2X or 3X 2-digit numbers: 30%

(e) Use minicomputer to subtract 2-digit numbers: 25%

About 25% of the students could not do any of the minicomputer questions and about 50% of the students could not use the minicomputer for any kind of computing.

Second Grade.
(a) Show 2 and 3-digit numbers on the minicomputer: 76%

(b) Read 2 and 3-digitnumbers from the minicomputer: 65%

(c) Show numbers on the Minicomputer (e.g. 6, 14, 24, 60) with exactly three checkers: 63%

(d) Read numbers shown in non-standard form (e.g. more than one checker per square), ones board only : 72%

The lowest ability group (percentile rank <20 on Reading or IQ test) averaged about 46% correct on these items.

Third Grade.
(a) Show 2 and 3-digit numbers on the minicomputer: 81%

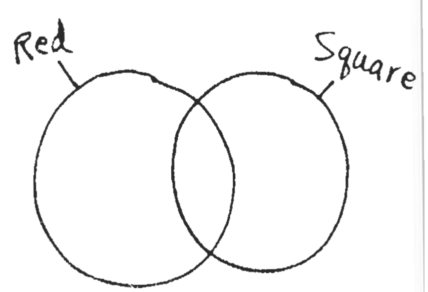
(b) Read decimal numbers from the minicomputer: 40%

(c) Show decimal numbers on the minicomputer: 15%

(d) Adding negative checkers to a display to show a certain number on the minicomputer: 50%

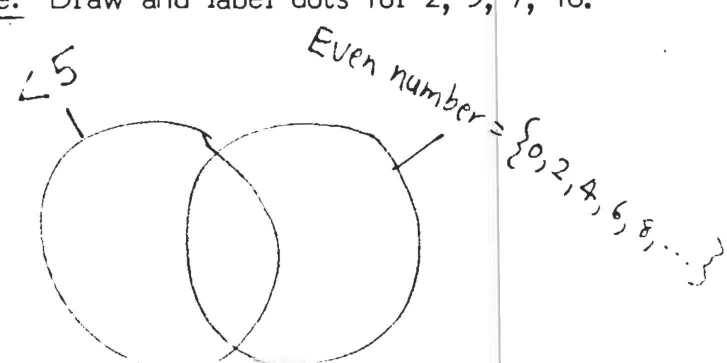
String Pictures.

First Grade. Draw a dot in the picture to show where the red square goes



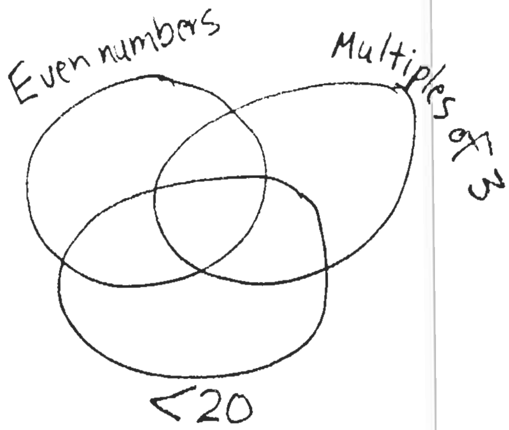
Percent correct = 39

Second Grade. Draw and label dots for 2, 3, 7, 10.



Average percent correct = 64.

Third Grade. Draw and label dots for ...(16 numbers given)



Average percent correct = 42.

At the end of first grade, substantial numbers of students (at least 25%) were unable to answer very straightforward questions about the minicomputer, arrow diagrams and string pictures. By the end of second grade, with harder questions, even the lowest ability students were able to get about half the items correct. By third grade, students were engaged in complicated problems which required real facility with the languages; about half of the third graders were successful on these more difficult problems.

Analysis of Results by Number Type

In order to analyze the performance of CSMP versus Non-CSMP students with respect to type of number, items dealing with fractions and items dealing with decimals were analyzed separately. Table 34 shows percent correct (adjusted for reading or vocabulary) for fractions, decimals and other MANS items.

Table 34
Percent Correct, Fraction and Decimal Items

	Number of Items	Percent Correct CSMP	Non-CSMP
Fourth Grade			
Fractions	15	57	47
All other MANS items	234	58	51
Fifth Grade			
Fractions	64	63	59
Decimals	29	71	57
All other MANS items	221	62	57
Six Grade			
Fractions	57	73	70
Decimals	31	71	61
All other MANS items	336	67	61

On fraction items, CSMP students had a large advantage in fourth grade (larger than MANS items overall) and a small advantage in fifth and sixth grades (smaller than MANS items overall).

On decimal items, CSMP students had much higher scores than Non-CSMP students in both fifth and sixth grades (larger than MANS items overall).

Figure 46 shows results for fraction items analyzed by ability level of student, as measured by reading or vocabulary (and dividing students into groups according to published norms).

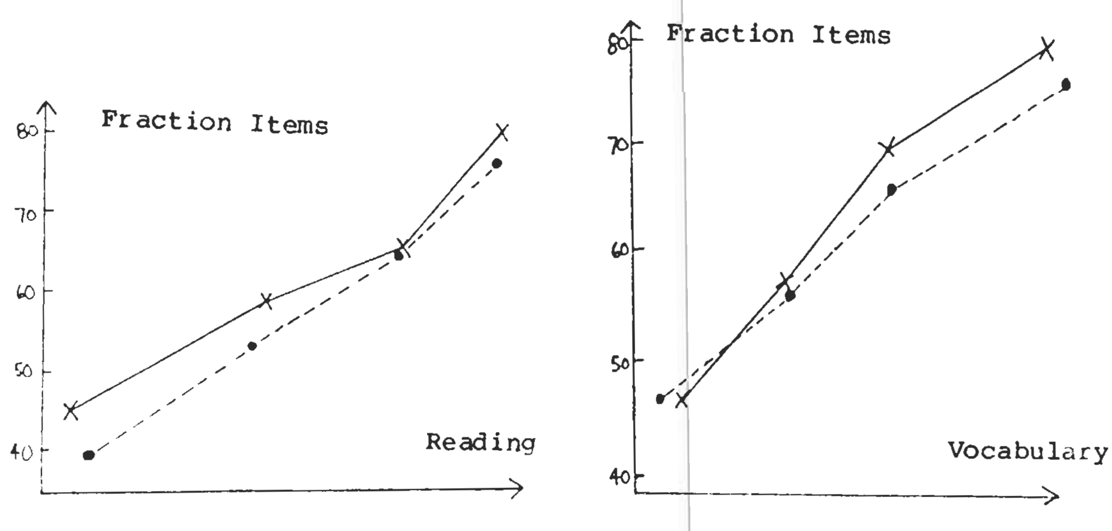


Figure 46. Percent Correct, Fraction Items, Extended Pilot Test
Fifth grade (left hand graph) and Sixth grade (right)
(X = CSMP Students, ● = Non-CSMP Students)

In fifth grade, the CSMP advantage was due mostly to the superior performance (compared to Non-CSMP) of low ability students, while in sixth grade it was due to the superior performance of high ability students. In both years, the gap between CSMP and Non-CSMP performance was smaller than it was for all MANS items combined (compare with Figures 23 and 24).

Figure 47 shows similar results for decimal items.

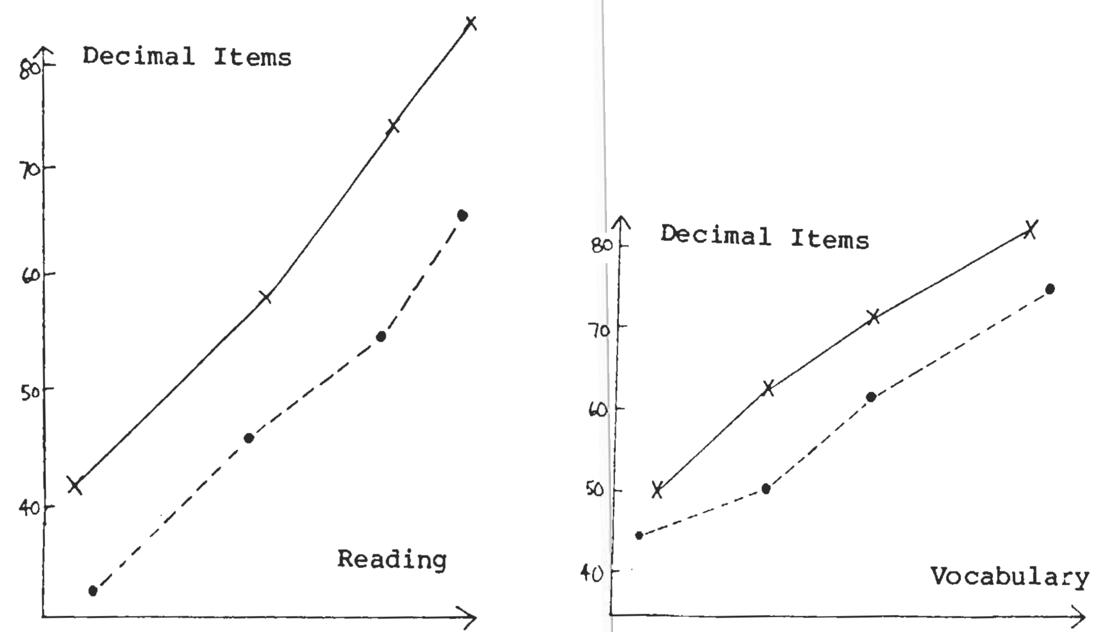


Figure 47. Percent Correct, Decimal Items, Extended Pilot Test
Fifth grade (left hand graph) and Sixth grade (right)
(X = CSMP Students, ● = Non-CSMP Students)

These results are more consistent and the CSMP - Non-CSMP gaps are larger than they are for all MANS items combined.

VIII. SUMMARY

On February 6 and 7, the CSMP Evaluation Review Panel met in St. Louis. This was the only meeting this group held; its charge was to "review the implications of the CSMP evaluation data for mathematics education and to make recommendations based on these implications." The members of the panel are listed below; their report begins on the next page, and continues through page 162. After that there is a brief discussion of the results.

CSMP Evaluation Review Panel

Theresa Denman,
Mathematics Supervisor, Grades K-5,
Detroit Public Schools

Robert Dilworth,
Professor of Mathematics,
California Institute of Technology

Edward Esty,
Senior Associate, Office of Educational Research and Improvement,
Department of Education

Shirley Hill,
Professor of Education,
University of Missouri at Kansas City

Ernest House,
Professor, Center for Instructional Research and Evaluation,
University of Illinois

Stanley Smith,
Coordinator, Office of Mathematics K-12,
Baltimore County Public Schools

Jane Swafford,
Dean of Graduate Studies,
Northern Michigan University

Marie Vitale,
Acting Director of Secondary Education,
Ann Arbor Public Schools

In fifth grade, the CSMP advantage was due mostly to the superior performance (compared to Non-CSMP) of low ability students, while in sixth grade it was due to the superior performance of high ability students. In both years, the gap between CSMP and Non-CSMP performance was smaller than it was for all MANS items combined (compare with Figures 23 and 24).

Figure 47 shows similar results for decimal items.

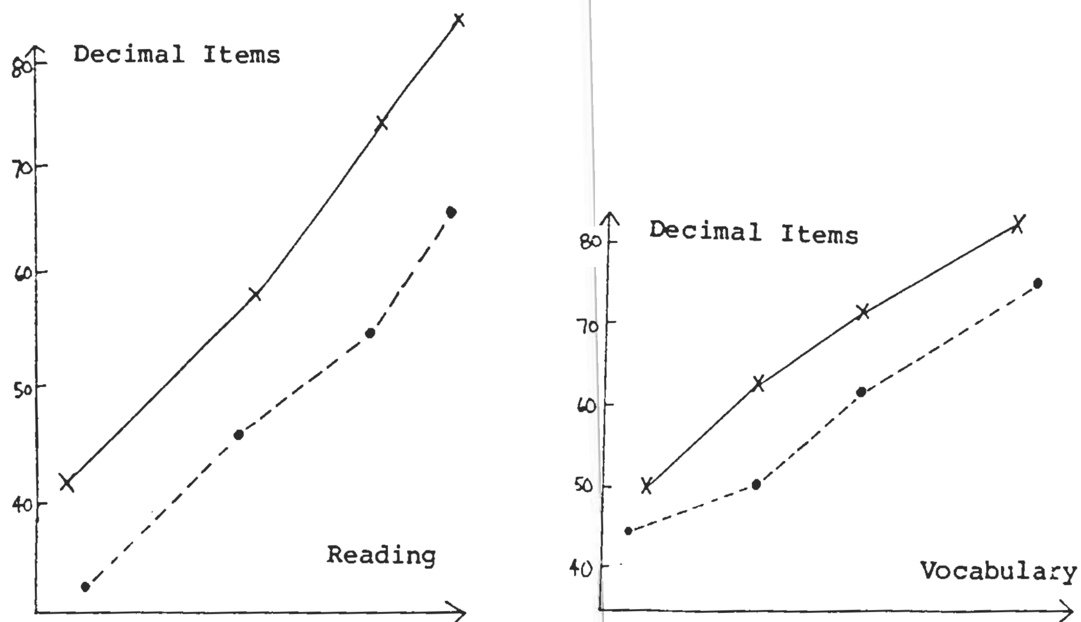


Figure 47. Percent Correct, Decimal Items, Extended Pilot Test
Fifth grade (left hand graph) and Sixth grade (right)
(X = CSMP Students, ● = Non-CSMP Students)

These results are more consistent and the CSMP - Non-CSMP gaps are larger than they are for all MANS items combined.

Conclusions and Recommendations Of The Evaluation Review Panel

Overview

The Comprehensive School Mathematics Program (CSMP) is a dramatic curricular innovation in elementary school mathematics. During its development, conscious decisions were made about how mathematics should be taught. The most important of these were the following:

> Mathematically important ideas should be introduced to children early and often, in ways that are appropriate to their interests and level of sophistication. The concepts (but not the terminology) of set, relation and function should have pre-eminent place in the curriculum. Certain content areas, such as probability, combinatorics, and geometry should be introduced into the curriculum in a practical, integrated manner.
>
> The development of rich problem solving activities should have a prominent place in the curriculum. These activities should generate topics, guide the sequencing of content, and provide the vehicle for the development of computation skills.
>
> The curriculum should be organized into a spiral form which would combine brief exposures to a topic (separated by several days before the topic appears again) with a thorough integration of topics from day to day.
>
> Whole group lessons should occupy a larger and more important role in mathematics class and teachers should be provided with highly detailed lesson plans which lay out both the content and pedagogical development of lessons. Furthermore, training in both the content and pedagogy of the program should be made available to teachers.

These beliefs about the teaching of mathematics were translated with remarkable integrity into the eventual curriculum materials. CSMP is a model of one very distinctive way of teaching mathematics and is one of the few that can be studied in detail by mathematics education researchers and teachers. Its implementation and evaluation in schools is, in a sense, an experimental test of these distinctive features.

Immediate gains in student learning of the kind emphasized in CSMP, particularly problem solving, should not be expected and are unrealistic. Some of CSMP's most important effects will be subtle and diffuse, for example, residual effects on teachers beyond the formal implementation of CSMP, the appearance in textbooks of the CSMP pedagogical techniques, problems and languages, and the use of CSMP as a valuable tool in methods and content courses offered for pre-service training. To promote these ideas, publishers and authors of mathematics texts should be encouraged to incorporate ideas and problems from the CSMP curriculum and teacher training institutions should be made aware of the program and its special characteristics for preparing teachers in mathematics.

CSMP's Effects on Students

1. The most important conclusion is that CSMP does teach problem solving skills better than the standard textbook curricula. It cannot be determined whether this result is due to a) the special CSMP "languages" (arrows, strings and Minicomputer), b) the CSMP content and curricular organization, including especially its spiral approach, or c) the classroom methods espoused in the teacher training and prescribed in the Teacher's Guides. Nevertheless, this finding is a demonstration that problem solving skills can be taught successfully by immersing students in a mathematically rich environment of problems and activities instead of requiring them to learn the different strategies in a highly organized, almost algorithmic, form.

2. The original CSMP belief that merely doing computations as part of the problem activities will develop computational skills as well as the traditional program does is not justified by test data. CSMP students fall somewhat behind their peers, particularly in the upper grades with the multiplication and division algorithms, unless teachers supplement the program with computation practice. However, modest supplementation of CSMP has been shown to eliminate this difference. This supplementation can be done unobtrusively without detracting from the strengths of the program, though it does add somewhat to the length of time normally allocated to mathematics. This finding indicates that regular practice in computation is necessary for the development of computation skills but such practice need not be in the form of long repetitive blocks of drill work.

3. The CSMP belief that emphasizing problems in a group setting and posing problems directly in the CSMP languages will develop adequate skills in word problems is justified by test data. Furthermore, CSMP students are better able to solve more complex, multi-step word problems, particularly those requiring inverse operations. This finding indicates that the ability to do one-step, computationally-oriented word problems of the type emphasized in standardized tests (an objective of dubious value) need not require the heavy emphasis on practicing these problems that exists in many classrooms.

4. There are two ways in which The evaluation results, particularly in the upper grades, probably underestimate the CSMP effects on students. First, these results are based on usage of experimental materials by teachers who had little CSMP experience. CSMP student effects should be appreciatively larger when more experienced teachers use the revised program.

   Second, CSMP students probably know more mathematics than the evaluation results indicate. These results were based on process oriented tests in which specific CSMP terminology and content were consciously avoided, in order to be fair to Non-CSMP students. Thus, tests in the less traditional content areas had to be very general, almost intuitive, and "non-technical". As such, they produced somewhat mixed results, for example, higher scores for CSMP students on tests of probability and pre-algebra but occasionally lower scores on geometry tests. However, it is to be expected that CSMP students will perform much better than Non-CSMP students on tests of content that is highly specific to CSMP, for example, the concepts of randomness in probability, and parallel projections in geometry.

   There is a need for additional evaluation of the program to investigate these two considerations.

5. CSMP has positive effects on students at all ability levels. Although the magnitude of the gains is sometimes larger for higher ability students than for lower ability students, the general result refutes many educators' belief that the teaching of mathematics to low ability students should concentrate almost exclusively on the basics, in a direct instructional mode, with heavy emphasis on rote, "how-to" methods of learning computational skills. The CSMP experience has shown that these students benefit from CSMP's spiral, problem solving approach just as other students do; in particular, the pictorial languages of CSMP allow young students with limited verbal skills to visualize mathematical concepts that would otherwise be inaccessible to them.

6. The CSMP feature which may be most widely applicable is the spiral organization of the curriculum. The CSMP organization and scheduling of topics is unusual in the degree to which concepts are integrated across different topics and repeated in short segments separated by several days. The gap between segments provides time for the material to "sink in"; later segments provide a natural review of earlier segments (which is very different from the massive review often required at the end of an extended period of study on a particular topic). CSMP teachers report that students generally like this approach. Nevertheless, it raises questions concerning, for example, the mastery of concepts which are prerequisites in future lessons, the need for reteaching concepts because of forgetting, the adequacy of the spiral approach in maintaining skills, and the ability of the teacher to deal with varying levels of understanding of a concept without recourse to tests built into the curriculum. The overall effect of CSMP's spiral curriculum, in combination with CSMP's other distinctive characteristics, is positive, but not enough is known about how the mechanics of the spiral curriculum affect student learning at different points in time. Because CSMP is unique in its use of this kind of spiral approach, research directed towards its specific effects would be beneficial to the whole educational community.

CSMP's Implementation

1. CSMP maintained the integrity of its point of view throughout the development, sometimes at the cost of reduced marketability of the product. The program costs more to adopt than a textbook, requires teacher training, and needs a skilled and influential coordinator to explain its unique approach. Nevertheless, it has been used successfully in a variety of contexts, and districts have been able to make local adaptations of the program while still retaining CSMP's distinctive and positive features. These adaptations should be encouraged; they mold the program to fit local needs and increase districts' sense of ownership of the program. Nowhere are adaptations more apparent than in the area of teacher training; many districts have been forced to scale down the CSMP-recommended training effort and have shown ingenuity in doing this successfully in many different ways. The fact that districts have continued to use the CSMP materials, in spite of a drastic curtailment in services available to them, supports the developers' decision to maintain the distinctive features of the program.

2. The role of the local coordinator in implementing and managing the program in school districts is vital to the success of CSMP; without a skilled and influential person at the helm a solid implementation of CSMP is almost impossible. Increasingly, curriculum reform has come to be seen as locally initiated and local districts are reluctant to import whole programs directly. CSMP's success in a district depends eventually on the acceptance by district teachers and administrators of CSMP's "point of view", for example, the spiral approach, emphasis on whole class instruction, and rich problem situations. But prior to implementation, the coordinator needs to gain consensus for the need to improve mathematics education in the district in ways that are consistent with the CSMP approach. Thus, curricular reform begins locally; external programs may be ready and available for schools to use, but they must be perceived as something <u>needed by</u> the district rather than merely <u>offered to</u> the district.

3. The role of teacher training in the program is crucial. There is not enough evidence available to directly trace the effects of training on student outcomes but the experience of learning CSMP and teaching it in the classroom will probably have a lasting effect on teachers regardless of the formal curriculum they use. Both the mathematical knowledge of teachers and their skill in teaching students to think should be enhanced.

   An important part of learning to teach CSMP, perhaps the most important part, comes from the teachers' day-to-day experiences as they teach the lessons. The highly prescriptive nature of the CSMP Teacher's Guides are very unusual in the extent to which they specify for each lesson both the sequence of tasks and the questioning techniques. Throughout the Guides, and in teacher training workshops, teachers are expected to engage in the same kinds of problem solving activities as their students will be encountering. It is important to determine the extent to which teachers have improved the way they present lessons, ask questions, and deal with student responses in Non-CSMP contexts. If this aspect of the Guides promotes valuable and generalizable teaching skills, then similarly detailed model lessons may be an effective way of improving teaching generally.

4. Not enough is known about the relationship between teacher characteristics and crucial aspects of the program. The objective teacher characteristics investigated during the evaluation of the program, such as mathematics background and teaching experience, appear to be relatively unimportant to the program's success in the classroom. Nevertheless, teacher success is undoubtedly related to the teacher's attitudes toward the CSMP philosophy and motivation for teaching it. For example, it seems likely that the way a teacher goes about leading the class towards the solution of problems affects the degree to which the students will adopt problem solving attitudes towards mathematics. This issue should be investigated and the results disseminated to coordinators.

Comments

1. The status quo of mathematics education makes curricular innovation almost impossible. Content and sequencing of topics have always been heavily influenced by the very traditional, computationally oriented view of mathematics held by many school administrators, principals, and teachers. Recent increased use of commercial standardized tests, and state and locally mandated competency tests, together with public dissemination of the results of these tests, has narrowed the traditional focus further so that, to a large extent, these tests effectively control the curriculum. (An example of the effect of this influence is the decision by some CSMP teachers to teach the traditional subtraction algorithm in second grade as usual, in spite of the fact that CSMP employs a different algorithm and intentionally delays its presentation until third grade. This decision naturally disrupts later learning.)

   This accountability movement has placed increased pressure on teachers to have students achieve these goals, even to the exclusion of other less well measured goals such as problem solving, or less well understood content such as probability. In the future, successful curricular innovations are likely to be limited to those which can provide advance proof of those positive student effects which are valued by the public as represented by school boards and administrators.

2. The CSMP curriculum is compatible with some recent trends in mathematics and mathematics education:

   the call for increased problem solving in the curriculum together with continued poor performance nationally (indicated by recent data on "non-routine" problems from the National Assessment for Educational Progress),

   the recommended increase in mathematics requirements for high school graduation,

   the recognized need to provide teachers with more mathematics training,

   the burgeoning use of computers in schools, and

   the increased interest in discrete mathematics and algorithmic thinking in mathematics.

   CSMP's value will increase as these trends continue.

Discussion

It is very hard to be neutral about CSMP, and not many people are, including teachers. Even the strongest critics must admit that CSMP students are better than Non-CSMP students in some kinds of mathematical thinking, regardless of any possible shortcomings in computational skills. And even the strongest proponents must admit that CSMP is hard to implement. It is worth considering what aspects of the program are most important in producing student learning (and should be saved and exported) and what aspects of the program make it hard to implement (and should be eliminated). To the extent that the answers to these questions are the same, there is a dilemma. But in the author's opinion (and the rest of this report is all opinion), it is possible to keep the baby and throw out at least a little of the bath water.

1. Teachers don't usually complete a full year's work in the curriculum now, partly because the lessons are too long and occasionally require a second day, and partly because teachers take class time to supplement the program for computation practice. Partly for this reason, and partly because they just don't see the point, teachers drop lessons in Probability and Geometry. Therefore, drop these strands, or at least reduce them by 2/3 or put them in a separate optional block which is not part of the schedule. Reduce the longer lessons by eliminating the last third of the lessons.

2. Teachers supplement the program with computational practice and this supplementing does improve student skills in multi-digit algorithms. Therefore, build time for computation practice into the schedule, add worksheets specifically designed for this which can be sent home as homework, and, as an important psychological change, admit to the teachers in the Guides and training materials that there isn't enough computation practice and that it is all right to spend time doing it.

3. Teachers complain that the spiral is too loose, i.e., too much time passes between one instance and another of a given concept; students forget what happened last week (or, sometimes, last month). Therefore, close up the spiral to some extent by reorganizing the lessons into blocks. Some care is required in making this change. One advantage of the spiral is that the constantly changing lessons make mathematics class more interesting.

4. As part of the same change, build in tests at the end of each of these blocks. The curriculum does not now contain tests or behavioral objectives and most teachers would like to have them for grading purposes and, a more difficult problem, to determine which children need extra help before the class goes on. While it may often be true that proceeding to the next lesson while children still don't understand the last one is good pedagogy, it is obvious that there are cases where the teachers should stop and review. Therefore, these tests should contain standards, at least as rough guidelines to help teachers make this most difficult decision. This will not trivialize the curriculum and teachers would be free to ignore the tests if they wish.

5. In the same spirit, individual lessons should be accompanied by objectives in fairly concrete terms and in some order of priority. Many teachers don't need this help; they can figure it out for themselves. For other teachers it would be very helpful, particularly during the many occasions when they must make choices about what to do in the few minutes left, whether to do another example or not, or whether it's alright to drop this portion of the lesson.

6. One recommendation, which has already been met, is the development of a self training manual for teachers. This will be enormously useful to coordinators, especially in districts with heavy CSMP usage, where new teachers have to be trained every single year, perhaps one or two a time, because of normal turnover.

7. So far, all the recommendations seem fairly safe. If adopted, they will not destroy those aspects of CSMP which produce such good thinking skills, namely the CSMP languages, the mathematical situations so nicely developed in the Teacher's Guide, and the student materials with their wonderful, colorful problems. But CSMP's cost does prevent its widespread use; its consumable materials prevent it from looking like a book (and being an adopted "textbook").

   Therefore, put all the workbooks and worksheets into a single, reusable, hard cover book. Systematically reduce the use of color so that many of the problems can be put in reproducible master form for local duplication. This is a drastic suggestion and would admittedly have a negative effect on the CSMP languages, the mathematical situations and the student materials.

All of these recommendations are attempts to normalize the program, at least in appearance, without seriously damaging its best characteristics. The conceptual underpinnings of the program, the mathematics and pedagogy, are very healthy and would easily survive these changes.

Appendix A

CSMP Evaluation Review Panel
1974-1984

Ernest House, Chairman
University of Illinois

Robert Dilworth
California Institute of Technology

Peter Hilton
State University of New York at Binghamton

Stanley Smith
Baltimore County Public Schools

Leonard Cahen (1974-1983)
Arizona State University

Andrew Porter (1983-1984)
Institute for Research and Teaching
Michigan State University

Appendix B

List of Evaluation Reports

The present report is the 51st formal evaluation report dealing with CSMP. Two other reports are summary reports, and were completed in 1983 under McREL auspices:

Summaries of Evaluation Reports, CSMP
User's Manual for MANS Student Data Tape

The first of these reports provides a one-page summary of all Evaluation Reports, 1974-1983, and all Joint Research Studies, 1981-82. The second describes the layout, on magnetic tape, of all class, student and item data from 1979-82, as well as a complete listing (76 pages in all) of all MANS items from the Extended Pilot Tests, grades 4-6, and Joint Research Studies, grades 2-5.

The next page lists the titles of the 48 volumes of the Evaluation Report Series from the CSMP Extended Pilot Test. Each Evaluation Report is labelled M - X - N,

where M is the year of Pilot Study (1973-74 = Year 1, ...1981-82 = Year 9)

X is the type of data being reported: A = overview or summary
B = student achievement
C = non-test data

N is the number within a given year and type of data

Evaluation Report Series

1974	1-A-1	Overview, Design and Instrumentation
	1-A-2	External Review of CSMP Materials
	1-A-3	Final Summary Report Year 1
	1-B-1	Mid-Year Test Data: CSMP First Grade Content
	1-B-2	End-of-Year Test Data: CSMP First Grade Content
	1-B-3	End-of-Year Test Data: Standard First Grade Content
	1-B-4	End of year Test Data: CSMP Kindergarten Content
	1-B-5	Test Data on Some General Cognitive Skills
	1-B-6	Summary Test Data: Detroit Schools
	1-C-1	Teacher Training Report
	1-C-2	Observations of CSMP First Grade Classes
	1-C-3	Mid-Year Data from Teacher Questionnaires
	1-C-4	End-of-Year Data from Teacher Questionnaires
	1-C-5	Interviews with CSMP Kindergarten Teachers
	1-C-6	Analysis of Teacher Logs
1975	2-A-1	Final Summary Report Year 2
	2-B-1	Second Grade Test Data
	2-B-2	Readministration of First Grade Test Items
	2-B-3	Student Interviews
	2-C-1	Teacher Questionnaire Data
	2-C-2	Teacher Interviews, Second Grade
	2-C-3	Teacher Interviews, First Grade
1976	3-B-1	Second and Third Grade Test Data Year 3
	3-C-1	Teacher Questionnaire Data Year 3
1977	4-A-1	Final Summary Report Year 4
	4-B-1	Standardized Test Data, Third Grade
	4-B-2	Mathematics Applied to Novel Situations (MANS) Test Data
	4-B-3	Individually Administered Problems, Third Grade
	4-C-1	Teacher Questionnaire Data, Third Grade
1978	5-B-1	Fourth Grade MANS Test Data
	5-B-2	Individually Administered Problems, Fourth Grade
	5-C-1	Teacher Questionnaire and Interview Data, Fourth Grade
1979	6-B-1	Comparative Test Data: Fourth Grade
	6-B-2	Preliminary Test Data: Fifth Grade
	6-C-1	Teacher Questionnaire Data: Grades 3-5
1980	7-B-1	Fifth Grade Evaluation: Volume I, Summary
	7-B-2	Fifth Grade Evaluation: Volume II, Test Data
	7-B-3	Fifth Grade Evaluation: Volume III, Non-Test Data
	7-B-4	Re-evaluation of Second Grade, Revised MANS Tests
	7-B-5	Achievement of Former CSMP Students at Fourth Grade
	7-B-6	Student Achievement, Rapid Implementation Model
1981	8-B-1	Sixth Grade Evaluation, Preliminary Study
	8-B-2	Evaluation of Revised Second Grade, MANS Blue Level
	8-B-3	Evaluation of Revised Third Grade, MANS Green Level
	8-B-4	Three Evaluations of Gifted Student Use
	8-C-1	Preliminary Study of CSMP "Graduates"
1982	9-B-1	Sixth Grade MANS Test Data
	9-C-1	Sixth Grade Evaluation: Teacher Questionnaires

Appendix C

Submission to the Joint Dissemination Review Panel

Approved March 13, 1984

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 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.

	Installation (Year 1)			Subsequent Years		
	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 succeding 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 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:

Type of Community		Geographic Location	
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	
	CSMP	Non-CSMP	CSMP	Non-CSMP
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 CSMP	Number of Classes Non-CSMP	Number of Items	Adjusted Means CSMP	Adjusted Means Non-CSMP	Signif at .05	In Favor of
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]	6	6	55	36.2	35.2		CSMP
	Stanford Ach	6	6	percentile ranks	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	*	CSMP
	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

[1]Total Math; This test does not have a separate computation section
[2]40 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.

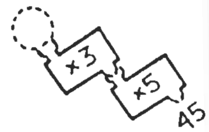
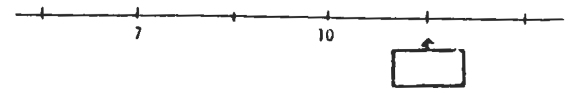
Reliability The reliability/internal consistency (KR20 corrected by Spearman-Brown for an equivalent 20-item test) was calculated for each of the 85 individual Mans tests. The reliability was above .80 for 72 of these tests; between .75 and .80 for 10 tests; and below .75 for 3 tests (.68, .71, and .72). 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.

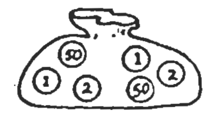
Correlations with Other Measures The median correlations between Reading scores and Total MANS scores were .60, .57, and .56 for grades 2, 4 and 6 respectively. The median correlations between standardized computation scores and Total Mans scores was .63. The median correlations between Total Mans and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

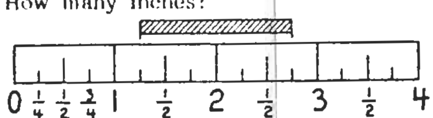
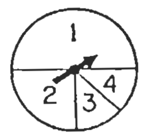
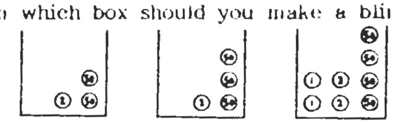
Student and Teacher Ratings Mean teacher ratings of importance of individual MANS tests, collected in 4th and 5th grades, were 4.3 and 4.1 on a 5-point scale. Mean rating of how well students liked individual MANS Tests, collected only in 4th grade, was 3.0 on a 4-point scale.

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, Relation-ships	2 - 10 8 - 4 10 - 5 6 - ? Which is Larger? 585 + 250 or 580+290 What is the missing number? 28, 25, ? , 19, 16, 13	Will 700 be in any of the boxes? . . | 60 | 67 | 74 | 81 | 88 | . . What are the missing numbers? ?, 50, ?, 200, 400, 800 	Which is larger? 5/2 or 5/4 0.9 or 0.11111 Name a fraction (decimal) that is: larger than 1/3 but smaller than 1/8 larger than 0.2 but smaller than 0.3 Label the missing number 

Multiple Answers	Write # Sentences about 8
	8 = 9 - 1
	8 = 3 + 4 + 1
	8 = 2 x 4
			Take out 3 balls together
			Add to get total score
			Give all possible scores
			

Mental Arith	300 - ? = 250	Hit = gain 5, miss = lose 1 start with : 3 below zero end with : 5 above zero # of misses : 2 # of hits : ?	12 x 75 = 900 13 x 75 = ? 1/2 x ? = 40 0.75 - 0.5 = ? scratch work not allowed
Word Problems	One-stage, with pictures, and read to the students	Two-stage Miscellaneous data Word problem approximations	Three-stage Word problems with fractions, decimals Novel word problems
Estimation	90 - 12 is in which interval 0 - 10 - 50 - 100 - 500?	602 is about ? as large as 298? 2, 5, or 10 times	8 1/2 ÷ 8 is: <1 or =1 or >1? Which interval contains 1,002.5 ÷ 21.5? 0 - 1 - 10 - 20 - 50 - 100?
Number Represen-tations	Write "two thousand, eleven" 100 more than 901 is ?	How many inches?  Name the 2nd largest 4-digit number using only 2, 5, 7, 8?	Which are equivalent to 1/3? 2/6 11/31 3/15 4/12 50/150? Which are equivalent to 3/4? 0.750 0.075 0.75 7.5 75.0? Put an arrow at 1.35 in. 
Predicting	(4th and 6th grades only)	Spin 100 times How often is the score 2? 	Spin both spinners together How often is the total > 9?  From which box should you make a blind draw? 
Pre-algebra	(sixth grade only)		(1)┬(3) = 1 + 2 + 3, so (1)┬(99) − (2)┬(99) = ? If k + 2 + k + 7 = 13, then k = ? Apply transformations - geometric rotations and/or symbol reversals - to various figures If q = 5, then 2 x q^2 = ?

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-McGinitie 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 Size (difference in adj. means - 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.

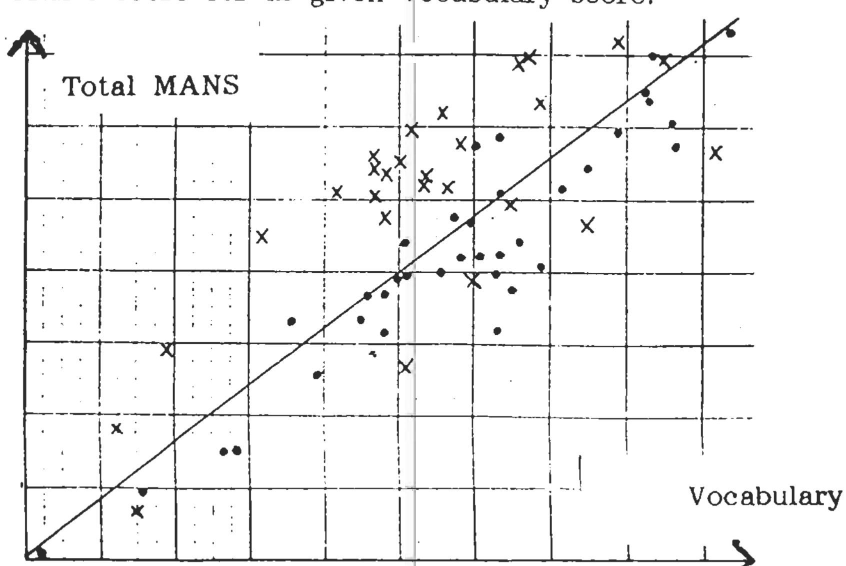


Fig 1. 6th Grade Class Means
(x = CSMP class, o = Non-CSMP)

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.

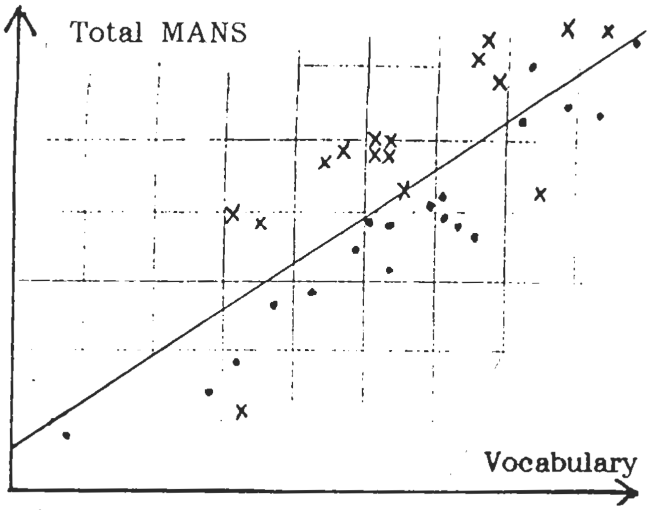


Fig 2. 6th Grade School Means
(x = CSMP school, o = Non-CSMP)

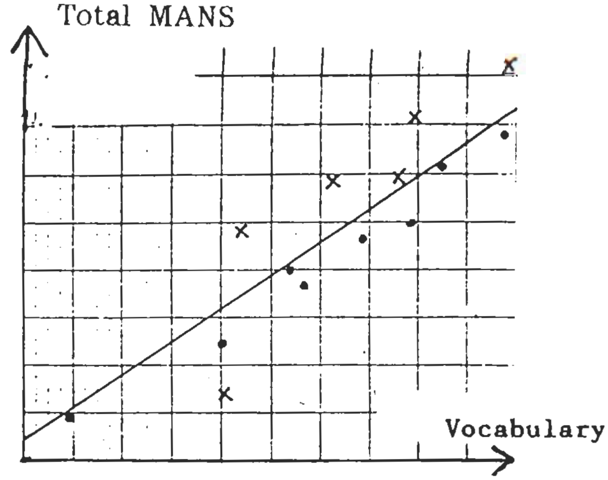


Fig 3. 6th Grade District Means
(x = CSMP district, o = Non-CSMP)

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.

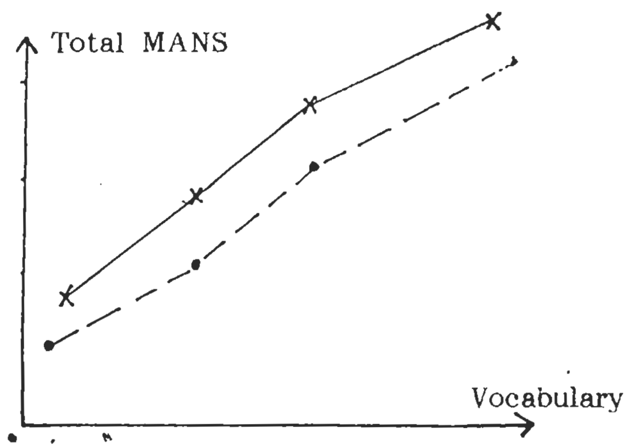


Fig. 4. Student Means Grouped by Reading
(x = CSMP students, o = Non-CSMP)

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 instuctional 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 CSMP	Adjusted Means Non-CSMP	Signif[1] at	Number of Categories Tested	Number of Categories Significant ($p < .05$)
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.

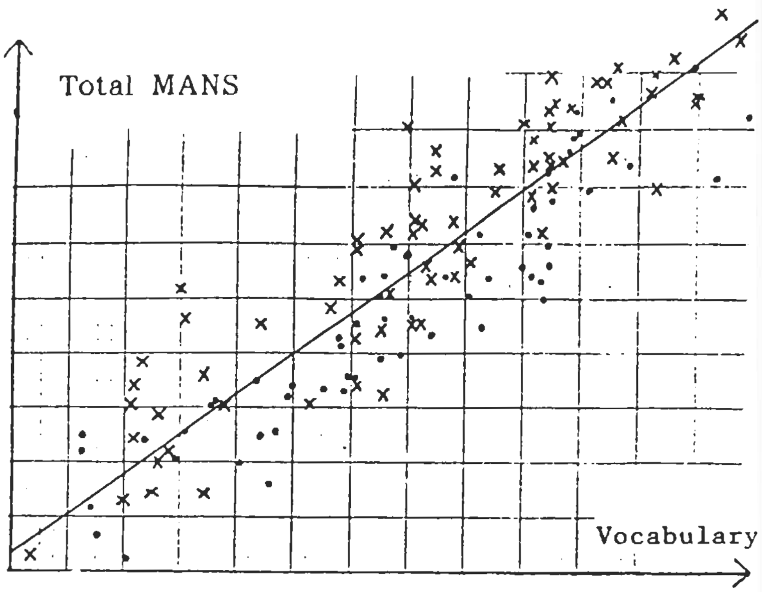


Fig. 5. Second Grade Class Means

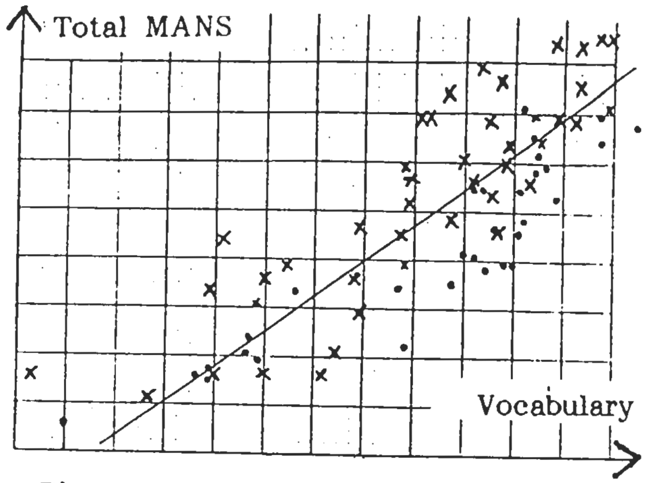


Fig. 6. Third Grade Class Means

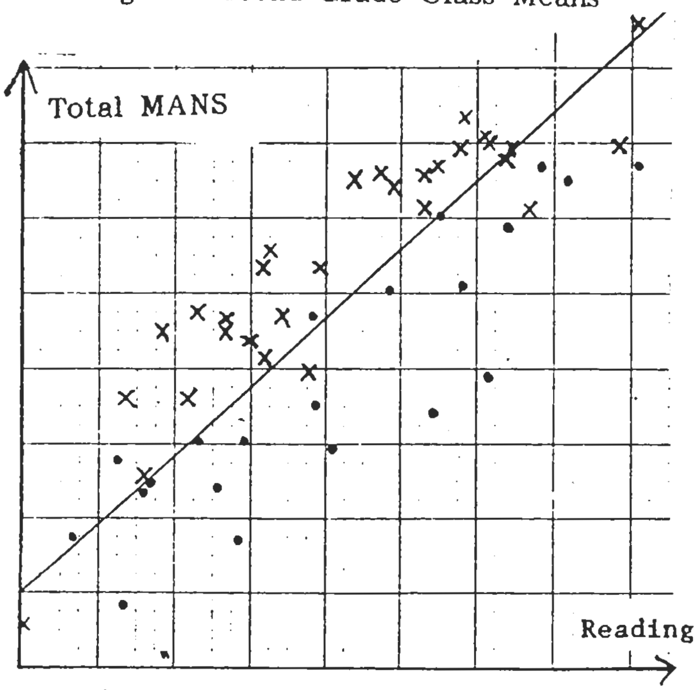


Fig. 7. Fourth Grade Class Means

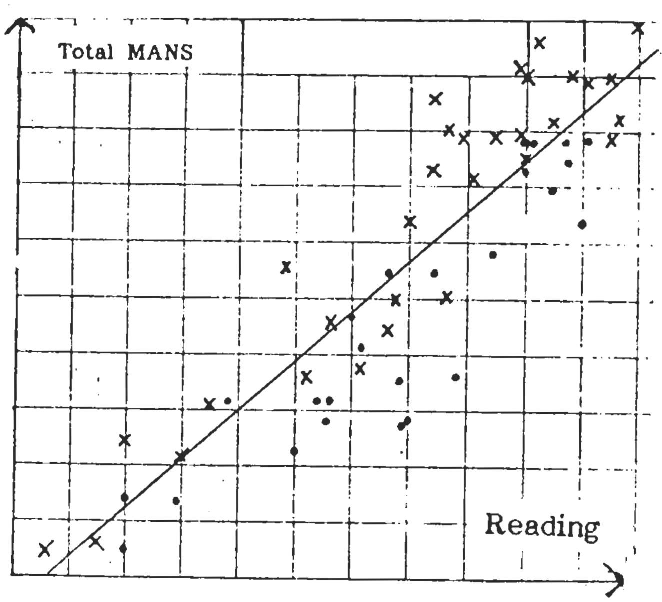


Fig. 8. Fifth Grade Class Means

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.

Appendix D

Description of CSMP Materials

This Appendix gives a partial list of materials developed by CSMP. Most items, but not all, were still available as of this report.

Curriculum Materials

CSMP instructional materials are available in classroom sets for each of grades K-6. Included are Teacher's Guides, workbooks, worksheets, storybooks, teacher display items, and a variety of manipulatives.

There are storybooks for three age groups: 5-6 (13 books), 8-12 (9 books) and 10-14 (5 books). There are story workbooks for two age groups: 7-11 (3 books) and 9-14 (11 books). The books are usually 16 or 32 pages, printed on newsprint. There is a Spanish edition of CSMP for grades 1 and 2.

The Elements of Mathematics is a textbook series for gifted seventh-twelfth graders. There are three descriptive booklets, 16 chapters (and Answer Keys) for Book O, Intuitive Background, and Books 1-12 (plus Answer Keys). Three books, proceedings of CSMP International Conferences, address the teaching of (1) probability and statistics, (2) algebra, and (3) geometry at the pre-college level.

TOPS: A program in the Teaching of Problem Solving contains about 100 detailed activities, organized as a supplement to the standard curriculum, grades 3-8. The activities are based on CSMP material.

Descriptive Materials for Potential Adopters

The CSMP Brochure - Contains initial information in detail.

The CSMP Curriculum Flyer - A one page presentation of reasons for using CSMP.

Filmstrip - CSMP: A Problem-Solving Curriculum for the 1980's.

TOPS* Announcement - General information about the Teaching of Problem Solving, activities that grew out of CSMP.

CSMP in Action - A Manual consisting largely of transcripts of actual lessons.

Preview Packet - A glossary of CSMP pictorial languages, sample lessons from all strand areas and all grade levels, and representative student materials.

CSMP Gifted Education - A pamphlet explaining CSMP usage with gifted students.

CSMP Compensatory Education - A pamphlet explaining why CSMP usage with low achievers.

CSMP Supplemental Usage - Description of materials which may be used without prior CSMP training.

CSMP Implementation Workshops - A pamphlet describing the workshops, their location, and a sample workshop day.

CSMP Readability Study - A pamphlet describing in detail the results of a CSMP readability study.

CSMP Social Fairness Report - A pamphlet describing in detail the results of a study to determine racial, gender and age equity.

CSMP Pre-Service/In-Service Packet - A resource booklet for mathematics educators with pre-or in-service responsibilities.

Profile of the Comprehensive School Mathematics Program - A 10-page document prepared for the National Commission on Excellence in Education.

Scope and Sequence - A K-3 Scope and Sequence Chart and a pamphlet entitled CSMP Summary of Content, Grades 4-6.

Information about Minipackage Samplers - A description of three CSMP minipackages.

Information About Examination Sets of Materials - A lending library is operated for the use of official search committees.

Detailed Scope and Sequence for Grades K-3

Information for Title IV-C Adopters

Chapter I Resource Handbook

Minipackages - Descriptions of Mini-computer games, attribute games, and the language of arrows in the study of relations.

CSMP Probability and Statistics - A collection of papers on the teaching of probability and statistics in the curriculum.

MANS (Mathematics Applied to Novel Situations) Test Information Packet

Sample Sets of Instructional Materials

Math Play Therapy (2 volumes). A description of the CSMP activities and games used with fourth and fifth graders classified as "slow learners".

Appendix E

List of School Districts Participating in MANS Testing

Arizona, Globe

District of Columbia

Georgia, Polk County

Hawaii, Wahiowa

Kentucky, Jefferson County

Louisiana, Mississippi State
New Orleans

Maine, Portland

Maryland, Baltimore County

Michigan, Ann Arbor
Bedford
Detroit
Livonia
Marquette

Missouri, Archdiocese of St. Louis
Ferguson-Florissant

Missouri, Ladue
Maplewood-Richmond Heights
Normandy
St. Louis
University City

New York, Bronx
Guilderland
Hartsdale
New Hartford

Pennsylvania, Philadelphia

Tennessee, Clarksville

Virginia, Albermarle County

Wisconsin, Glendale
Janesville
Madison

Wyoming, Gilette

Appendix F

Individually Administered Problems

At two grade levels, third and fourth grades, sets of problems were constructed and administered individually to samples of students in CSMP and comparable Non-CSMP classes. The studies were conducted in St. Louis area schools during the first year of the Extended Pilot Tests.

At each grade level, two sets of problems were developed, each requiring 30-45 minutes for a single administration. Sampling was based on a stratified random sampling plan based on scores on an ability test, the Kuhlman Andersen test. Half of each group of selected students were given one set of problems, the other half of the group took the other set of problems. The numbers of students tested are shown below.

Numbers of Students and Classes
Participating in Individually Administered Problems

		Number of Students Tested[1]		Number of Classes Represented	
		CSMP	Non-CSMP	CSMP	Non-CSMP
Third Grade:	Problem Set A	17	16	5	4
	Problem Set B	18	18	5	7
Fourth Grade:	Problem Set A	30	30	5	5
	Problem Set B	24	24	6	6

[1] In third grade, these entries represent pairs of students; the interviews were conducted with two students acting as a team.

For each individual problem, an extensive protocal was developed, piloted and revised. Students were asked to explain their answers, or to show why a sample problem was correct or incorrect. Each interview was tape recorded and coded. In order to investigate CSMP - Non-CSMP differences, an analysis of the responses was carried out by assigning scores to the type and quality of response

In third grade there were several problems on which CSMP students did better than their Non-CSMP counterparts:

> Students were shown a set of completed calculations which "a student at another school" had done (e.g. 6 X 13 = 53). They were then asked to rapidly indicate which answers "could be right" and which ones were "probably wrong". Finally students were asked to go back to each probably-wrong answer and tell why they thought the given answer was wrong.
>
> CSMP students made a higher average number of correct decisions (70% versus 64%) and their explanations of wrong answers were more likely to be acceptable (89% versus 77%). The largest differences between CSMP and Non-CSMP students occurred for students of about average ability.

Students were shown a partial calendar with "69 cents" written under each day of the week and told that Bill gets 69 cents every day this week. They were then asked to describe the fastest way, on a calculator, to figure out "how much Bill would earn by the end of the week."

CSMP students were more likely to suggest a multiplication process (88% versus 53%) and less likely to suggest an addition process.

Students were asked to quickly estimate the number of dollar bills that would be needed to purchase seven items whose costs were as shown below, "but we don't want to take any more (money) then we'll need":

$1.22
1.81
1.51
1.53
1.33
1.33
1.39

A higher proportion of CSMP students (50% versus 34%) gave good answers, defined as 10,11 or 12 and a lower proportion (12% versus 25%) gave poor answers, i.e., <8 or >14.

Students were shown an undifferentiated set of "people pieces", which were simplified figures that were either tall or short, fat or thin, boy or girl, and red or blue. They were then asked to put them in piles so that all the pieces in a pile were similar in some way and so that the piles were all different from one other. They performed this classification in as many different ways as they could.

CSMP students were able to make more complex sorts than Non-CSMP students, the average "best effort" being 3.0 dimensions simultaneously (versus 2.2 dimensions for Non-CSMP students).

Students were asked to figure out the Interviewer's "secret" rule for the people pieces, by offering individual pieces to which the interviewer would respond with a "yes" or "no," according to whether the offered piece fit the secret rule. Examples of the secret rule were "blue" and "fat and tall."

CSMP students needed to offer fewer pieces to figure out the rule. In four trials, the average total number of pieces needed was 14.8 for CSMP students versus 19.7 for Non-CSMP students.

On the remaining third grade problems, described briefly below, there were virtually no differences between CSMP and Non-CSMP students.

Estimating the sum of the <u>ten</u> (emphasized to students) numbers:

$5 + 5 + 4 + 6 + 3 + 3 + 6 + 6 + 5 + 4.$

Estimate the largest and smallest answer could be.

Quickly estimate the answer to $\frac{6}{5} + \frac{9}{8} + \frac{2}{1} + \frac{5}{4}$.

Given 1,573. Write the number obtained by reversing the '7' and the '3'. Is the new number larger or smaller? By about how much? Then reverse the '5' and '3' and the '1' and '3' (but without writing the new number) and answer the same questions.

Figure out 6 X ? = 138 on a calculator using the 'X' button. Use repeated trials until the correct answer is obtained.

Students were shown the "people pieces" problem described earlier, except that this time, a standardized sequence of pieces that another student had supposedly done was shown together with Interviewer responses about whether they had fit the secret rule. Students had to figure out what the secret rule was.

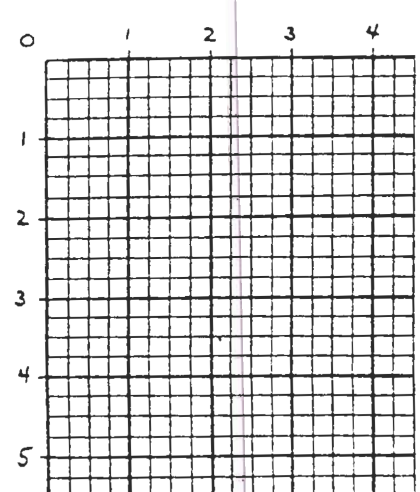
Students had to determine the Interviewers "secret rule" with the people pieces, based on being shown a sequence of pieces that did not fit the rule.

The total mean score across all items was 50.3 for CSMP students versus 42.5 for Non-CSMP students. The largest difference occurred at the average or slightly above average ability levels.

Fourth Grade. On two of the six problems in fourth grade, CSMP students had significantly higher scores using Analysis of Covariance on class means:

Students secretly drew a number out of a hat (but the interviewer knew that the number was 24) and answered a series of questions about their secret number. The questions dealt with concepts of order, whole numbers, negative numbers, multiples and divisors. The students were also asked whether the question itself was a good one. (For example, after finding out that the number was less than 100, a question about whether it was less than 200 was not a good question.) The adjusted mean scores, out of eleven, were 9.0 for CSMP versus 7.4 for Non-CSMP.

Students were given sheets of graph paper, with different ways of labelling the lines and some lines heavier than others. An example is shown below.



CSMP students were better able than Non-CSMP students to figure out how many little squares were shown, were more likely to use a length-times-width method, and were more likely to use the guide numbers in the margins versus a one-at-a-time counting process. They were also better able to do related problems of figuring out the area when pieces were combined or when one of the figures had a "hole" in it. Finally they were better able to figure out how many squares were on a partly hidden role of paper marked off at every second square.

On the other four problems, CSMP students had higher adjusted scores but the differences were not significant.

Students were given a calculation to do mentally (e.g. subtract 244 from 543). CSMP students got more problems correct (51% versus 48%) and were more likely (33% versus 22%); to use a method other than brute force, for example, 543 - 244 is 1 less than 300, i.e., 299).

Students were shown a computation problem (e.g. 277 + 277) and then shown three other computation problems (e.g. 277 + 177) and asked if and how the answer to each of those three would help with the original problem.

Students were shown a series of subtraction problems (e.g. 260 - 211) and asked to quickly indicate which interval (0 - 10 - 50 - 100 - 500 - 1000) contained the answer.

Students were asked to identify the interviewer's secret number, which was between 0 and 99, by asking a series of "yes" or "no" questions.

The adjusted total scores across all items were 32.1 for CSMP versus 27.6 for Non-CSMP.

Appendix G

Abstracts of MANS Tests

This appendix gives an abstract and sample item(s) for individual MANS tests used in any of grades 2-6 in the revised MANS tests (revised 1981-1982). The tests are grouped by category, and the categories appear in the following order.

Process Categories:

C: Computation
E: Estimation
M: Mental Arithmetic
N: Number Representations
R: Relations & Number Patterns
U: Elucidation
W: Word Problems

Special Topic Categories:

A: Algebra
G: Geometry
L: Logic
O: Organization of Data
P: Probability

Category C: Computation

C1 Whole Number Computation

Abstract: Given straightforward computation problems involving whole numbers, produce exact answers (by calculating on paper if necessary). The items do not have the multiple choice response format but are similar in range and difficulty to those found in the standardized achievement tests of the appropriate grade level.

Grade Levels: 2, 3, 4, 5, 6

Examples (from Grade 4):

$$\begin{array}{r} 352 \\ +683 \\ \hline \end{array} \qquad \begin{array}{r} 675 \\ -469 \\ \hline \end{array} \qquad \begin{array}{r} 143 \\ \times\ 5 \\ \hline \end{array} \qquad 6\overline{)492}$$

C2 Fraction Computation

Abstract: Given straightforward computation items involving simple fractions, produce exact answers (by calculating on paper if necessary). Though the items do not have the multiple choice response format, they are similar in range and difficulty to those found in the standardized achievement tests of the appropriate grade level.

Grade Levels: 4, 5, 6

Examples (from Grade 5):

$$\frac{3}{5} - \frac{1}{5} = \square \qquad \frac{1}{2} + \square = 1 \qquad \frac{1}{2} \times \frac{1}{2} = \square$$

C3 Decimal Computation

Abstract: Given straightforward computation items involving one and two place decimals, produce exact answers (by calculating on paper if necessary). Though the items do not have the multiple choice response format, they are similar in range and difficulty to those found in the standardized achievement tests of the appropriate grade level.

Grade Level: 6

Examples: $0.5 + 0.25 = \square$ $\quad 5 - 1.5 = \square$ $\quad 0.5 \times 0.5 = \square$

Category E: Estimation

E1 2 or 5 or 10 Times

Abstract: Given two numbers, quickly estimate whether the first is about 2 or 5 or 10 times as large as the second. A sample is worked collectively.

Grade Levels: 3, 4

Examples (from Grade 3): 65 is about _____ times as large as 12

98 is about _____ times as large as 51

E2 Estimating Intervals: Addition

Abstract: Given a computation problem involving whole number addition, and 5 fixed intervals (0-10, 10-50, 50-100, 100-500, 500-1000), determine which interval contains the answer to the problem, and put an x in the interval. By instruction, format and short time limits, students are discouraged from computing exact answers. Two or three sample items are done collectively.

Grade Levels: 2, 3, 4, 5

Examples (from Grade 2):

51 + 53	0	10	50	100	500	1000
189 + 273	0	10	50	100	500	1000

E3 Estimating Intervals: Subtraction

Abstract: The scale is similar to E2 (except that it involves whole number subtraction) and follows it directly in the test booklets.

Grade Levels: 2, 3, 4

Examples (from Grade 3):

93 - 86	0	10	50	100	500	1000
147 - 99	0	10	50	100	500	1000

E4 Estimating Intervals: Multiplication

Abstract: The scale is similar to E2 and E3 (but is devoted to multiplication with whole numbers for the most part) and follows them in the test booklets.

Grade Levels: 2, 3, 4, 5, 6

Examples (from Grade 4):

40 x 10	0	10	50	100	500	1000
4 x 29	0	10	50	100	500	1000

E5 Estimating Intervals: Division

Abstract: The scale is similar to E2, E3 and E4 (but is devoted to division with whole numbers for the most part) There are only four fixed intervals (0-1, 1-10, 10-20, 20-100) in the response format. It follows E4 in the test booklets.

Grade Level: 5, 6

Examples:

$1 \div 15$	0	1	10	20	100
$101 \div 9$	0	1	10	20	100

E6 Estimating Fractions <, =, >1

Abstract: Given a calculation (+, -, or :) of two numbers (at least one of which is a fraction or mixed number), quickly estimate whether the answer would be less than, equal to or more than 1. Students are encouraged to work quickly and not to compute exact answers before making their choices. A completed sample item is provided.

Grade Level: 6

Examples:

	Less than 1	CHECK ONE Exactly 1	More than 1
$1\frac{5}{8} - \frac{1}{128}$			
$2\frac{1}{2} \div 3$			

M1 Whole Number Open Sentences

Abstract: Given an open sentence, where the box may be either on the right or the left of the equal sign, where the numbers are large and easy to work with, and where only one operation is used, put the number in the box which makes the sentence true. By instruction and prompting, students are discouraged from "computing the long way" and are not allowed to do any figuring on paper.

Grade Levels: 2, 3, 4, 5, 6

Examples (from Grade 3)

$500 + \square = 800$

$\square - 150 = 50$

$2 \times 200 = \square$

M2 Above and Below Zero

Abstract: Given a starting score (which could be above or below zero), and how much the score went up or down, select the correct final score (multiple choice).

Grade Levels: 2, 3

Examples (from Grade 3)

Score at the start: 3 below zero
Then: Lost 4
Score at the end: 7 below zero 1 below zero 1 above zero 7 above zero

Score at the start: 2 above zero
Then: Lost 4
Score at the end: 6 below zero 2 below zero Zero 2 above zero

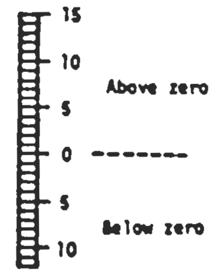
M3 Negative Hits and Misses

Abstract: Given the description of a "game" with two rules (a) each hit means a gain of 5 points and b) each miss means a loss of 1 point) and partial information on the outcome of turns, the student must deduce the missing information. Two sample items are completed collectively.

Grade Levels: 4, 5, 6

Examples:

	Started with a score of	Number of Hits	Number of Misses	Ended with a score of
Pam:	4 above zero		6	3 above zero
John:		2	0	15 below zero



(provided, but not mentioned in instructions)

M4 Fraction Open Sentences

Abstract: Given an open sentence involving at least one fraction, and one of the four arithmetic operations, complete the sentence.

Grade Level: 6

Examples: $\frac{3}{5} \div \square = 1$ $\qquad 1 - \square = \frac{3}{4}$

M5 Decimal Open Sentences

Abstract: Given an open sentence involving at least one decimal number and one of the four arithmetic operations, complete the sentence.

Grade Level: 6

Examples: $0.5 + \square = 1$ $\qquad 0.75 - \square = 0.5$

Category N: Number Representations

N1 Writing Whole Numbers

Abstract: Part I: The student must write numbers as they are read aloud by the tester.
Part II: Given a number, written in the test booklet, the student must write the number which is 1 (or 10 or 100) more than it. A sample item is worked collectively.

Grade Level: 2

Examples: Part I: Tester says, "Eight hundred twenty" (repeats)

Tester says, "Seven thousand sixty five" (repeats)

Part II: What number is 1 more than 999? ________

What number is 10 more than 495? ________

N2 1, 10, 100 or 1000 More

Abstract: Given two numbers, decide whether the first number is about 1, 10, 100 or 1000 more than the second number. (None is exactly right.) Two sample items are worked collectively.

Grade Level: 3

Examples:

	1	
	10	
4,265 is about	100	more than 4,254
	1000	

	1	
	10	
1,001 is about	100	more than 998
	1000	

N3 Constructing Numbers

Abstract: Given the use of only four digits (2, 5, 7 and 8) and the rule that no digit be used more than once, construct numbers like the smallest (or largest), the second smallest (or largest) or the closest to a given number. The constructed numbers are to be of either 2, 3 or 4 digits and sometimes restricted to a given range of numbers. Collectively, to clarify the rules, two incorrect answers and the correct one are examined for two sample problems.

Grade Level: 4

Examples: What is the second largest four digit number? _______

What is the smallest three digit number between 730 and 850? _______

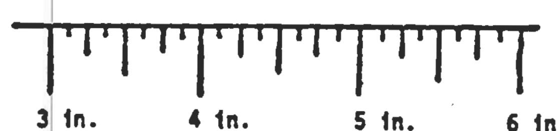
What four digit number between 2,000 and 3,000 is closest to 2,800? _______

N4 Representing Fractions

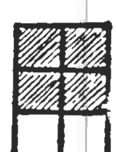
Abstract: The scale has five short subsections each containing one of two kinds of items: a fraction or mixed number is given in standard form and must be represented in another specific way or else that process is reversed and the response format is multiple choice. Instruction is largely in the form of a written question or command at the beginning of each subsection.

Grade Level: 4

Examples: Put an arrow at $4\frac{1}{4}$ inches.



How much is shaded?



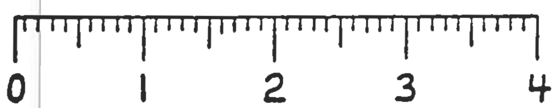
$\frac{1}{2}$ $\frac{1}{4}$ $\frac{2}{3}$ $\frac{3}{4}$

N5 Representing Fractions and Decimals

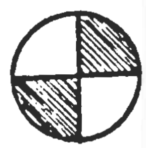
Abstract: The scale has five short subsections each containing one of two kinds of items: either a mixed number or decimal is given in standard form and must be represented in another specific way or else that process is reversed and the response format is multiple choice. Instruction is largely in the form of a written question or command at the beginning of each subsection.

Grade Level: 5, 6

Examples: Put an arrow at 1.35 inches.



How much is shaded?



$\frac{1}{3}$ $\frac{1}{2}$ $\frac{2}{3}$ none of these

(A completed sample was given.)

N6 Equivalent Fractions and Decimals

Abstract: Given a fraction (or decimal) determine which members of a set of fractions (or decimals) are equivalent to it. A sample set of four completed items is shown.

Grade Level: 5, 6

Examples: Circle all the fractions that are equal to the one in the box.

$\boxed{\frac{2}{3}}$ $\frac{9}{12}$ $\frac{4}{6}$ $\frac{3}{2}$ $\frac{10}{15}$

Category R: Relationship & Number Patterns

R1 Solving Number Rules

Abstract: Given 3 clues (i.e., pairs of numbers) in a game, determine what the secret method is (i.e., the unique rule relating each of the pairs of numbers) and then use the rule to calculate the missing number from the fourth pair.

Grade Levels: 2, 3, 4, 5, 6

Examples (from Grade 3):

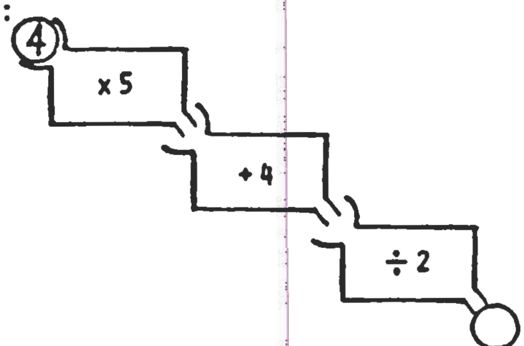
	Maria's Game		Jim's Game	
	Class said:	Maria's answer:	Class said:	Jim's answer:
First clue:	5	10	2	6
Second clue:	7	12	5	9
Third clue:	8	13	10	14
Question:	2	$\square$	$\square$	12

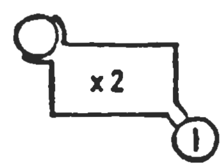
R2 Using Number Machines

Abstract: Given labelled "number machines" in sequence and either the initial or the terminating number, determine the other number. There is an introduction showing that "number machines" take in numbers; add, subtract, multiply or divide by a fixed quantity; and give out the resultant number. Then three sample items (each with a "number machine" sequence) are worked collectively.

Grade Levels: 3, 4, 5, 6

Examples (from Grade 4):





R3 Sequences

Abstract: Given an incomplete portion of an additive sequence of numbers, determine the missing number. One sample item is worked collectively.

Grade Level: 2

Examples: 28, 25, _____, 19, 16, 13

1, $1\frac{1}{2}$, 2, _____, 3, $3\frac{1}{2}$, 4

R4 Which Result is Larger

Abstract: Given two quantities (usually similar computation problems using +, -, or x) mark the one which yields the larger result, or mark them both if they are equal. By instruction, format and time limits, students are discouraged from computing exact answers. The correct response should be more easily determined by inspection than by computation. Two sample items are worked collectively.

Grade Levels: 2, 3

Examples (from Grade 2): 585 + 250 ☐ 3 x 31 ☐

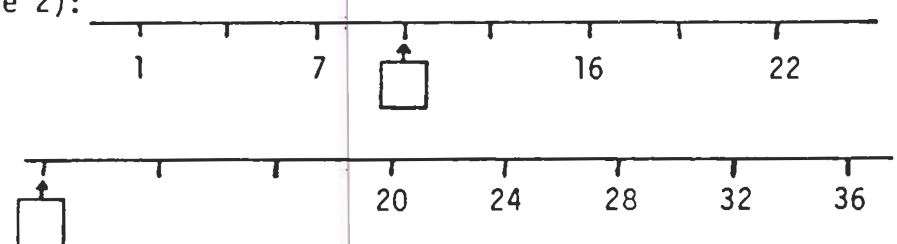
580 + 290 ☐ 31 x 3 ☐

R5 Labelling Number Lines

Abstract: Given partially labelled number lines, with varying increments, determine certain missing numbers. A sample item is worked collectively.

Grade Levels: 2, 3, 4, 5, 6

Examples (from Grade 2):

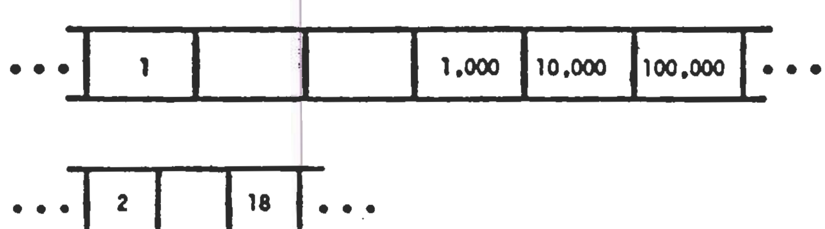


R6 Multiplication Series

Abstract: Given an incomplete portion of a multiplicative series of numbers, determine the constant multiplier involved in order to complete the portion shown. Portions of several series are shown altogether with one, two or three numbers missing from each. A sample series is examined and completed collectively.

Grade Level: 4

Examples:



R7 Which Fraction is Larger

Abstract: Given two non-whole numbers written in fractional form (a proper fraction, an improper fraction or a mixed number), circle the larger one. A completed sample item is shown.

Grade Level: 5, 6

Examples: $\frac{3}{4}$ or $1\frac{1}{4}$

$3\frac{1}{2}$ or $\frac{5}{2}$

R8 Which Decimal is Larger

Abstract: Given two non-whole numbers written in decimal form, circle the larger one. A completed sample item is shown.

Grade Level: 5 , 6

Examples: 4.999 or 5.1

1.5 or 0.58

R9 Fractions Between Two Others

Abstract: Given two fractions, write another which is larger than the first and smaller than the second.

Grade Level: 6

Examples: ____ is larger than $\frac{1}{3}$, but smaller than $\frac{7}{8}$

____ is larger than $\frac{1}{4}$, but smaller than $\frac{1}{2}$

R10 Decimals Between Two Others

Abstract: Given two decimal numbers, write another which is larger than the first and smaller than the second.

Examples: ____ is larger than 1.25, but smaller than 2.0

____ is larger than 0.42, but smaller than 0.43

Category U: Elucidation

U1 Number Sentences About 8

Abstract: Students are to produce as many different "sentences about 8" as possible, always in the form "8 = ...". Four correct answers to similar exercises about 9 are examined collectively. (9 = 10 - 1, 9 = 1 + 5 + 3, 9 = 3 x 3, 9 = 18 - 2).

Grade Level: 2

Example: My number sentences about 8.

8 = ________________________

8 = ________________________

U2 Producing Many Answers

Abstract: Given several different situations each of which poses a problem for which there are many correct solutions, produce as many of them as possible. For each situation, some potential solutions are accepted or rejected for not following the given rules as inappropriate.

Grade Level: 3, 4, 5, 6

Examples (from Grade 3):

Rules: Take out two balls.
Add the two numbers to get a score.

What are the possible scores? 6, 2, 35

Rules: Write all the two digit numbers you can.
Use only the digits 1, 2, 3.

Give all the numbers that follow the rules. 34, 22

U3 Getting to 12

Abstract: Given a starting point (0), a goal (12) and two rules, invent as many ways of reaching the goal as possible. The rules are that only the numbers 2, 3, 5 & 7 can be used along with addition, subtraction, multiplication or division. Two sample solutions (see below) are worked collectively.

Grade Level: 6

Examples:

Sample 1: 0 + 7 = 7
7 x 2 = 14
14 - 2 = 12

Sample 2: 0 + 5 = 5
5 + 3 = 8
8 + 2 = 4
4 x 3 = 12

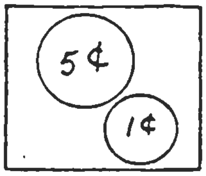
Category W: Word Problems

W1 One Step Word Problems

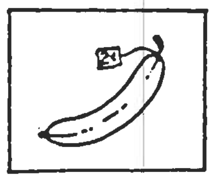
Abstract: Solve word problems in which the story (including the question) is read by the tester while the student looks at a series of cartoons and/or follows the story in the captions beneath the cartoons. Seven items require one-step solutions; two items require two.

Grade Level: 2

Examples:



Jill spent 6¢ to buy some bananas.



Bananas cost 2¢ each.

How many bananas did she buy? ____________

Jim found 3 marbles but he lost 4.

And now he has 5 marbles.

How many marbles did he have to begin with? ____________

W2 Two Stage Word Problems

Abstract: Solve word problems in which the solutions require two operations. The numbers in the problems are relatively small; the computational and reading requirements are simple.

Grade Levels: 3, 4, 5, 6

Examples (from Grade 4):

Pam gets 50¢ each week.
She always spends 30¢ and saves the rest.
How much will she save in 4 weeks? ________

Tom has 3¢ more than Ann.
Tom has 5¢ less than John.
If John has 20¢, how much does Ann have? ______

W3 Miscellaneous Word Problems

Abstract: Solve word problems which are unusual for third graders in one of several ways: requires three-stage solution, requires working backward from a given final state to an unknown initial state, requires more logical analysis than straight computation, involves proportional ratios, involves extraneous data.

Grade Level: 3

Examples: At first, Sally had some marbles.
Then, she lost 3 of them.
Then, she found 2 marbles.
After that, she still had 8 marbles left.
How many did she have at first? ____________

Sam has to move 10 boxes.
He can carry 3 boxes each trip.
How many trips will he need to make? ____________

W4 Extraneous Information

Abstract: Solve word problems in which extraneous information is given. Once the relevant information is selected, the solutions are simple one-step problems involving small whole numbers.

Grade Level: 4

Examples: A belt costs $4.
A shirt costs $5.
A hat costs $10.
How much more does a hat cost than a belt? ________

Peter has $10.
He needs 4 pounds of candy.
Candy is $2 per pound.
He is buying candy for 6 people.
How much will the candy cost altogether? ________

W5 Fractional Sugar

Abstract: Solve word problems each of which start with   cups of sugar. The one-step solutions all require simple computions (+, -, x or -) with fractions or mixed numbers.

Grade Level: 4
Examples: Tina has $4\frac{1}{2}$ cups.
She buys $5\frac{1}{2}$ more cups.
How much sugar will she have then? ________

Kari has $4\frac{1}{2}$ cups.
She gives away half of it.
How many cups of sugar will she have left? ________

W6 Three Stage Word Problems

Abstract: Solve word problems in which the solution requires three operations. The problem is stated in 3 to 5 short sentences and the numbers given in the problems are relatively small.

Grade Level: 5, 6

Examples: Shirts cost \$10 each and ties cost \$5 each.
Altogether Joe spent \$35 for shirts and ties.
He bought 2 shirts.
How many ties did he buy? ______________________

Bill loads 6 boxes in 2 hours.
John loads 4 boxes in 2 hours.
Together, how many boxes do they load in 6 hours? _____

W7 Decimal Gas

Abstract: Solve word problems each of which start with 6.5 gallons of gas. The one-step solutions all require simple computations (+, -, x, or -) with decimals.

Grade Level: 5

Examples: Peter has 6.5 gallons.
Then he spills 1.2 gallons.
How much gas will he have left? ________________

Ron has 6.5 gallons.
Next week he will use ten times this much.
How much gas will he use next week? _____________

W8 Novel Word Problems

Abstract: Solve word problems which are novel for sixth graders in one or two of the following ways: involves fractions or decimals, requires more-than-three-stage solution, answer choices are approximate, requires solving for two unknowns, requires the use of data which is common knowledge but not given in the problem. Response format is multiple choice.

Grade Level: 6

Examples:

Ellen saw pepper plants on sale at 3 plants for 40¢.
She bought 12 plants.
She usually bought 3 plants for 50¢.
How much did she save?

20¢ 40¢ 48¢ \$1.60 \$2.00

George's father gives him 2¢ for every hour he spends in school.
About how much would he have given George for the month of October?

\$.50 \$1.00 \$3.00 \$6.00 \$10.00

Category A: Algebra

A1 Algebraic Symbols

Abstract: Given the numerical value of a letter (or letters) produce the numerical value of an expression involving that letter (those letters). In written instructions, two sample items are worked out and implied multiplication (e.g. in 3bc or in d^4) is explained. This scale follows A2 in the test booklet.

Grade Level: 6

Examples: If g = 4 and h = 3 then 5gh = ____

If p = 2 then p^5 = ____

A2 Solving Equations

Abstract: Given simple equations in one unknown, solve for the unknown. Three sample items are worked collectively, including one with a parenthesis.

Grade Level: 6

Examples: (7 x h) + 1 = 15, so h = ____

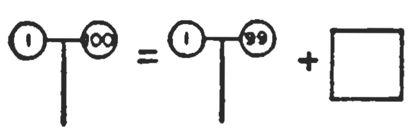
(n + 1) ÷ 3 = 6, so n = ____

A3 Summation Operator

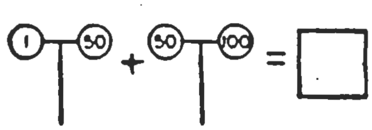
Abstract: Given an open sentence involving one or more summations of consecutive integers, select the answer that completes the sentence. A symbol for such summations (○T○) is introduced and explained (②T⑥ = 2+3+4+5+6) and two items are worked collectively.

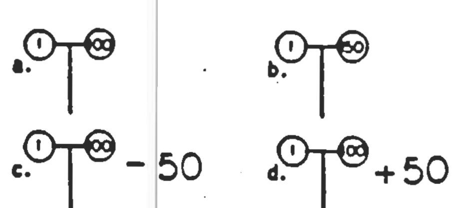
Grade Level: 6

Examples:



a. 1 b. 99 c. 100 d. 199





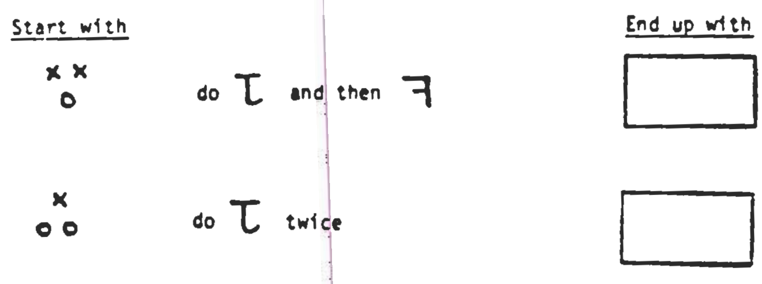
A4 Transformations

Abstract: Given two different transformations (ꟻ which turns a design clockwise by 90^{o} and τ which reverses the number of symbols at the top and bottom of a design), the scale consists of two different sections: requiring the application of either ꟻ or τ to a design, requiring several applications of ꟻ and/or τ to a design. Several sample items are worked collectively in each section.

Grade Level: 6

Examples: Section I:

Section II:



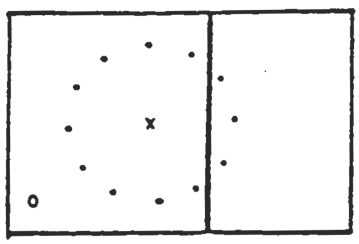
Category G: Geometry

G1 Geometric Loci

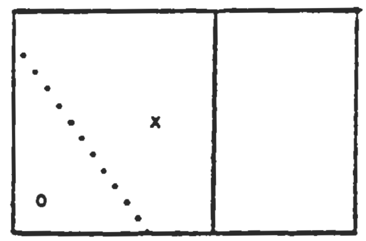
Abstract: Determine which picture is described by a given statement, where several pictures are given, each of which has identically placed elements (an 'x,' an 'o' and a line) but a different set of dots, determine which picture a given statement describes. First statement is read by the tester.

Grade Level: 4

Examples: A



E



In which picture are all the dots the same distance from the x? A B C D

In which picture is each dot just as close to x as to o? A B E F

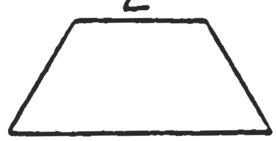
G2 Geometric Congruencies

Abstract: Given a regular geometric shape dividethe shape into a certain number of congruent parts. The word "congruent"is not used. Three correct and three incorrect solutions to a sample problem are examined collectively.

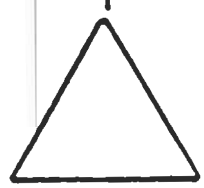
Grade Level: 5

Examples:

2



4



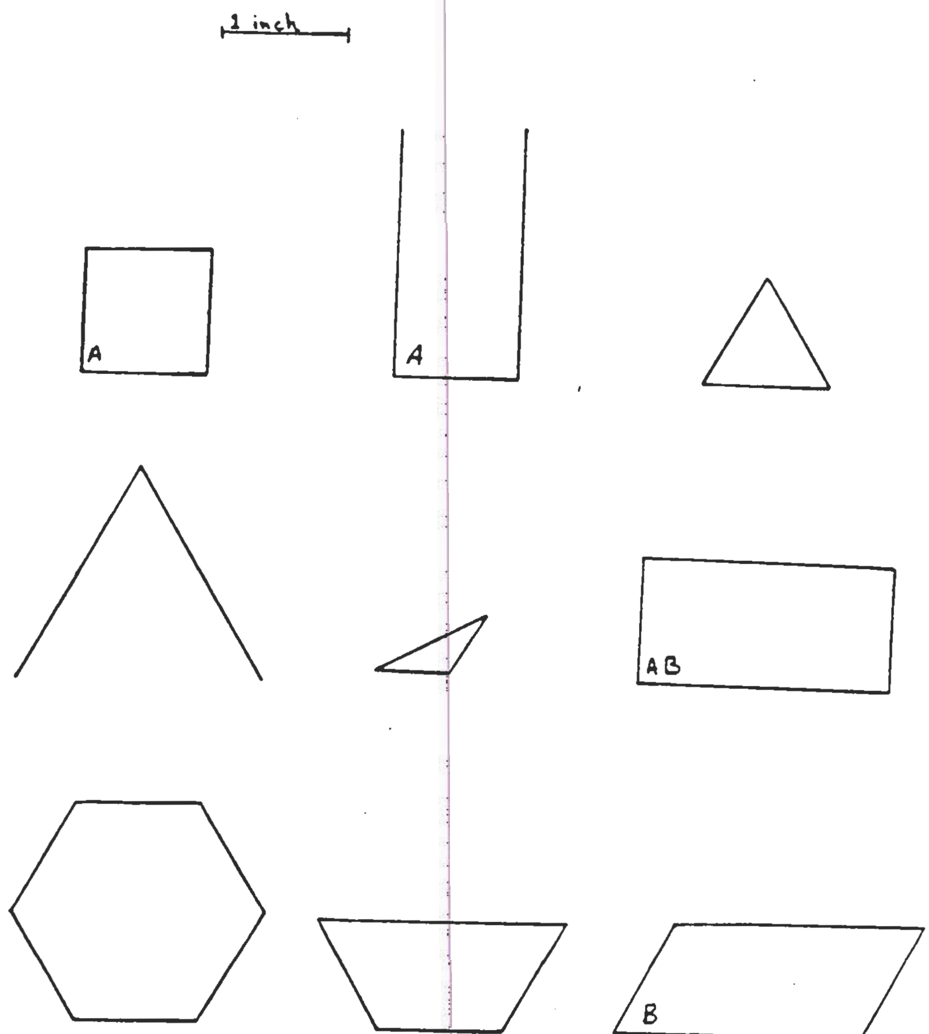
Category G: Geometry

G3 Geometric Categories

Abstract: Given nine different geometric figures, identify a set of 2 to 7 figures that are alike in some way, describe the distinguishing characteristic and label the figures accordingly. Go through this process as many times as possible. Two examples are worked collectively.

Grade Level: 6

Examples:



Sample 1 All the figures with "A" have square angles.

Sample 2 All the figures with "B" have only two sides that are one inch long.

All the figures with "C" ____________________

All the figures with "D" ____________________

etc.

Category L: Logic

L1 Logical Identification

Abstract: Given a specific set of individuals, a specific set of characteristics, the fact that each individual has a distinct combination of characteristics, and several facts about some of the characteristics of some of the individuals, identify the characteristics of each individual. A smaller sample problem is worked collectively.

Grade Level: 6

Example :

These are the 4 boys: Bill Tom Ed Pete

These are the 4 leagues: indoor soccer outdoor soccer indoor hockey outdoor hockey

These are the facts: Each boy plays in a different league.
Bill plays indoors
Tom <u>doesn't</u> play hockey
Ed <u>doesn't</u> play outdoors and he <u>doesn't</u> play soccer.

What league does each boy play in? (Circle your answers.)

Bill:	indoor soccer	outdoor soccer	indoor hockey	outdoor hockey
Tom:	indoor soccer	outdoor soccer	indoor hockey	outdoor hockey
Ed:	indoor soccer	outdoor soccer	indoor hockey	outdoor hockey
Pete:	indoor soccer	outdoor soccer	indoor hockey	outdoor hockey

L2 Making Sentences False

Abstract: Given a picture of a set of blocks and a true sentence about them, make the sentence false by changing the blocks. In the first two items, three suggested changes in the blocks are given and the student need only <u>mark</u> which ones would falsify the sentence. In the last three items, the student must <u>write</u> a change in the blocks. An item of the first type is worked collectively.

Grade Level: 6

Examples:

JOE'S BLOCKS

"There are triangles above the line and squares below the line."

a. Take away the triangles.
b. Take away the squares below the line.
c. Add squares above the line.

"Triangles go above the line or circles go below the line."
(You write what Joe could do to make the sentence false.)

____

Category O: Organization of Data

01 Graphing Weight

Abstract: Given a graph in which weight (axis labelled at 10 pound increments for each 5 graph units) is plotted against age (axis labelled at 2 year increments for each 2 graph units), determine age per given weights and vice versa. One sample item is worked collectively.

Grade Level: 5

Examples: How much did Bill weigh at 4 1/2 years of age? _____

How old was Bill when he reached 90 pounds? ________

02 Interpolating from a Table

Abstract: Given a table of prices for pipe of 4 different widths and 4 different lengths, interpolate or extrapolate to obtain the price on a pipe of given dimensions: at least one of which is not shown in the table. Two sample items are worked collectively.

Grade Level: 6

Examples:

Cost of Pipe

		Length			
		100'	300'	600'	1,000'
Width	4"	$50	$150	$300	$500
	8"	$70	$210	$420	$700
	12"	$90	$270	$540	$900
	16"	$110	$330	$660	$1100

HOW MUCH DOES IT COST TO BUY PIPE WHICH IS:

6" x 100' = $______

20" x 1000' = ______

Category P: Probability

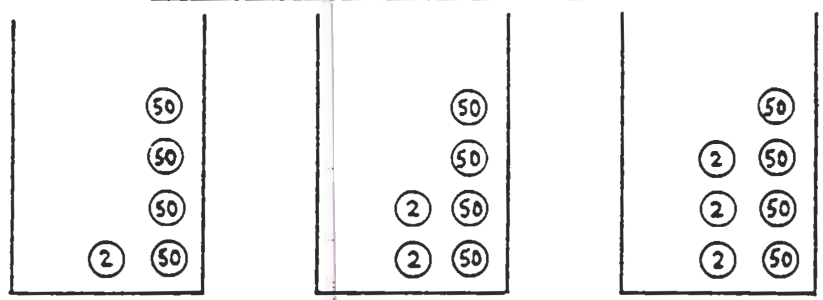
P1 Choosing the Best Box

Abstract: Given three boxes containing different combinations of 1, 2 and 50-cent"balls", determine from which box it would be best to make a blind draw.

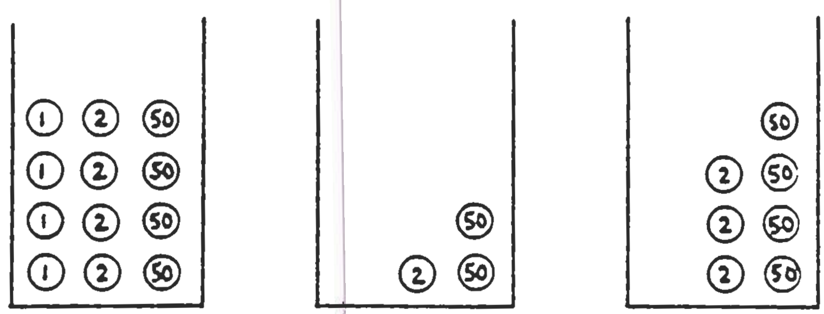
Grade Level: 5, 6

Examples:

WHICH BOX WOULD YOU CHOOSE?



WHICH BOX WOULD YOU CHOOSE?



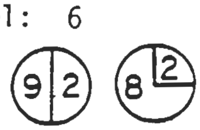
P2 Dependent Outcomes

Abstract: Given two (or three) spinners and an amount (10) to be achieved or exceeded to win, select (from five standard choices) how often a player would win. Collectively it is shown how a player could win or could lose with a specific set of spinners.

Grade Level: 6

two forms, approximately 3.5 minutes.

Examples:



never	less than half the time	half the time	more than half the time	always

never	less than half the time	half the time	more than half the time	always

Appendix H

Summary of External Review of CSMP Materials, 1974
Dr. Shirley Hill, University of Missouri at Kansas City

If this sample of mathematicians' opinions is in any way representative then I cannot help but comment that the mathematical community is a long way from any consensus concerning what mathematics is important and what should be taught. (The possibility that there could be agreement on how it is taught is so remote as to exclude hope.) The difficulty of summarizing the five reports is exacerbated by the apparent fact that the reviewers' perceptions of their roles and the purpose of their evaluations differed greatly. The reports seem to be addressed to different audiences and vary widely in degree of specificity, in focus and in the framework of time and vision (farsighted, shortsighted, nearsighted, hindsight, foresight, the "now," the future, the past, etc.) within which any value judgment is imbedded. Thus I strongly urge any reader of this summary at least to skim each of the individual reports.

The overall impression of the materials was favorable; three reviewers expressed quite favorable evaluations directly, the reaction of another was mixed, and the impression of the fifth cannot be said to be favorable, though it was not explicitly negative.

One point of general agreement in the reports was on the soundness of the mathematical content. The material is seen to be mathematically sound without any egregious technical or conceptual errors. There were differences of opinion concerning matters of preference and taste in the development of the mathematical ideas.

It was at least implicit in every report that it was impossible to separate completely in an evaluation of this kind, matters of mathematics and matters of pedagogy. Certainly most of the differences in preference concerning the way the mathematics was presented had little to do with mathematical soundness but rather related to questions of learning, development, concept formation and the like. Many of these are empirical questions. I think that it is fair to say that most of the very specific comments and specific criticisms concern psychological and pedagogical issues.

An example of a curricular element which is a mix of mathematical and pedagogical issues is the use of the minicomputer. This is the single point of complete agreement among all reports. There is too much reliance on the minicomputer. Three reviewers vehemently opposed its use as an aid altogether; the other two seriously question its value in light of the very great investment of time. (Both of these reviewers agree that the effectiveness of the device with respect to computational skills is an empirical question) All five reviewers are dubious to very negative on the minicomputer's mixture of a binary and decimal base.

Are the materials innovative, current, timely? Comments ranged from "it is more of the same" to "the material is refreshingly full of new ideas." The majority were of the opinion that the materials were timely and current and in many instances excitingly new. One reviewer found much new material of which he could approve but too much "old" material from the era of "new math." One found some "good sections" but little mathematics and much "obsessive ritual."

The question of relevance is tricky, as everyone knows. "Relevance" has no meaning except in the context of one's objectives, values, indeed one's philosophy. I can only infer that there are differences among the reviewers in the philosophical basis of their views of mathematics - what it is and what it does. Thus it is impossible to summarize the comments relating to perceived relevance of the material. There simply is no constant base for the opinions expressed. Certainly I can ascertain no consistent set of criteria for relevance.

Let me offer some examples of these differences. One reviewer sees the authors of the materials as "oriented to pure mathematics" and working in the "format of the past twenty years," while another feels that the extent of "student's participation" and spontaneity is encouraging, apparently viewing the materials as having moved beyond "the precocious discussions of systems and structure" of the past decade.

One reviewer sees too much carryover of material from the "new math" (I defy anyone to provide a clear-cut definition of that unfortunate term) and views such material as faddish while another, believing in the need for more historical perspective in distinguishing trends from fundamentals, compliments the authors on maintaining a balanced program that is timely and relevant today without discarding all the achievements of recent years.

The majority of reviewers saw the materials as modern, relevant to today's trends in mathematics and its applications with potential for developing competent future mathematical users and problem-solvers.

I will end by mentioning some specific things mentioned in more than one review. All reviewers praised the inclusion of extensive study of probability. Most liked the material on relations and functions, on graphing and arrow diagrams, on combinatorics.

Three reviewers specifically pointed to the "spiral" development and saw this as a positive feature. These three reviewers also believed the balance between concepts and applications was good. Two specifically pointed out that the activities stimulate active problem-solving and logical reasoning.

Most reviewers were critical of the material on sets, set operations, and Venn diagrams. Two opposed the material on the properties of arithmetic operations. Two felt there should be more reliance on manipulative, physical materials.

As mentioned earlier, all reviewers were negative (in varying degrees) about the minicomputer.