EDUCATION AND TREATMENT OF CHILDREN Vol. 23, No. 2, MAY 2000

Effective Mathematics Instruction The Importance of Curriculum

Donald B. Crawford
Western Washington University
Bellingham, Washington

Vicki E. Snider
University of Wisconsin--Eau Claire
Eau Claire, Wisconsin

Abstract

A two-year study was conducted in two fourth grade classrooms to evaluate the effectiveness of two mathematics curricula. During the first year, Teacher A used a Direct Instruction program, *Connecting Math Concepts* (CMC), and Teacher B used a traditional math basal textbook published by Scott, Foresman. During the first year, the CMC group scored significantly higher on the computation subtest of the National Achievement Test and on curriculum-based tests constructed from the CMC and basal programs as well as on a multiplication facts test. The next year Teacher B also used the CMC program, and achievement in Teacher B's classroom was significantly greater than the previous year, on the curriculum-based tests and the multiplication facts test. This suggests that the curriculum was the critical variable responsible for higher student achievement. Implications for textbook adoption and selection are discussed.

★ ★ ★

What is the best route to improved mathematics outcomes? In response to criticisms of math texts, some have suggested that textbooks be set aside in favor of more life-like, problem-solving mathematics instruction. Despite attempts to deemphasize the role of textbooks through projects and interdisciplinary thematic instruction, textbooks still "dominate instruction in elementary and secondary schools" (Farr, Tulley, & Powell, 1987, p. 59). Even though good teachers provide instructional opportunities that go beyond the textbook, 75% to 90% of classroom instruction is organized around textbooks (Tyson & Woodward, 1989; Woodward & Elliott, 1990). With this fact in mind, it seems wise to explore the possibilities of improving textbooks rather than abandoning them (Grossen & Carnine, 1996). Osborn, Jones, and Stein (1985) argued that "improving textbook programs used in American schools is an essential step toward improving American schooling" (p. 10).

Improved textbooks, if available, would need to be selected by state or

Address: Donald D. Crawford, Department of Special Education, Miller Hall 318, Western Washington University, Bellingham, WA 98225-9090.

Pages 122-142

local textbook adoption committees However, current textbook adoption processes are not research based. According to Tyson-Bernstein (1988), selections are more likely to be guided by political and economic factors than by qualities that are known to benefit students. Moreover, adoption committees are often poorly trained for the task of analyzing textbooks. Evaluation of textbooks is not as objective as regulations and policy statements suggest and it is based primarily on professional judgment rather than any objective rating system (Five & Cook, 1994) Pedagogy and educational research are seldom mentioned as factors that influence decisions (Courtland et al., 1983; Powell, 1985).

Even if adoption committee members were better trained, they would find that (a) field-test data and program evaluation data are infrequently reported and (b) textbooks are very similar. Texbook similarity has occurred because 22 states, most notably California and Texas, have statewide adoption procedures that require centralized textbook adoption. Because adoption by large states is critical for profits, publishers tailor their textbooks to meet the requirements of these states. The result is that the textbooks published by different companies are "very careful to be comprehensive in their coverage of topics, but they are seemingly indifferent to the conceptual coherence of the content and the pedagogical effectiveness of activities that are recommended therein" (Carnine, 1991, p 263).

Table 1
Comparisons Between Math Curricula

	Scott Foresman	**Connecting Math Concepts**
Organization	Spiraling; one objective for each lesson	Strands; multiple objectives for each lesson
Strategies	Implicit; suggestions given to the teacher	Explicit; presentations are scripted for the teacher
Problem Solving	General strategies; a variety of problems in each cliapter	Specific strategies for comparison, classification, fractions, and multistep problems
Mastery	Spiral design teaches for exposure; chapter tests; cumulative tests for chapters 1-6, 1-9, and 1-12; extra problems at back of book	Mastery emphasized; review of previous problem types in each lesson; cumulative tests every 10 lessons; criteria for "passing" given for each test and remedies provided
Facts instruction	Chapter 1 (+ and -); Chapter 5 (x); Chapter 8 (.); no mastery activities	Lessons 1-90 (x and .); mastery activities and timed tests
Field testing	None described	Extensive field testing and revision

In the current study, two curricula were compared in preparation for an adoption decision in a small school district. The two curricula that were experimentally compared are *Invitation to Mathematics* (1988) published by Scott Foresman and *Connecting Math Concepts.* (Engelmann, Engelmann, & Carnine, 1993). The two differ in important ways as summarized in Table 1.

Scott Foresman

Organization

The most prominent feature of the Scott Foresman (SF) mathematics curricula is its spiral design. In a spiral design, texthooks are organized into 10-20 chapter or unit topics and each topic is revisited each year (Stein, Silbert & Carnine, 1997). The intent of the spiral design is that topics will be treated with increasing depth and sophistication each year, building on the previous year's learning. However, as Miller and Mercer explain, "... in reality the result seems to be superficial coverage of many different skills" (1997, p. 51). The International Mathematics and Science Study (TIMSS) Curriculum and Textbook Analysis project reported that their analysis of 628 texthooks from roughly 50 countries showed that American math textbooks covered more topics than almost any other country in the world, but little was covered in depth (Schmidt, McKnight, & Raizen, 1996). Referring to the spiral designed curriculum, the National Council of Teachers of Mathematics (NCTM, 1989) noted the need to change the "repetition of topics, approach, and level of presentation in grade after grade" (p. 66). Porter (1989) found over 70% of the math concepts in the elementary curriculum received cursory instruction, defined as less than 30 minutes instructional time during; the entire school year.

As an example of the spiraling design, the concept that fractions are part of a whole occupies only one lesson in the Scott, Foresman (SF) fourth grade text. Conducting that lesson will give students less than 30 minutes instruction during all of 4th grade, which is insufficient for many below average to average students to understand the concept. A curriculum that spirals may result in teaching for exposure. "Skill mastery is unlikely, because new skills are introduced too quickly in an attempt to 'get through the book'" (Miller & Mercer, 1997, p. 51).

Another problem is that the rate at which new concepts are introduced is often either too fast or too slow. One objective is stated for each lesson in SF and each lesson has a new objective. For example, Objective 119 has to do with addition of fractions with like denominators and Objective 120 with addition of fractions with unlike denominators. Assuming the daily math period is the same length of time, there will be too much time for Objective 119 leading to wasted instructional time, and there will not be enough time to introduce, let alone master, Objective 120.

Strategies

The SF curricula relies upon discovery learning rather than explicit presentation of specific strategies for solving problems. In discovery leaming, the teacher sets up a situation in which students are to discover important concepts through the use of inductive reasoning. The SF teacher's manual provides suggestions for using "concrete materials" to help students understand important concepts and poses questions or situations to motivate students. Some teaching suggestions are offered such as "write the example on the board," "point out...," "explain..," and "encourage students to discuss."

Problem Solving

Scott Foresman includes problem-solving activities in each chapter. Throughout the text, they remind students to use a general problem solving strategy that is prompted by a penguin holding a sign. The general problem-solving strategies include the following: make a table, find a pattern, use physical modes, use logical reasoning, work backward, list all possibilities, try and check, draw a diagram, and make a graph.

Mastery

The teacher's guide for the SF curricula does not include any discussion of mastery. The text is organized so that a new concept is introduced daily, so some students may not receive enough practice to develop mastery. In addition, there is minimal review after the unit in which the skills are introduced. For example, learning how to tell the time from a clock is presented on pages 90-91 of the fourth grade SF text, but is never seen again in the remaining 256 pages of the text. Fractions are presented in Chapter 11 of the 4th grade, but students do not see any fractions again until Chapter 10 of the 5th grade textbook.

The fourth grade SF text has a test at the end of each chapter, occasional "maintenance" problems and cumulative tests for Chapters 1-6, 1-9, and 1-12. There are also extra problems at the back of book for both calculation and problem-solving However, the cumulative review involves a limited number of problems, it is not integrated and it occurs infrequently. Most lessons do include an error analysis, which alerts teachers to common errors and makes suggestions for extra practice.

Facts Instruction

Chapter 1 of the SF text includes a review of single digit addition and subtraction. Chapter 5 reviews multiplication and Chapter 8 introduces division. There are no mastery activities or timed tests of math facts.

Field Testing

No field testing was described in the SF teacher's manual or in the literature.

Connecting Math Concepts

Organization

The curriculum design of the Direct Instruction program, *Connecting Math Concepts* (CMC) is very different from SF. The fact that each lesson is organized around multiple concepts and skills, rather than around a single unit, is the most unusual aspect of the design of CMC. Each concept/skill is addressed for only 5- to 10-minutes in any given day's lesson, but is revisited day-after-day for many lessons. Organizing lessons so that concepts/skills are revisited for a few minutes a day over many days is referred to as a "strand" organization.

Many important curriculum goals are made easier by organizing lessons into strands, such as sequencing of dozens of preskills, cumulatively introducing skills, and treating topics in depth. The presentation of key concepts in strands which run through several lessons allow the concepts to be arranged in a logical scope and sequence. Thus preskills can be taught prior to being integrated into more complex mathematical concepts. For example, before students are taught to find equivalent fractions (e g., 1/3 - 3/9) by multiplying by a fraction of one (e.g, x 3/3), they learn the necessary multiplication facts and that multiplying by a fraction equal to one doesn't change the value of a number and how to write fractions equal to one.

With a strand design the variations in amount of time needed for learning each concept are easily accommodated by adjusting the number of minutes and the number of consecutive days spent teaching it. The strand design thus accommodates the problems associated with variability in the time needed to learn each skill/concept (and the related issue of the rate of introduction of concepts). (See Stein, Carnine, & Dixon, 1998) or (Carnine, Jones, & Dixon, 1994) for a more thorough discussion of the design features of CMC).

Strategies

CMC uses explicit strategies to teach both basic operations and problem solving. Within a strand the amount of structure is gradually decreased each day moving from the initial teaching presentation through guided practice to independent practice. Because each lesson contains many strands, a balance between new learning and practice is maintained.

The teacher's manual provides specific teaching procedures including wording and error correction procedures. The purpose of the teaching scripts is to assure that the explicit strategies are presented in clear and unambiguous language of instruction. This represents a radical departure from traditional texts and can be daunting for teachers who are used to a more freewheeling delivery style.

Problem solving

Students work on solving comparison and classification problems using addition and subtraction, multistep problems and problems involving fractions including probability. Strategies are systematically presented in a word problems strand throughout the year. Explicit instruction is provided to help students discriminate among problem types so that they know when to apply each strategy.

Mastery

An important advantage of strand organization is that it enables gradual mastery of concepts by repeating and extending information over many lessons rather than teaching by a single exposure. Strands allow distributed review of skill in which only a few problems are presented daily over a long period of time. This distribution over time allows systematic, brief review of concepts until they are integrated with other more complex mathematical procedures. Such distributed practice allows students to become both accurate and rapid in their responses. Distributing practice across several days facilitates mastery better than massing practice in one day's lesson (Dempster, 1991), and it is easy to schedule when lessons are designed in strands.

Mastery is critical to success in CMC because the strand design requires that students use everything that has been taught. If students are weak on a particular skill, they will most certainly have trouble later in the program when that skill becomes a component in a more complex operation. The teacher's guide provides criteria for "passing" each test and makes specific suggestions for remediation. In addition, teachers are given specific suggestions for correcting different types of errors and specific procedures for "firming" students on concepts within each lesson.

Facts Instruction

Addition and subtraction facts are assumed to have been mastered as they were practiced in previous levels. Mastery of multiplication and division facts is emphasized in lessons 1-90 including a variety of oral mastery activities and timed tests.

Field-Testing and Research

Uniquely, CMC has received extensive field-testing prior to being marketed (Engelmann, Englemann, & Carnine, 1993). For example, evaluations of the third grade text prior to publication indicated that low-income minority students scored two years above grade level after using CMC (Carnine & Engelmann, 1991). In another prepublication investigation, high performing third graders were better able to solve word problems and make connections between various math concepts than other high performing students (Carnine & Engelmann, 1991). Since its publication empirical data supporting its effectiveness continues to accumulate.

The CMC program and a basal math program were separately used at two schools in Camden, NJ with beginning first graders from predominantly educationally at-risk backgrounds. After two years, CMC-taught students scored significantly higher on math computation than the basal group on both the California Test of Basic Skills and the Metropolitan Achievement Tests (Brent & DiObilda, 1993).

In an educationally at-risk elementary school in Kalamazoo, MI, two third-grade classes, and a fifth-grade class participated in a pilot study using CMC. Students taught with CMC: (a) achieved average to above-average rates of progress in both math calculation and application on the Kaufman Tests of Education Achievement; (b) came close to the 50th percentile on the Iowa Tests of Basic Skills although other students displayed significant percentile declines between second and third grade; and (c) were more sophisticated at math problem solving (Vreeland, et a1., 1994). As a consequence, CMC was implemented in other math classes. At the end of the second implementation year, the outcomes were overwhelmingly positive (Vreeland, et al., 1994).

CMC was implemented in the first and fourth grades in eight elementary schools in a Pennsylvania school district, during the 1992-93 school year (Wellington, 1994), Teacher-designed posttests for the fourth grade showed significant differences between students educated with CMC compared to the "traditional basals," consequently CMC was adopted district-wide in grades 1-5. Although Wellington (1994) did not find significant differences in favor of the CMC group in the first grade, Tarver and Jung; (1995) found that first grade students using CMC outperformed students using *Math Their Way* and cognitively guided instruction (CGI) on the CTBS Math subtests. Furthermore, statistically significant differences increased after the second year of implementation when the CMC group scored more than one grade level above the CGI group on the CTBS. Furthermore, 20% of the CMC students scored at the ceiling of the test although none of the CGI students did so.

The study reported herein began as the result of two teachers' willingness to conduct a pilot study for one year to evaluate CMC, in which one used CMC (Teacher A) and the second (Teacher B) served as a control group by continuing to use the same Scott, Foresman textbook the two had been using for years. Then the second year, when Teacher B wanted to use CMC also, we had an opportunity to examine the effect of the curriculum independent of teacher effect by comparing with Teacher B's previous scores.

Conducted the year prior to the district's consideration of a new math basal curriculum, this study's purpose was to provide local empirical data useful for making a decision to adopt and to add to a growing body of small-scale program evaluations supporting the effectiveness of CMC. Within this framework, CMC was compared to a traditional basal, namely the Scott, Foresman (SF) text *Invitation to Mathematics* on several norm-referenced and curriculum-based measures. The addition of the second year's data from Teacher B enabled us to make the same comparisons independent of the effect of the teacher. To our knowledge the question of the effectiveness of CMC has not been previously examined by comparing the achievement of a teacher's own classes before and after implementation.

Methodology

Materials

The instructional materials used were *Connecting Math Concepts* (Engelmann, et al, 1993) Level D and *Invitation to Mathematics* (1988) published by Scott, Foresman. Both curricula were designed for 4th grade students. The CMC curriculum included a teacher's guide, teacher presentation books A and B, an answer key, and student textbooks and workbooks SF included a teacher's guide and a student textbook. The content of both curricula contained considerable overlap, but was not identical. The SF text included chapters on addition and subtraction facts, numbers and place value, addition and subtraction, measurement, multiplication facts, multiplication, geometry, division facts, division, decimals, fractions, and graphing. Each chapter in the SF text interspersed a few activities on using problem solving strategies. CMC included strands on multiplication and division facts, calculator skills, whole number operations, mental arithmetic, column multiplication, column subtraction, division, equations and relationships, place value, fractions, ratios and proportions, number families, word problems, geometry, functions, and probability. Despite the differences in content and organization, both programs covered math concepts generally considered to be important in 4th grade--addition and subtraction of multi-digit numbers, multiplication and division facts and procedures, fractions, and problem solving with whole numbers.

Other subjects and activities sometimes interfered with time set aside for math instruction, and some lessons took longer than the one lesson per day typically expected. As a result, neither group had time to complete their respective curricula. During year 1, the CMC group completed 90 out of 120 lessons, and the SF group completed 10 out of 12 chapters. During year 2, Teacher A completed 105 lessons in the CMC curriculum and Teacher B completed 95 lessons.

Measures

Four measures were used, with the same versions used for both pretest and posttest. The *National Achievement Test* (1989) (NAT) is a timed standardized test battery designed to be administered to groups of students. The math section consists of three subtests--computation, concepts, and problem solving, with concepts and problem solving combined for scoring. Its norms are based on a stratified random sample of 150, 000 kindergarten to twelfth-grade students in public schools in five geographical areas. Test-retest reliability on the mathematics subtests and total for Level F were all reported as r = .90 or higher (Wick, 1990).

A cumulative curriculum-based measure was drawn from each of the two curricula. The first test, (SF test), which was published by Scott, Foresman to go along with the *Invitation to Mathematics* text, was the complete Cumulative Test for Chapters 1-12 and was intended to be comprehensive as well as cumulative. The SF test consisted of 22 multiple-choice items (four choices) which assessed the range of concepts presented in the 4th grade SF textbook. For the CMC measure the first author designed a test that consisted of 55 production items for which students computed answers to problems, including both computational and word problems. The CMC test was comprehensive as well as cumulative; problems were examples of the entire range of problems found in the last quarter of the CMC program. Problems were chosen from the last quarter of the program because the various preskills taught in the early part of the program are integrated in problem types seen in the last quarter of the program. Students did not use calculators on any of the tests.

Fluency in recall of basic multiplication facts is an essential objective for fourth grade and was assessed by an experimenter-designed test. The three-minute, timed-test consisted of 72 simple multiplication facts, and students completed as many as possible within the time limit. Students took the test three times on three different days and their mean score was recorded.

Participants

All the students in the 4th grade of a school in a small community in Wisconsin were randomly assigned to one of two fourth grade classrooms prior to the study. Data were collected on all who had permission to participate. In year 1, 23 students were in each class (two students in each class did not wish to participate, so their data were excluded from the analysis). Percentile scores on the NAT indicated the classes began the year slightly below average in math skills with the CMC class in the 43rd percentile and the SF class in the 44th.

In year 2, there were 19 students in each 4th grade class. Both classes were heterogeneous and included the full range of abilities including learning disabled and gifted students. The attrition rate was low. In year 1, one CMC student left in midyear and returned later in the year. In year 2, a total of four of Teacher B's students moved out and four moved in. Scores for these students were not included in the data analysis.

Teachers

Teacher A, a female with 14 years of experience had taught 4th grade in the school for the previous three years. Teacher B, a male with 11 years of experience, had taught 4th grade in the school for his entire career. Both teachers were considered by their peers and administrators to be caring and competent.

Teacher A taught the CMC curriculum during Years 1 and 2. She had taught from SF during the previous years and had no previous experience with CMC or any other Direct Instruction programs. She received 4 hours of training at a workshop in August and about three hours of additional training from the experimenters. Teacher B had 11 years of experience with the SF text and taught from it during year 1. He received minimal training before teaching CMC during year 2. He attended a three-hour inservice training session and observed two demonstration CMC lessons.

Procedures

The second author informally and intermittently observed both classrooms a few times during both years. No observational data more formal than anecdotal notes were collected. Because this experiment was designed as a test of curriculum rather than a test of fidelity to a given set of teaching procedures we did not require the teachers to teach in any specific way.

In year 1, each teacher spent approximately 45 minutes per day engaged in math instruction, however his or her presentation differed considerably. Teacher A used the scripted presentation in the CMC teacher presentation book. She frequently asked questions to which the whole class responded, but she did not use a signal to elicit unison responding. If she got a weak response she would ask the question again to part of the class (e.g., to one row or to all the girls) or ask individuals to raise their hands if they knew the answer. There were high levels of teacher-pupil interaction, but not every student was academically engaged. Generally, one lesson was covered per day and the first 10 minutes were set aside to correct the previous day's homework. Then a structured, teacher-guided presentation followed, during which the students responded orally or by writing answers to the teacher's questions. Student answers received immediate feedback and errors were corrected immediately. If there was time, students began their homework during the remaining minutes.

During year 1 Teacher B's math period was divided into three 15-minute parts. First, students checked their homework as B gave the answers. Then students told B their scores, which he recorded. Second, B lectured or demonstrated a concept, and some students volunteered to answer questions from time-to-time. The teacher presentation was extemporaneous and included explanations, demonstrations, and references to text objectives. Third, students were assigned textbook problems and given time for independent work.

During year 2, when both Teacher A and B implemented CMC, their presentations became more similar, but differences remained. Teacher A, who sought additional training in direct instruction presentation techniques during the summer, had a very polished presentation. She now signaled to initiate unison responding, proceeded at a brisk pace and engaged all students during the entire lesson. Her demeanor was dynamic and enthusiastic. Teacher B followed the script, but deviated somewhat from the CMC delivery procedures. He added his own explanations, asked questions of the whole class, but did not require all students to demonstrate their understanding by responding overtly. He continued his practice of using the first part of the period to go over homework.

Both teachers had emphasized mastery of multiplication facts for several years. They continued to do so both years of the study and assessed progress through timed tests. Teacher A introduced the fact families (e.g., 2x1, 2x2, 2x3, etc.) in the same order in which they were introduced in CMC (9s, 3s, 4s, 7s, 6s, 8s, 5s, 1s, 2s). Students in A's class took a one-minute, 24 item timed test in one fact family and, when they passed, they went on to the next. Parents were encouraged to practice with their children at home, but little class time was allotted, outside of what was in the textbook, to practice facts. Teacher B did not follow the CMC order for introducing fact families. Instead the facts were given in numerical order (1s, 2s, 3s, etc.) and the facts were not broken down into smaller units for mastery. Students took a one-minute test of all facts on an irregular basis.

During year 1, the teachers administered the multiplication facts and curriculum-based pretests to their own classes during the first week and a half of October. During year 2 these tests were given in late August. Both years the NAT was administered as a pretest during district-wide testing in early October. All posttests, including the NAT, were administered during the same two-week period in early May. Teachers A and B switched places in giving the curriculum-based posttests to each other's class to prevent any coaching. The same directions were read to both classes.

Results

There were two questions that we wanted to answer: (a) Did implementation of the CMC curriculum when taught by A improve achievement when compared to Teacher B in year one and, if so, on which measures' and (b) Did implementation of the CMC curriculum by Teacher B in year 2 improve achievement when compared to B's year 1 achievements and, if so, on which measures?

Comparison Between Teachers in Year 1

Based on analysis of variance, there were no significant pretest differences between students in the two curriculum groups on the computation, concepts and problem solving subtests of the NAT nor on the total test scores. Nor did any significant pretest differences show up on any of the curriculum-based measures (see Table 2).

Table 2

Pre and Posttest Raw Score Means on all Measures for Year 1

	Pretest		*Posttest*	
Groups	M	*SD*	*M*	*SD*
	National Achievement Test (NAT) Computation			
CMC Teacher A	26	8	35	4
SF Teacher B	26	8	29	7
	NAT Concepts & problem-solving			
CMC Teacher A	31	12	37	13
SF Teacher B	32	9	39	12
	NAT Total Test			
CMC Teacher A	56	19	72	16
SF Teacher B	58	15	69	18
	Connecting Math Concepts Test			
CMC Teacher A	6	6	41	8
SF Teacher B	7	3	15	8
	Scott Foresman Test			
CMC Teacher A	12	3	19	2
SF Teacher B	13	4	16	4
	Facts Fluency			
CMC Teacher A	15	7	66	7
SF Teacher B	22	11	48	12

Significant posttest differences in favor of the CMC group were found in mean raw scores on both of the curriculum-based tests as well as on the multiplication facts (see Table 1). The differences on the CMC curriculum posttest were quite large, (CMC = 41 or 74% correct), SF = 15 or 27% correct), with F (1, 40) = 104.4, p = 0.0001 (see Figure 1). Against typical expectations, the CMC group even outscored the SF group on the test based on the SF curriculum, (19 or 87% *vs.* 16 or 72%) with F (1, 40) = 11.2, p = 0.002. The CMC group also scored significantly higher on rapid recall of multiplication facts. Of 72 items, the mean correctly answered in 3 minutes for the CMC group was 66 compared to 48 for the SF group with F (1, 40) = 33.3, p = 0.0001 for the multiplication facts posttest.

Posttest comparisons on the computation subtest of the NAT, (Table 1), indicated a significant difference in favor of the CMC group, F (1,40) = 8.32, p = 0.006. On the other hand, neither the scores for the concepts and problem-solving portion of the NAT nor the total NAT showed any significant group differences. The total NAT scores put the CMC group at the 51st percentile and the SF group at the 46th percentile, but this difference was not statistically significant.

In summary, during year 1 the groups began the year with approximately equivalent scores on the pretests. The slight differences that were apparent favored the SF group. At posttest, statistically significant differences favored the CMC group on the NAT computation subtests, curriculum-based measures and multiplication fact fluency, but not on the NAT concepts and problem-solving subtest nor on the NAT total score.

Comparison Before and After Implementation for Teacher B

Our second question was, did achievement improve when the CMC curriculum was implemented with B in year 2 relative to the SF outcomes in year 1? Comparing the pretest scores of year 1 and year 2 for Teacher B's students, no significant differences were found on the subtest and total scores of the NAT nor on the pretest scores of the SF and the multiplication facts measures. However, significant pretest differences on the CMC measure occurred, E (1,34) =7.49, p= .009, with the pretest scores lower for year 2 than for year 1. In response to this potential confounding, a two-factor analysis of variance for repeated measures was used rather than doing simple posttest comparisons to evaluate the effect of the curriculum. Because identical forms of each of the tests were used during pre and post testing the raw scores are comparable.

For the NAT total score, the main effect for time, (pretest *vs.* posttest scores) was significant, F (1, 34) = 43.8, p = .0001, but the interaction between type of curriculum and time was not. The students learned each year, but the higher mean in the second year as against the first year on the NAT was not significant. The same pattern held for both the NAT

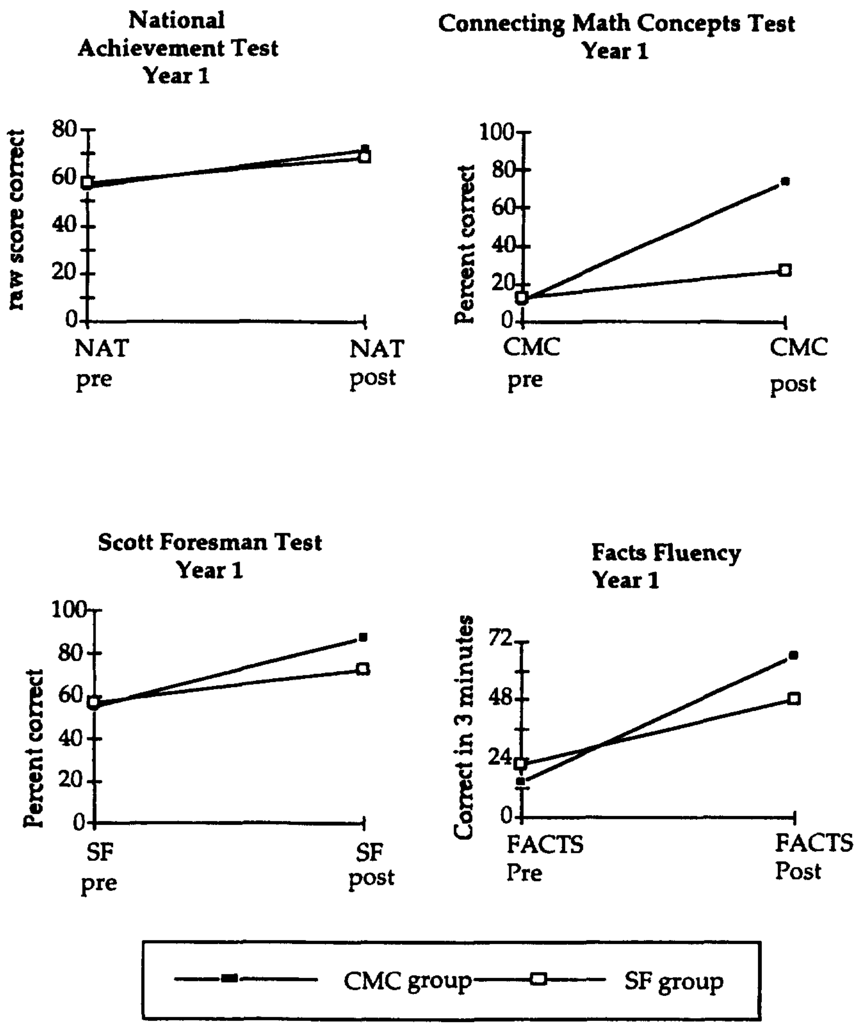

Figure 1. Year 1 Pre and posttest means on all achievement measures.

subtests, computation and problem solving.

On the other hand, for the CMC measure the curriculum x time interaction was highly reliable, $F(1, 34) = 40.5$, $p = .0001$, with the gains greater after the second year than the first year. This finding is not surprising because the CMC measure was based on the CMC curriculum. Before implementation of the CMC, the students taught by B using the SF text began the year with a mean score of 7 (12% correct) on the CMC measure and improved to 15 (27%), a gain of 15 percentage points. After implementation of CMC in year 2, B's students jumped from means scores of 4 to 33 (6% to 60% correct), a gain of 54 percentage points. Table 3 shows pre and posttest means and standard deviations for Teacher B's students on all measures for both years.

Table 3

Pre and Posttest Raw Score Means on all Measures for Teacher B Before and After Implementing CMC

	Pretest		*Posttest*	
Years	*M*	*SD*	*M*	*SD*
National Achievement Test (NAT) Computation				
Before	26	9	39	12
After	26	8	33	7
NAT Concepts & problem-solving				
Before	32	9	39	12
After	30	11	41	13
NAT Total Test				
Before	58	15	69	18
After	56	17	74	19
Connecting Math Concepts Test				
Before	7	3	15	8
After	4	3	33	14
Scott Foresman Test				
Before	13	4	16	4
After	10	4	18	3
Facts Fluency				
Before	22	11	48	12
After	17	10	54	12

For the SF measure, the curriculum x time interaction was also strong, $F\ (1, 34) = 15.82$, $p = .0004$, with the second year gains higher. In year one, B's students pretested at a mean score of 13 (57% correct) and improved to 16 at posttest (72% correct), a gain of 15 percentage points with the SF program. After implementation of CMC in year 2, B's students mean scores improved from 10 to 18, (47% to 82% correct), a gain of 35 percentage points. It should be remembered that the SF measure was not aligned with the CMC curriculum.

Finally, the interaction for the multiplication facts fluency test was also significant, $F\ (1, 34) = 10.5$, $p = .002$, with the second year gains greater. Before implementation of CMC, B's class average pretested at 22 multiplication facts in 3 minutes to and improved on a posttest to 48 facts, a gain of 26 facts due to the SF program. After implementation of CMC, B's students went from 17 multiplication facts in 3 minutes at pretest to 54 facts in 3 minutes, a gain of 37 facts. In summary, improvements on the two curriculum-based measures and the fact fluency measure were much greater after implementation of the Direct Instruction CMC curriculum in year 2 than before its implementation in year 1. The pre and posttest means (raw scores for the NAT, percent correct for curriculum-based measures, and facts completed in 3 minutes for facts fluency) for both years are shown in Figure 2. Crossover effects occurred for all four test measures, with a steeper slope or increased improvement in year 2.

Discussion

This study suggests that students learning from CMC perform better than students learning from SF when measured by curriculum-based tests. At the end of year 1, students in Teacher A's CMC classroom performed better on curriculum-based measures than students in the other, SF, fourth grade class. The two classes were not significantly different at pretest which suggests that the differences in achievement were not due to characteristics of the groups, but rather to implementation of a different curricula used by Teacher A. The data for students in Teacher B's classes, comparing year 1 using SF to year 2 using CMC as the math curriculum, suggests that the critical variable in student achievement was the curriculum, not the teacher. The significant curriculum x time interaction on both curriculum-based measures and the multiplication facts test favored CMC over SF. Although the higher achievement of students in Teacher A's group after year 1 could initially be attributed, all or in part to the teacher, the replication of higher student achievement in year 2 with Teacher B makes that interpretation less credible.

We believe this implementation of CMC was less than optimal because (a) students began the program in fourth grade rather than in first grade

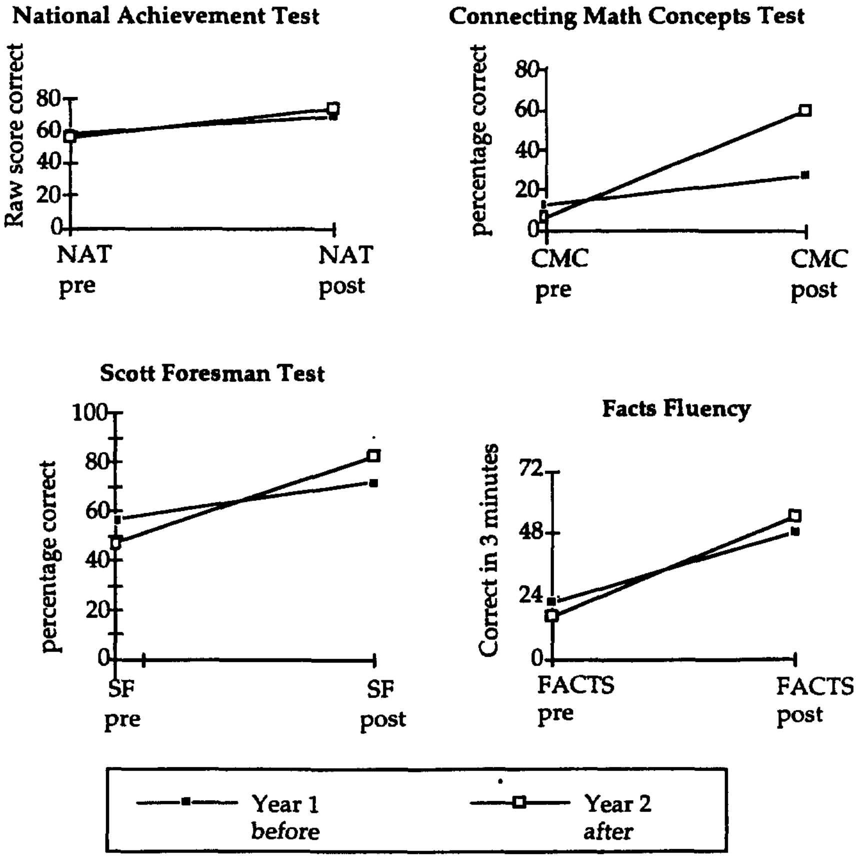

Figure 2. Teacher B Pretest and Posttest means before and after implementation of CMC.

and (b) students could not be placed in homogeneous instructional groups. A unique feature of the CMC program is that it's designed around integrated strands rather than in a spiraling fashion. Each concept is introduced, developed, extended, and systematically reviewed beginning in Level A and culminating in Level F (6th grade). This design sequence means that students who enter the program at the later levels may lack the necessary preskills developed in previous levels of CMC. This study with fourth graders indicated that even when students enter Level D, without the benefit of instruction at previous levels, they could reach higher levels of achievement in certain domains. However, more students could have reached mastery if instruction were begun in the primary grades.

Another drawback in this implementation had to do with heterogeneous ability levels of the groups. Heterogeneity was an issue for both curricula. However, the emphasis on mastery in CMC created a special challenge for teachers using CMC. To monitor progress CMC tests are given every ten lessons and mastery criteria for each skill tested are provided. Because of the integrated nature of the strands, students who do not master an early skill will have trouble later on. Unlike traditional basals, concepts do not "go away," forcing teachers to continue to reteach until all students master the skills. This emphasis on mastery created a challenge for teachers that was exacerbated in this case by the fact that students had not gone through the previous three levels of CMC.

Why didn't the improved learning for students using the CMC program demonstrated on the curriculum-based tests show up on the NAT? Our guess is that a more optimal implementation of CMC would have increased achievement in the CMC group, which may have shown up on the NAT. In general, the tighter focus of curriculum-based measures such as those used in this study makes them more sensitive to the effects of instruction than any published, norm-referenced test. Standardized tests have limited usefulness for program evaluation when the sample is small, as it was in this study (Carver, 1974; Marston, Fuchs, & Deno, 1985). Nevertheless, we included the NAT as a dependent measure because it is curriculum-neutral. The differences all favored the CMC program.

That no significant differences occurred either between teachers or across years on the NAT should be interpreted in the light of several other factors. One, the results do not indicate that the SF curriculum outperformed CMC, only that the NAT did not detect a difference between the groups, despite the differences found in the curriculum-based measures. Two, performance on published norm-referenced tests such as the NAT are more highly correlated to reading comprehension scores than with computation scores (Carver, 1974; Tindal & Marston, 1990). Three, the NAT concepts and problem solving items were not well-aligned with either curriculum. The types of problems on the NAT were complex, unique, non-algorithmic problems for which neither program could provide instruction. Performance on such problems has less to do with instruction than with raw ability. Four, significant differences on the calculation subtest of the NAT favored the CMC program during year 1 (see Snider and Crawford, 1996 for a detailed discussion of those results). Because less instructional time is devoted to computation skills after 4th grade, the strong calculation skills displayed by the CMC group would seem to be a worthy outcome. Five, although the NAT showed no differences in problem solving skills between curriculum groups or between program years, another source of data suggests otherwise. During year 1, on the eight word problems on the curriculum-based test, the CMC group outscored the SF group with an overall mean of 56% correct compared to 32%. An analysis of variance found this difference to be significant, with $F = 10.8$, $p = .002$.

Commonly teachers do not expect the textbook to be a significant assist in teaching math facts to fluency. Teachers learn how to do that on their own. Anecdotally Teacher A told us that Teacher B's students usually "learned their facts better than my students." The fact that Teacher A's class, using CMC, outperformed Teacher B's class is notable. The fact that Teacher B's students, using CMC, significantly outperformed his previous years class is even more interesting. The CMC program works on multiplication facts beginning in lesson 1 and continuing through lesson 68. The multiplication facts are taught in the following order in Level 1:): 5s, 9s, 3s, 4s, 7s, 6s. (Fact teaching begins in Level C.) For each series of facts, students work from a number map that shows a unique pattern for that fact family. As each fact family is introduced, the students use these facts to solve other problems in their classwork and homework. Aside from the motivation and opportunity for students to learn math facts offered by the teachers, the systematic integration and practice of those facts in CMC probably helped the students achieve a higher level of mastery.

Both teachers reported anecdotally that the high-performing students seemed to respond most positively to the CMC curricula. One of Teacher A's highest performing students, when asked about the program, wrote, "I wish we'd have math books like this every year.... it's easier to leam in this book because they have that part of a page that explains and that's easier than just having to pick up on whatever. "

It may be somewhat counter-intuitive that an explicit, structured program would be well received by more able students. We often assume that more capable students benefit most from a less structured approach that gives them the freedom to discover and explore, whereas more didactic approaches ought to be reserved for low-performing students. It could be that high-performing students do well and respond well to highly-structured approaches when they are sufficiently challenging. These reports are interesting enough to bear further investigation after collection of objective data.

This study is a model of the kind of small research project that districts should undertake prior to spending thousands of dollars adopting a new textbook. The cost of implementing a new program on a small scale is minimal and the data are easy to collect. After seeing the data from year 1, Teacher B and two first and third grade teachers also volunteered to pilot the program. In the following year, the program was adopted district-wide. Too often in education, important curricular decisions that affect thousands of children are made because of ideology (Dixon & Carnine, 1994) or faddism (Slavin, 1990). This study provides a model for how educators can make data- based decisions.

The data from the second year provide evidence that curriculum is a critical factor in student achievement. These data show that this teacher was able to achieve substantially better student outcomes by changing the curriculum that he was using. This is not to say that teacher skill, dedication and compassion are not important; but rather that given these qualities, teachers can produce better educational outcomes if they also have access to "tools that work" (Carnine, 1992). This study adds to a growing body of research indicating that CMC is a powerful tool that can enable teachers to help students understand and apply mathematical concepts.

References

Brent, G., & DiObilda, N. (1Y93). Curriculum alignment versus direct instruction: Effects on stable and mobile urban children. *Journal of Educational* Research, 86(6), 333-338.

Carnine, D. (1991). Curricular interventions for teaching higher order thinking to all students: Introduction to the Special Series. *Journal of Learning Disabilities, 24*. 261- 269.

Carnine, D. (1992). Expanding the notion of teachers' rights: Access to tools that work. *Journal of Applied Behavior Analysis 25* 13-19

Carnine, D., & Engelmann, S. (1991). Making connections in third grade mathematics--Connecting Math Concepts. *ADI News 10(4)*,17-27.

Carnine, D., Jones, R. D., & Dixon, R. (1994). Mathematics educational tools for diverse learners. *School Psychologv Review*. 23. 406-427.

Carver, R. P. (1974). Two dimensions of tests: Psychometric and edumetric. *American Psychologist, 29*, 512-518.

Courtland, M. C., Parr, R., Harris, P., Tarr, J., & Treece, L. (1983). *A case study of the Indiana state reading textbook adoption process*. Bloomington, IN; Center for Reading and Language Studies.

Dempster, F. N. (1991). Synthesis of research on reviews and tests. *Educational Leadership. 48*(7), 71-76.

Dixon, R. & Carnine, D. (1994). Ideologies, practices, and their implications for special education. *Journal of Special Education. 28 (3)*. 356-67.

Engelmann, S., Engelmann, O., & Carnine, D. (1993). *Connecting Math Concepts level D*. Science Research Associates: Chicago, IL.

Farr, R., Tulley, M. A., & Powell, D. (1987). The evaluation and selection of readers. *Elementary School Journal, 87*(3), 267-282.

Fiore, T. A., & Cook, R. A. (1994). Adopting textbooks and other instructional materials: Policies and practices that address diverse learners. *Remedial and Special Education, 15*, 333-347.

Grossen, B. & Carnine, D. (1996) Considerate instruction helps students with disabilities achieve world class standards. *Teaching Exceptional Children, 28(4)* 77- 81. *Invitation to Mathematics*. (1988). Glenview, IL.: Scott, Foresman, & Co.

Marston, D., Fuchs, L., & Deno, S, L (1985) Measuring pupil progress: A comparison of standardized achievement tests and curriculum-based measures. *Diagnostique, 11, 77-90.*

Miller, S. P. & Mercer, C D. (1997). Educational aspects of mathematics disabilities *Journal of Learning Disabilities. 30* (1), 47-56.

National Council of Teachers of Mathematics (1989) *Curriculum and evaluation standards for school mathematics.* Reston, VA NCTM.

National Achievement Test. (1989). Chicago, IL: American Testronics.

Osborn, J., Jones, B. F., & Stein, M. L. (1985). The case for improving textbook programs: An issue of quality. *Educational Leadership. 42,* 9-16.

Porter, A. (1989). A curriculum out of balance: The case of elementary school mathematics. *Educational Researcher. 18*(5), 9-15.

Powell, D. (1985). Selection of reading textbooks at the district level: Is this a rational process? *Book Research Ouarterly,* 24-35

Schmidt, W. H., McKnight, C. C., and Raizen, S. A. (1996). *A splintered vision: An investigation of U.S. science and mathematics education.* Boston, MA; Kluwer Academic Publishers.

Slavin, R. E. (1990). On making a difference. *Educational Researcher. 19 (3)* 30- 34, 44

Snider, V E., & Crawford D. B. (1996). Action research: Implementing Connecting Math Concepts, *Effective School Practices. 15,*17-26.

Stein, M., Carnine, D., & Dixon, R. (1998). Direct Instruction: Integrating curriculum design and effective teaching practice. *Intervention in School and Clinic. 33.* 227-234.

Stein, M., Silbert, J., & Carnine, D. (1997). *Designing Effective Mathematics Instruction: A direct instruction approach.* (3rd ed.) Columbus, OH Merrill.

Tarver, S., & Jung, J. (1995), A comparison of mathematics achievement and mathematics attitudes of first and second graders instructed with either a discovery-learning mathematics curriculum or a direct instruction curriculum. *Effective School Practices. 14*(1), *49-57.*

Tindal, G. A. & Marston, D. B. (1990). *Classroom-based assessment; Evaluating instructional outcomes.* Columbus, OH: Merrill.

Tyson, H., & Woodward, A. (1989). Why students aren't learning very much from textbooks. *Educational Leadership 47,* 14-17.

Tyson-Bernstein, H. (1988). *A conspiracy of good intentions. America's textbook fiasco.* Council for Basic Education: Washington, D.C.

Vreeland, M., Vail, J., Bradley, L., Buetow, C., Cipriano, K., Green, C., Henshaw, P., Huth, E. (1994). Accelerating cognitive growth: The Edison School Math Project. *Effective School Practices, 13*(2), 64-69.

Wellington, J. (1994). Evaluating a mathematics program for adoption: Connecting Math Concepts. *Effective School Practices, 13*(2), 70-75.

Wick, J. W. (1990). *National Achievement Test Technical Manual.* Chicago; American Testronics.

Woodward, A., & Elliot, D. L. (1990). Textbook use and teacher professionalism. In D. L. Elliott & A. Woodward (Eds,), *Textbooks and schooling in the United States* (89th Yearbook of the National Society for the Study of Education, Part 1, pp. 178-193). Chicago; National Society for the Study of Education.