The Longitudinal/Cross-Sectional Study of the Impact of Teaching Mathematics using *Mathematics in Context* on Student Achievement

---

Monograph 1

2004

Purpose, Plans, Goals, and Conduct of the Study

Thomas A. Romberg
University of Wisconsin-Madison

Mary C. Shafer
Northern Illinois University

Wisconsin Center for Education Research
School of Education - University of Wisconsin-Madison

The Longitudinal/Cross-Sectional Study of the Impact of Teaching Mathematics using *Mathematics in Context* on Student Achievement was carried out by the staff of the Wisconsin Center for Education Research, University of Wisconsin-Madison with the support of the National Science Foundation Grant No. REC 0553889. The analysis of the data gathered in this study was conducted by the staff of the Wisconsin Center for Education Research, University of Wisconsin-Madison and funded by the National Science Foundation Grant No. REC 0087511. Additional support for completing the monograph series was provided by Northern Illinois University.

Romberg, T. A., & Shafer, M. C. (Editors). (2004). *Purpose, Plans, Goals, and Conduct of the Study*. (Longitudinal/Cross-Sectional Study of the Impact of Teaching Mathematics using *Mathematics in Context* on Student Achievement: Monograph 1). Madison, WI: University of Wisconsin-Madison.

Romberg, T. A. (2004). Standards-based reform and *Mathematics in Context*. In T. A. Romberg & M. C. Shafer (Editors), *Purpose, Plans, Goals, and Conduct of the Study*. (Longitudinal/Cross-Sectional Study of the Impact of Teaching Mathematics using *Mathematics in Context* on Student Achievement: Monograph 1), 7–17. Madison, WI: University of Wisconsin-Madison.

Romberg, T. A., & Shafer, M. C. (2004). The design of the longitudinal/cross-sectional study. In T. A. Romberg & M. C. Shafer (Editors), *Purpose, Plans, Goals, and Conduct of the Study*. ((Longitudinal/Cross-Sectional Study of the Impact of Teaching Mathematics using *Mathematics in Context* on Student Achievement: Monograph 1)), 18–30. Madison, WI: University of Wisconsin-Madison.

Shafer, M. C. (2004). Instrumentation, sampling, and operational plan. In T. A. Romberg & M. C. Shafer (Editors), *Purpose, Plans, Goals, and Conduct of the Study*. (Longitudinal/Cross-Sectional Study of the Impact of Teaching Mathematics using *Mathematics in Context* on Student Achievement: Monograph 1), 31–50. Madison, WI: University of Wisconsin-Madison.

Shafer, M. C. (2004). Conduct of the study. In T. A. Romberg & M. C. Shafer (Editors), *Purpose, Plans, Goals, and Conduct of the Study*. (Longitudinal/Cross-Sectional Study of the Impact of Teaching Mathematics using *Mathematics in Context* on Student Achievement: Monograph 1), 51–93. Madison, WI: University of Wisconsin-Madison.

The L/CSS Monograph Series

This is the first of eight monographs derived from the National Science Foundation-funded Longitudinal/Cross-Sectional Study of the impact of teaching mathematics using *Mathematics in Context* (National Center for Research in Mathematical Sciences Education & Freudenthal Institute, 1997-1998) on student achievement.

In 1992 the National Science Foundation (NSF) funded several projects to develop new sets of instructional materials that reflected the reform vision of school mathematics espoused by the National Council of Teachers of Mathematics (NCTM, 1989). One of the funded projects was to the National Center for Research in Mathematical Sciences Education (NCRMSE) at the University of Wisconsin–Madison. The project was organized to develop a comprehensive mathematics curriculum for the Grades 5–8 (NSF Grant No. ESI-9054928). Assisted by the staff of the Freudenthal Institute (FI) at the University of Utrecht in The Netherlands, the *Mathematics in Context* (MiC) curriculum materials were created and field-tested prior to being published in 1997-1998 by Encyclopaedia Britannica.

In 1996, as the development of the MiC materials was nearing completion, a proposal was submitted to the National Science Foundation to investigate how teachers were changing their instructional practices in schools whose staffs were using *Mathematics in Context* and how such changed practices affected student achievement. Two NSF grants were awarded to the University of Wisconsin–Madison: first, to conduct a three-year study of the impact of *Mathematics in Context* on student mathematical performance (NSF Grant No. REC-9553889); and second, to analyze the data gathered in that study (NSF Grant No. REC-0087511). This monograph series presents the rationale, development, and conduct of the study of the implementation of the MiC materials in classrooms across the nation, and the results portray the impact of the use of that curriculum on student achievement.

As students and teachers begin to use any of the new mathematics materials, district administrators, mathematics educators, teachers, parents, and funding agencies express cogent needs to demonstrate that the curricula have a positive impact on students' understanding of mathematics. They often want to know the bottom line—the results on measures of achievement that confirm improved student mathematical performance. However, while improved student performance is critical, we contend that just relying on outcome measures to judge the impact of a standards-based program is insufficient. In fact, it is not enough to consider student outcomes in the absence of the effects of the culture in which student learning is situated, the instruction students experience, and their opportunity to learn comprehensive mathematics content in depth and with understanding. The dynamic interplay of all these variables has an impact on student learning, and as such, these variables must be considered in making judgments about the impact of any standards-based curriculum.

This monograph series tells the complex story of the variations in how the MiC materials were used by teachers and students in classrooms that vary in location and ecological culture, and the impact of that variation on the achievement of their students. The story unfolds in eight monographs. This initial monograph provides the background information for the study.

L/CSS Monograph Series on the Impact of Teaching *Mathematics in Context* on Student Achievement

Introduction to Monograph 1

The initial monograph provides the background information for the study. In Chapter 1 the study is situated in the reform efforts in mathematics education. This chapter contains brief overviews of the National Council of Teachers of Mathematics [NCTM] reform vision, the instructional design of the MiC materials, and evidence that MiC was a quality program worthy of its use in school classrooms and its impact on student performance being studied. In Chapter 2 the design of the Longitudinal/Cross-Sectional Study is described. It contains a brief description of the research questions, the structural research model, and the proposed design for gathering data and analyzing the results. Chapter 3 presents the actual instrumentation, sampling plan, and operational plan that were followed. This monograph concludes with Chapter 4 that contains a detailed description of the conduct of the study as it was carried out over the four years of negotiating and working with teachers and administrators in four school districts.

CHAPTER 1: STANDARDS-BASED REFORM AND *MATHEMATICS IN CONTEXT*

Thomas A. Romberg

To situate the study in the reform efforts in mathematics education, this chapter contains brief overviews of the National Council of Teachers of Mathematics [NCTM] reform vision, the instructional design of the MiC materials, and evidence that MiC was a quality program worthy of its use in school classrooms and its impact on student performance being studied.

The Reform Vision for Middle School Mathematics

To appreciate the degree of change envisioned in the reform documents for school mathematics one must first see the recommendations for change as a consequence of the ongoing evolution throughout the world's societies from "The Industrial Age" into "The Information Age." The potential impact of this social evolution on school mathematics was brought vividly to the attention of the American public with the publication of *A Nation at Risk* (National Commission on Excellence in Education, 1983) and *Educating Americans for the 21st Century* (National Science Board Commission on Precollege Education in Mathematics, Science, and Technology, 1983). The authors of those documents claimed that all students needed to be knowledgeable about the mathematical, scientific, and technological aspects of the emerging information age, and our schools were failing to prepare students for their futures.

The mathematical sciences community saw the failure of current school mathematics to properly prepare students for the future because of the traditional, if simplistic, view of learning and teaching commonly practiced in American schools. First, the need to teach students a rich variety of topics is in sharp contrast to the traditional perspective about mathematics. Edward Barbeau (1989) stated:

> Most of the population perceives mathematics as a fixed body of knowledge long set into final form. Its subject matter is the manipulation of numbers and the proving of geometrical deductions. It is a cold and austere discipline which provides no scope for judgment or creativity (p. 2).

This perception reflects the mathematics now studied in school. Second, the common institutional structure for organizing school mathematics that has evolved over the past century fails to challenge students or relate the mathematics they are studying to its uses in the world. As the Mathematical Sciences Education Board (MSEB, 1990) has expressed it:

> To the Romans, a curriculum was a rutted course that guided the path of two-wheeled chariots. Today's mathematics curriculum—a course of study—follows a deeply rutted path directed more by events of the past than by the changing needs of the present. Vast numbers of specific learning objectives, each associated with pedagogical strategies, serves as mileposts along the trail mapped by texts from kindergarten to twelfth grade. Problems are solved not by observing and responding to the natural landscape through which the mathematics curriculum passes, but by mastering time-tested routines, conveniently placed along the path (p. 4).

For Grades K–8, the mathematics curricula include the arithmetic of whole numbers, fractions, decimals, and percents during the first seven or eight years of the school for all students. This is followed in Grade 8 for some middle-school students with a yearlong course in algebra for those students that have successfully completed the hurdles along the arithmetic path. This course structure that we have inherited is deeply embedded not only in our curricular structure, but in our society's expectations regarding schools.

Third, the way students are led along this path also needs to be changed. Instruction in most American classrooms follows a very mechanistic sequence. For example, Mark Weller (1991), based on extended observations, described a typical mathematics lesson as follows:

> A period in this algebra class consisted of three segments . . . .The first five to ten minutes of each period usually began with the correction of the previous day's work. Emphasis was placed upon organization as evidenced by her daily instructions: "Place today's assignment directly behind yesterday's work; don't get them out of order if you want full credit."
>
> The second segment of the period, which lasted from ten to seventeen minutes, was devoted to the presentation of new material. Example problems, which had been placed on the chalkboard prior to the beginning of school, were used throughout the day. After each period, the answers were erased so the problems would not have to be rewritten for each class.
>
> The remaining third of the period was devoted to study time during which students were to work on the day's assignment, which was always to be completed the following day. During this study period, [the teacher] would either look at individual student's work or assist those who had questions or would sit at her desk at the back of the room evaluating notebooks and/or calling students to her desk to inquire about delinquent assignments.
>
> Emphasis was placed upon ascertaining the one correct answer, a focus quite apparent to most students. [The teacher] provided a step-by-step explanation of her thinking processes as the exercises were computed. Students who had skipped steps but who had arrived at the same answer as the teacher were immediately corrected. Her students were expected to use the prescribed computational thinking process presented in the textbook. Practice problems were prototypes of those in the assignment students would eventually be completing. Students rarely raised questions or sought explanations regarding computation. The necessity of "thinking right," as determined by the teacher, was evident in this mathematics classroom (pp. 130-131).

This mechanistic approach to basic skills and concepts illustrates the difficulties with this traditional perspective about school mathematics. It is no wonder that this conception of mathematics isolates it from its uses and other disciplines, since what Weller has described is not so much a form of thinking as a substitute for thinking.

Finally, the technology of traditional instruction includes a teacher directed classroom with students sitting at desks in neat rows, a basal text which is a repository of problem lists, a mass of paper-and-pencil worksheets, and a set of performance tests. Although a few books include things to read, there is very little that is interesting to read. Thus, workbook mathematics gives students little reason to connect ideas in "today's" lesson with those of past lessons or with the real world. The tests currently used ask for answers that are judged right or wrong, but strategies and reasoning used to derive answers are not evaluated.

This portrayal of school mathematics—a tedious, uninteresting path to follow with lots of hurdles to clear—bears little resemblance to what a mathematician or user of mathematics does. What is clear is that students learn a collection of techniques that are useful for some purposes. In fact, when mastery of techniques is defined as knowledge, the acquisition of those techniques becomes an end in itself and the student spends his or her time absorbing what other people have done. In summary, traditional school mathematics fails to provide students with any sense of the importance of the discipline's historical or cultural importance, nor any sense of its usefulness. “We have inherited a mathematics curriculum conforming to the past, blind to the future, and bound by a tradition of minimum expectations” (MSEB, 1990, p. 4).

In contrast, NCTM’s position is that school mathematics should be viewed as a human activity which reflects the work of mathematicians—finding out why given techniques work, inventing new techniques, justifying assertions, and so forth. It also reflects how users of mathematics investigate a problem situation, decide on variables and how to use mathematics to quantify and relate the variables, carry out calculations, make predictions, and verify the utility of the predictions. To make visible this perspective in 1986 the Board of Directors of NCTM established the Commission on Standards for School Mathematics. The reform vision for school mathematics has been described in four documents prepared by NCTM: *Curriculum and Evaluation Standards for School Mathematics* (1989), *Professional Standards for Teaching Mathematics* (1991), *Assessment Standards for School Mathematics* (1995), and *Principles and Standards for School Mathematics* (2000). These documents argue that all aspects of school mathematics—content, teaching, and assessment—need to change, and that the changes must be reflected in the work of students and teachers in classrooms. To realize this vision there need to be shifts in classrooms:

- in content toward a rich variety of mathematical topics and problem situations for all students, and away from just arithmetic for most;
- in learning toward investigating, and away from memorizing and repeating;
- in teaching toward questioning and listening, and away from telling; and
- evaluation of students toward useful information about student performances, and away from externally mandated tests.

Furthermore, it has been assumed that if such changes occurred in classrooms, the students’ mathematical achievement would dramatically improve.

If mathematics is to serve students' needs to make sense of experience arising outside of mathematics itself, including making sense in the various sciences, it must be firmly rooted in and connected to that experience. And its systems of signs and symbols must be learned and experienced as genuine functioning languages—for expressing, communicating, reasoning, computing, abstracting, generalizing, and formalizing—that the student experiences as serving his or her real needs. Likewise, the systematic logical forms of reasoning and argument must be learned by serving personally and socially experienced needs for certainty and reliability—for establishing for the student what is true and what is not true.

Also, it was assumed that the problem situations that are likely to introduce students to the language and uses of mathematics are those that encourage “mathematization.” Such situations include those that are subject to measure and quantification, that embody quantifiable change and variation, that involve specifiable uncertainty, that involve our place in space and the spatial features of the world we inhabit and construct, and that involve symbolic algorithms and more abstract structures. The mathematical systems of signs and symbols extend the limited powers of the human mind in many directions, and they make possible a long-term cultural growth of the subject matter that crosses generations. Finally, such situations embody systematic forms of reasoning and argument to help establish the certainty, generality, consistency and reliability of one’s mathematical assertions.

In summary, the reform vision sees students learning to formulate problems and develop and apply strategies to their solution (both within and outside mathematics). In a range of contexts, they verify and interpret results and generalize solutions to new problem situations. In so doing, they apply mathematical modeling and become confident in their ability to address real-world problem situations. As they reason through their problem situations, students develop the habit of making and evaluating conjectures, and of constructing, following, and judging valid arguments. In the process they deduce and induce, apply spatial, proportional, algebraic, and graphic reasoning, construct proofs, and formulate counterexamples.

The new view of mathematics is, above all, integrative; it sees everything as part of a larger whole, with each part sharing reciprocal relationships with other parts. It stresses the acquisition of understanding by all, including the traditionally underprivileged, to the highest extent of their capability, rather than the selection and promotion of an elite. It is a philosophy that simultaneously stresses erudition and common sense, integration through application, and innovation through creativity. Most importantly, it stresses the creation of knowledge. Against this broad and ambitious view of mathematics, traditional school mathematics appears thin, artificial, and isolated.

The practical consequences of this vision sees students sitting around tables or group work stations, working on collections of problem-related activities, being coached by teachers, and their progress being judged via observation as well as via the quality of their products.

In summary, to implement NCTM’s vision of reform classroom instruction should exhibit this non-routine pattern of instruction that allows students to become mathematically literate. The complexity of instructional issues involved in creating and sustaining such classrooms include the interconnected roles of tasks on the one hand and how students and their teachers talk about mathematics on the other, how technological tools can help in the development of such instruction, the normative beliefs within a classroom about how one does mathematics, the organizational structures of the classroom, the role of professional development in helping teachers to develop their own classrooms that promote understanding, how the school, as an organization, supports (or impedes) the work of teachers in developing and sustaining these classrooms, and how non-school agents (such as parents), agencies (district), and their actions support (or impede) the development of these classrooms.

Overview of *Mathematics in Context*

When the National Science Foundation funded projects to develop new middle-school curricula that reflected NCTM's curriculum standards (1989), one of the products of that effort was Mathematics in Context (MiC). The materials consist of 40 curriculum units, 10 at each grade level 5-8, and include assessment materials; a teacher's guide for each unit; and two supplementary packets, *News in Numbers,* which provide extra opportunities for students to develop estimation skills in contexts similar to those in the curriculum units, and *Number Tools*, consisting of a collection of tasks designed to help students maintain basic skills.

The instructional design of the MiC materials was based on the Dutch "Realistic Mathematics Education" (RME), an approach to school mathematics deemed consistent with NCTM's vision. This approach to instruction grew from on the ideas of the mathematician Hans Freudenthal (1983) who believed that "students are entitled to recapitulate in a fashion the learning process of mankind" (p. ix). He stated:

> Mathematical structures are not presented to the learner so that they might be filled with realities . . . .They arise, instead, from reality itself, which is not a fixed datum, but expands continuously in individual and collective learning process" (Freudenthal, 1987, p. 280).

In MiC, Freudenthal's beliefs about what mathematics students should learn and how they should learn it are made operational. The Dutch instructional approach includes four components:

- goals that reflect Freudenthal's notions of students recapitulating the creation of the discipline,
- the design of a structured set of instructional activities in mathematical domains that reflect those goals,
- the provision to teachers of a guide to strategies that support students' investigation of reality, and
- the development of an assessment system that monitors both group and individual student progress.

Goals for Students

The underlying goal of instruction is that students need to participate in the mathematization of reality and can do that by exploring aspects of several mathematical domains. Students gradually shift from creating "models of" problem situations to "models for" mathematical reasoning and problem solving. By doing this they should come to understand both how mathematics has developed and how it is used in the world in which they live. For school mathematics, vision emphasizes the notion that mathematics is a plural noun in that there are several intertwined strands or domains, each an assemblage of ideas defined by the community of mathematicians, mathematics educators, and users of mathematics. In MiC, the instructional units are organized in two ways, first, by grade level, and second by mathematical strand. Over the course of this four-year curriculum, middle school students explore and connect the mathematical strands:

Number: Whole numbers, common fractions, ratio, decimal fractions, percents, integers;
Algebra: Creation of expressions, tables, graphs, and formulas from patterns and functions;

Geometry: Measurement, spatial visualization, synthetic geometry, coordinate and transformational geometry; and

Statistics and probability: Data visualization, chance, distribution and variability, and quantification of expectations.

Although each unit in MiC emphasizes specific topics within a particular mathematical strand, most units involve ideas from several strands and emphasize the interconnectedness of those ideas.

Instructional Activities

Designing instructional activities based on the domain-based perspective involves creating a collection of problem situations that engage students so that they explore each domain. To accomplish this key features and resources of the domain that are important for students to find, discover, use, or even invent for themselves need to be identified. This is not an easy task, as Webb and Romberg (1992) argued:

> Knowledge of a domain is viewed as forming a network of multiple possible paths and not partitioned into discrete segments . . . .Over time, the maturation of a student's functioning within a domain should be observed by noting the formation of new linkages, the variation in the situations the person is able to work with, the degree of abstraction that is applied, and the level of reasoning applied. (p. 47)

Second, activities that encourage students to explore the domain need to be identified and organized in a structured manner that allows for growth in what is learned. "Mathematical concepts, structures, and ideas have been invented as tools to organize the phenomena of the physical, social, and mental world" (Freudenthal, 1983, p. ix) implies that the student activities need to be authentic, or *realistic*, in that they reflect real phenomena, or actual situations, from which mathematics has developed the interpretation of which requires the use of mathematics. The implication is that a sequence of activities must be designed such that students grow in their knowledge and understanding of the ideas in a domain over time. Each activity, therefore, has to be justifiable in terms of potential endpoints in a learning sequence. In particular, it is assumed that students come to understand from their experiences solving problems. This involves making sense of a situation by seeing and extracting the mathematics embedded within it. For students, this involves learning to represent quantitative and spatial relationships in a broad range of situations, to express those relations using the terms, signs, and symbols of mathematics, to use procedures with those signs and symbols following understood rules to carry out numerical and symbolic calculations, and to make predictions and interpret results based on the use of those procedures. This assumption also implies that students need to understand the rationale for and use of the mathematical terms, signs, symbols, and rules.

This instructional approach has been described as "bottom-up," in that students construct models for themselves and these models serve as the basis for developing formal mathematical knowledge. To be more precise, at first, a model is constituted as a context-specific model of a situation. Later, the model is generalized over situations. Thus, the model changes in character; it becomes an entity on its own. In this new shape it can function as a basis, a model for mathematical reasoning on a formal level (Gravemeijer, 1994, p. 100). It was also assumed that the sequence of contextual activities should help students gradually develop methods for symbolizing problem situations. Thus, all activities are related to end goals, and are seen as a means of helping students in their transition from informal to formal semiotics.

Teaching

The role of the teacher is to select appropriate experiences and provide guidance so that a student's informal models evolve into models for increasingly abstract mathematical reasoning. The development of ways of symbolizing problem situations and the transition from informal to formal semiotics are important aspects of these instructional assumptions. This implies that instruction, as is too commonly done in mathematics classes, should not start with presenting students the formal terms, signs, symbols, and rules and later expecting them to use these formal ideas to solve problems. Instead the activities should lead students to the need for the formal semiotics of mathematics. The implication for students is that they should gradually develop ways of representing complex problems. Psychologically this process is called “progressive formalization.” Teachers using MiC in their classrooms are expected to provide opportunities for students to explore ideas via a variety of engaging problem situations in a classroom environment that rewards alternative solution strategies, encourages appropriate mathematical modeling, invites reflective thinking, and allows genuine sharing of information.

Assessment

The RME approach to assessment is closely aligned with instruction and is seen as part of daily instructional practice. However, rather than just including tasks that mimic content that has been covered, the Dutch include open tasks that expect students to relate concepts and procedures and use them to solve non-routine problems. Also, the Dutch assume that a high level of understanding in any domain will occur only as a consequence of a variety of experiences. For assessment, then, ascertaining how many artifacts of the domain a student can identify is not sufficient. Instead, assessment should focus on the ways students identify and use such artifacts to solve increasingly complex tasks. Such focus should provide reliable evidence of what a student is able to do in any domain at a point in time. With additional experiences (often over the course of several years), one would expect growth in the level or complexity of tasks a given student is able to solve.

Steps in Developing MiC

The development of the MiC materials took six years and involved the following steps. First, an international advisory committee of mathematics educators, mathematicians, scientists, curriculum supervisors, principals, and teachers was formed to ensure that MiC conformed to the goals and philosophy of NCTM’s *Standards*. Initially, this committee met and prepared a “blueprint” document to guide the development of the materials. Second, Freudenthal Institute staff prepared initial drafts of the 40 curriculum units based on the blueprint. Researchers at the University of Wisconsin–Madison then modified these units and developed them further to create a curriculum appropriate for U.S. students and teachers. Third, pilot versions of the individual units were tested in middle school classrooms in Wisconsin, where both students and teachers provided feedback and suggestions for revisions. Fourth, the units were then revised and field-tested at additional schools in other states and in Puerto Rico. The field tests included trials of a sequence of curriculum units at each grade level. Data collected during the field tests were used to again revise the units and prepare detailed guides for teachers for the final step, the commercial publication of the materials. The final product, *Mathematics in Context*, is a complete mathematics program, with the resources and support materials necessary for successful implementation: student booklets, teacher guides, and assessment materials.

In summary, MiC is a carefully designed and field-tested curriculum for Grades 5–8. Each MiC unit includes tasks and questions designed to engage students in mathematical thinking and discourse. Students are expected to explore mathematical relationships, develop their own strategies for solving problems, use appropriate problem-solving tools, work together cooperatively, and value each other’s strategies. They are encouraged to explain their thinking as well as their solutions. Teachers are expected to help students develop common understanding and usage of the terms, signs, symbols and rules of mathematics in order to assist students in articulating their thinking.

Quality of the MiC curriculum

When the commercial MiC materials were published and sold to school districts, we were confident that MiC was an exemplary program that would have important impact on students’ mathematical performance. The potential of the program has been validated by external reviews, growing sales of the program to schools, and initial data on the impact of using the program on student achievement. Brief summaries of this information we deemed important to justify the cost and effort involved in conducting the study reported in this monograph series.

External Reviews of MiC

Four external reviews of the program have been conducted. First, before publication several eminent mathematicians, educational psychologists, and educators from the University of Wisconsin substantiated the value of the materials via detailed content reviews. A summary of these reviews can be found in Romberg and Pedro (1996). For example, in his comments about the MiC number strand, Dr. Robert Wilson, a professor of mathematics, said:

> First let me say that I loved these units! Naturally there were some parts I liked more than others, and I have a few nits to pick, but I find the content, the organization, and (to the extent that I can tell at my distance from the middle-school student) the appeal of the presentation to be simply great!

Dr. Simon Hellerstein, a professor of mathematics and former chair of the department of mathematics, in a review of the units in the algebra strand, wrote—

> Incorporating pre-algebraic ideas and skills into the middle-school math curriculum, culminating in the eighth grade with some basic formal algebra, is a splendid undertaking. The end goals of the algebra strand as listed in teacher guide are commendable. In those schools where the goals are actually achieved, students will have been well served and better prepared to tackle high school algebra.

Dr. Michael Bleicher, a professor of mathematics, in his review of the units in the geometry strand noted that—

> The mathematics contained in these units far surpasses that in any previous middle-school curriculum with which I am familiar. . . .This is an ambitious program introducing a great deal of new mathematics into the school system. Even those topics that are not new have here a novel presentation.

A professor of cognitive psychology, Dr. Richard Lehrer, reviewed the units in the geometry strand for their pedagogical content:

> The activities begin with student experiences or with contexts that are apt to be interesting to a wide range of middle-school students. . . .The units on geometry are integrated with the mathematics developed in other units. For example, units on measurement make frequent reference to ideas developed on other units about data and uncertainty. This is an exceptional feature of this curriculum; usually, geometry is fragmented and isolated in the middle school. Moreover, the geometry units also are integrated and are not simply a collection of topics. The geometry units provide frequent opportunity for students to learn about the uses of mathematical modeling. I have never seen much of an emphasis on this form of applied mathematics so early. It's an outstanding feature of this curriculum. Finally, the assessment items developed for the units seem to measure much more than rote response. Instead, the items demand thoughtfulness and understanding. Student responses to them could conceivably be used as a guide to further instruction.

In summary, these reviewers found that MiC was impressive in that it provided students with a range and depth of mathematics rarely covered in middle-school mathematics curricula.

Second, the American Association for the Advancement of Science (AAAS) carried out a review of middle-school mathematics curricula in 1998. The criteria used were six benchmarks from the 2061 project (AAAS, 1993). The report “Benchmarks-Based Evaluations” (AAAS, 1998) indicated how well MiC attempts to address the substance, breadth, and sophistication of the ideas contained in each of the six mathematics benchmarks that were selected for the analysis. Overall, the reviewers for AAAS found that MiC covered almost all the content referred to in these six benchmarks, and that the instructional approach has considerable potential for learning.

Third, in 1998 the U.S. Department of Education undertook a “quality of program” review of a large number K–12 mathematics and science programs (U.S. Department of Education, 1999). Four separate panels reviewed MiC with respect to seven criteria on a 5-level scale (Exceptionally, Strongly, Adequately, Inadequately, Not at all). The criteria were:

1. The program’s learning goals are challenging, clear, and appropriate for the intended student population.

2. The program's content is aligned with its learning goals, and is accurate and appropriate for the intended student population.
3. The program's instructional design is appropriate, engaging, and appropriate for the intended student population.
4. The program's system of assessment is appropriate and designed to inform student learning and to guide teachers' instructional decisions.
5. The program can be successfully implemented, adopted, or adapted in multiple educational settings.
6. The program's learning goals reflect the vision promoted in national standards in mathematics.
7. The program addresses important individual and societal needs.

Overall, the ratings indicate that the program met all seven criteria "strongly" or "exceptionally." Only for Criterion 5 were there two ratings of "adequate." The justification for this rating was "Unless the entire package is used, a lot of the program's strength would be dissipated. Using individual units for enrichment, while attractive because of their design, would dilute the overall thrust of the program."

Finally, the staff of the Department of Applied Mathematics at the University of Washington prepared a report for the National Science Foundation (Adams, Tung, & Warfield, 2000) comparing three middle-school programs. MiC was one of the curricula examined. The project staff used the middle-school standards from *Principles and Standards for School Mathematics* (NCTM, 2000) as the criteria. They found that MiC fully meets most of the standards questions. Only in the content standards for number and geometry were low judgments made. For example, "Does the curriculum enable all students to develop and analyze algorithms for computing with fractions, decimals, and integers and develop fluency in their use?" The justification for a low score was that "the units are mathematically correct, are done well, teach for conceptual understanding, but are not at the right level to give fluency." Unfortunately, these reviewers failed to examine the supplementary *Number Tools* material designed to help students gain fluency. Similarly, a low score in geometry was the lack of using coordinate geometry to study properties of pairs of parallel or perpendicular sides of shapes. Their overall conclusion was that MiC corresponds well to the NCTM standards.

These four external reviews of MiC, although conducted for different purposes and using different criteria, each present the picture of a mathematically and pedagogically strong middle-school curriculum.

Impact on Student Achievement

While considerable formative information about student achievement on each unit was collected during the development of MiC, the only summary data collected was from publicly reported data sets and submitted by various MiC implementation sites (see Webb & Meyer, 2001 for details). For example, the most extensive data on student achievement comes from Ames, IA. Since field-testing of MiC units began in 1995, composite district scores on the Grade 7 *Iowa Test of Basic Skills* (Hoover, Hieronymus, Dunbar, & Frisbie, 1996) have remained above the 90th percentile and in spring 2000 were at the 99th percentile. Results on the *New Standards Reference Exam* (University of Pittsburgh & National Center on Education and the Economy, 1997) demonstrate exemplary achievement of Grade 8 students in mathematical skills, concepts, and problem solving after five years of using MiC. During the period 1997 to 2001, the percent of eighth-grade students meeting or exceeding the standard has progressed continuously beyond the national norm: from 79% to 82% in skills (33% national norm); from 57% to 73% in concepts (20% national norm); and from 50% to 65% in problem solving (11% national norm).

In summary, data reported by districts establish a positive trend in student achievement in schools using MiC. With regard to state assessments, student use of MiC results in substantial gains.

The Need to Study the Implementation and Impact of *Mathematics in Context*

The problem with the vision of school mathematics, as outlined in the previous paragraphs, is that they are ideas put forward by educational leaders, policymakers, and professors about what mathematical content, pedagogy, and assessments should be. Implementation of such ideals can be undermined by a number of factors. For example, not all persons agree with the goal of mathematical literacy for all, some influential persons believe that the current course of study works reasonably well (particularly for their children), etc. In fact, as Labaree points out that during the past century calls for reform have had “remarkably little effect on the character of teaching and leaning in American classrooms” (1999, p. 42). Instead of changing conventional practices the common response to calls for reform has been the “nominal” adoption of the reform ideas. Schools have used the reform labels but did not follow most of the practices advocated. It is often a political necessity for schools and teachers to claim they are using a “new” program even if classroom practices have not changed (Romberg, 1985). Thus, to document the impact of any reform efforts in classrooms one needs to examine the degree to which the reform vision actually has been implemented.

Thus, given the information from the reviews that the new materials and procedures were consistent with the NCTM’s vision, and the initial reports of student achievement were encouraging, we decided that a more careful examination of the programs use and its impact on student performance was warranted. In fact, the assumption “if changes in middle-school mathematics occurred that reflected NCTM’s vision, there would be dramatic improvement in student achievement” needed to be empirically examined.

CHAPTER 2: THE DESIGN OF THE LONGITUDINAL/CROSS-SECTIONAL STUDY

Thomas A. Romberg and Mary C. Shafer

In 1995 a proposal was submitted to the National Science Foundation (NSF) by the Wisconsin Center for Education Research at the University of Wisconsin–Madison to document the impact of the *Mathematics in Context* (MiC) on the mathematical performance of students over one to four years, and to compare their performance with that of students who have been studying in conventional programs. To gather such evidence a four-year combined longitudinal/cross-sectional comparative study was proposed. In that proposal three questions were raised.

1. What is the impact of the MiC instructional approach on student performance?

MiC is a four-year program designed for Grades 5–8, and its design included the deliberate structuring of activities so that over the four years students progressed from informal to formal ideas in four content strands. Thus, a longitudinal study designed to track student overall mathematical growth, and growth in each of the content stands was warranted.

2. How is this impact different from that of traditional instruction on student performance?

To answer this question we proposed to gather similar data for students in matched conventional mathematics classes as was to be gathered for students using MiC materials.

3. What variables associated with classroom instruction account for variation in student performance?

This question was raised for three reasons. First, although the results on measures of achievement that confirm improved student mathematical performance are very important, we contend that just relying on outcome measures to judge the value of a standards-based program is insufficient. Second, in the field-testing of the materials there was considerable variation in the observed patterns of instruction. Teachers often chose to augment MiC units with conventional worksheets, used only some MiC units, failed to emphasize key concepts, etc. (Romberg, 1997). Thus, it is not enough to consider outcomes in the absence of the effects of the instructional setting in which student learning is situated, and the students' opportunity to learn comprehensive mathematics content in depth and with understanding. The dynamic interplay of such variables has an impact on student learning. Third, because this observed variation occurred with teachers committed to reform in predominantly suburban schools, we were anxious to see how the program would be used during initial implementation in urban schools with minority students and in turn what the impact of its use would be on student performance.

Given these questions and concerns, the research design for the study had two primary components. First, to answer Questions 1 and 3 we proposed to use a structural research model to examine the impact of the multivariate classroom interventions anticipated when teachers use MiC materials in their classrooms. Structural modeling involves identifying the key variables hypothesized to be critical and locating causal paths between the variables. The methodological issues that arise involve attempts to capture, interpret, and report variation in the methods used to support curricular change, implementation of reform-based curricula, and student performance. Technically, this involves specification, measurement, estimation, and statistical inference, rather than the control of sources of variations. Comprehensive examination of such variables involves both qualitative and quantitative methods that explicate potential differences in student performance as a consequence of studying either reform-based or conventional curricula. This process of structural modeling together with the development of composite scales based on such a model we contend is an effective way to carry out curricular research in school settings.

To answer Question 2 we proposed to supplement the structural model with a quasi-experiment by comparing student performance in a sample of classrooms using the MiC materials with that in a similar sample of classrooms using traditional materials. The assignment of students to classrooms was nonrandom, and comparable information was collected on the variables in the structural model for both groups. This design allows us to build a case as to whether observed group differences on outcome measures were a consequence of using a particular curriculum as opposed to an inherent product of preexisting group differences described through some of the variables in the research model (Campbell & Stanley, 1963).

Note that there is a long-held belief that such comparative experiments provide the best evidence for making causal statements about alternate education practices. Unfortunately, what the naïve person may view as a simple comparative experiment turns out to be very complex when examining differences between instruction in different social contexts with considerable variation in prior, independent, intervening, outcome, and consequent variables. This happens because of the difficulty involved in arranging student and school settings to control potential sources of possible co-variation, a control more easily achieved in laboratory settings. The dynamic interplay of all these variables has an impact on student learning, and as such, these variables must be considered in any comparison of instructional programs in real classrooms.

In summary, the initial proposal included the following subsections:

- the structural research model,
- instrumentation,
- the proposed sites,
- the design for data gathering,
- the analysis procedures, and
- what we expected to learn.

In this monograph what was initially proposed for each sub-section of the proposal is summarized, followed by revisions that were negotiated with the National Science Foundation when the project was funded, and later modifications based on technical considerations. Also, note that what was actually done during the conduct of the study is presented in Chapter 4.

The Structural Research Model

The research model for this study is an adaptation of a structural model for monitoring changes in school mathematics (Romberg, 1987).[1] The model is composed of variables and their theoretical interrelationships (represented by arrows in the model). This model, illustrated in Figure 2.1, includes 14 variables in five categories (prior, independent, intervening, outcome, and consequent).

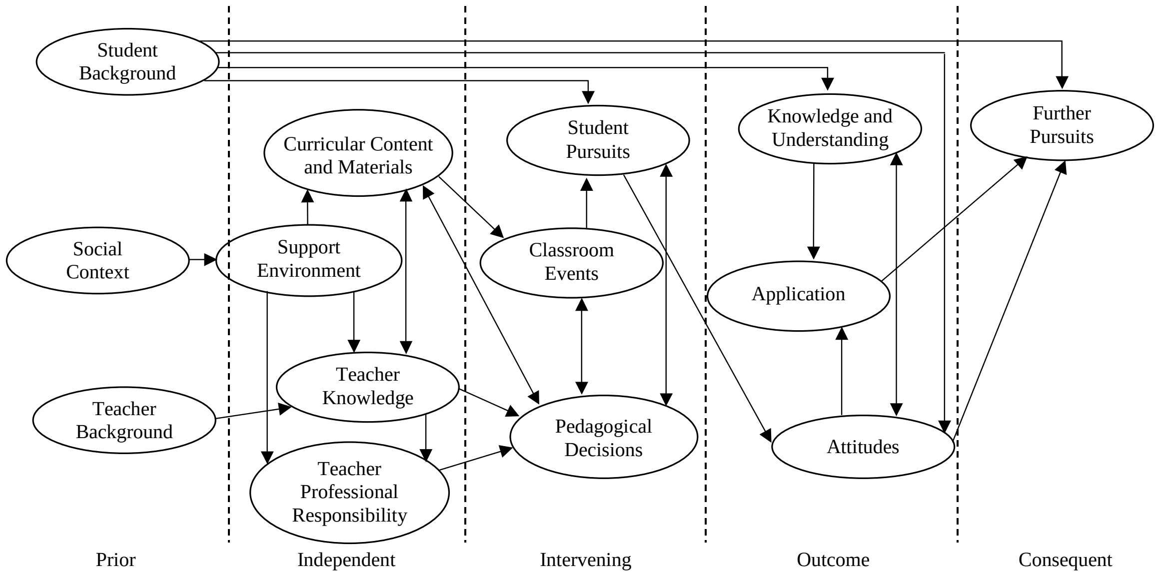

Figure 2-1. Revised model for the monitoring of school mathematics.

[1] This research model is a revised version of the original (Romberg, 1987). Names for some research variables were updated (e.g., "content" was replaced with "curricular content and materials"). Also, some hypothesized relationships (indicated by arrows) were added. For example, an arrow was added from pedagogical decisions to curricular content and materials. Although curricular materials have an impact on pedagogical decisions (e.g., the sequence of chapters or instructional units), pedagogical decisions also affect the materials used (e.g., adding supplementary worksheets for skill practice). Arrows were also added from curricular content and materials to teacher knowledge, from teacher knowledge to pedagogical decisions, and from classroom events to pedagogical decisions.

Prior Variables

By documenting baseline information, prior variables permit the exploration of change in mathematics instruction and the documentation of the extent of that change, contrasting the type and the degree of change, as data are gathered over time. This structural model takes into account three variables: background and prior knowledge of students, teacher background, and the social context or culture in which particular schools operate.

Student background. Among this background information is the students' gender, ethnicity, primary language, and previous standardized test scores. In addition, information about students' reasoning in mathematics and attitudes toward mathematics was gathered.

Teacher background. This information that was collected includes the teachers' mathematics preparation, teaching experience, conceptions about mathematics teaching and learning, and use of assessment procedures.

School context. This variable takes into account multifaceted information about the school environment: the demographic location of the school; the percentage of students by gender, ethnicity, use of the English language, and socioeconomic status; the vision for student learning in general and mathematics in particular; and the support environment in which teachers and students work.

Independent Variables

The independent variables represented in the structural model include curricular content and materials, the support environment available for students and teachers, teacher knowledge, and teacher professional responsibility.

Content and materials. This variable refers to the mathematical content students have an opportunity to learn in classrooms. For the middle school, the strands proposed by NCTM (1989) are number (whole numbers, common fractions, ratio, decimal fractions, percents, integers), algebra (creation of expressions, tables, graphs, and formulas from patterns and functions), geometry (measurement, spatial visualization, synthetic geometry, coordinate and transformational geometry), and statistics and probability (data visualization, chance, distribution and variability, and quantification of expectations). Instructional units and lessons center on specific topics within a particular mathematical strand. For this variable, both the actual content covered in classrooms and the types of materials used during instruction are considered.

The support environment. The support environment refers to that environment cultivated and experienced by school staff and students. Schools vary in their efforts to promote quality instructional experiences for all students and to develop and support professional communities for staff—two elements of school environment dependent on the complex interaction of both cultural and structural conditions (Newmann & Associates, 1996). Successful schools develop cultures that include both high expectations for student learning (as well as normative practices of staff that increase the likelihood of meeting those expectations) and structural features such as sustained time for collaborative discussions, observation, instructional planning, and staff development. The variable for support environment takes both school cultural and structural conditions into consideration.

Teacher knowledge. Reform-based curricula often prove challenging to teachers. In a synthesis of eight case studies involving teachers using the prepublication versions of MiC, Romberg (1997) outlined features of the mathematical content of MiC that were problematic for teachers. In MiC, topics traditionally reserved for high-school students are introduced to middle-school students in contextualized settings with emphasis on student reasoning rather than on procedures. The order of topics throughout the grade levels in MiC units is also different from that in conventional middle-school mathematics curricula (e.g., concepts related to percent are introduced in fifth and sixth grade in MiC rather than, more conventionally, in eighth grade). This change of order along with unfamiliar content sometimes meant that the elementary-school teachers in these studies were unfamiliar with teaching the topics in MiC and were also concerned that they were not covering material traditionally taught at their grade levels.

Furthermore, although MiC units primarily present mathematical ideas from one content strand, many activities involve topics from several strands (in order to emphasize the interrelated nature of mathematical ideas) and include some problems that can be approached in multiple ways (opening possibilities for students to draw on ideas from several content strands). This change provided challenges for the case study teachers who were accustomed to teaching mathematics as isolated pieces of knowledge. Even when these teachers were familiar with the mathematical content of the MiC units, they were at times still unsure how to continue lessons when the tasks, and/or students' approaches, involved a variety of strategies.

In MiC, as in other reform curricula, the rigor of the mathematical content, the variety of strategies students used as they worked to solve problems, and the blending of content strands places new, challenging demands on the teachers. The teacher knowledge variable in the structural model takes into account (a) the teachers' developing understanding of pedagogical methods for teaching broader, deeper, and sometimes new mathematical content, and (b) the opportunities teachers have to learn mathematics content in ways that are more compatible with the processes reform-based curricula require of students.

Teacher professional responsibility. This variable involves the ways in which teachers learn about new curricula and approaches in mathematics teaching and learning, specifically, teachers' opportunities to read professional literature, to participate in in-service sessions, and to attend professional meetings. The use of reform-based curricula can motivate changes in teaching mathematics (Romberg, 1997), but teacher–student and student–student interaction in MiC classrooms is often complex, influencing teachers to change their instructional practice and to address different classroom management issues, including facilitating effective student work in small groups, developing a classroom atmosphere that encourages a variety of student explorations of new concepts, and guiding sometimes overly enthusiastic classroom discussions.

In the eight case studies (Romberg, 1997) mentioned earlier, as teachers began to encourage students to rely on their own thinking, they learned more about their students' mathematical reasoning and abilities. They began to observe and listen to their students and recognized the need to develop methods of informal assessment. This shift from conventional instructional practice, authority, and expectations to a reform approach presented a serious challenge yet also promoted teachers' understanding of their students' learning.

Intervening Variables

Intervening variables are directly influenced by changes in the independent variables and significantly affect student outcomes. The intervening variables capture an array of complex teacher decisions in planning and interactive decision making that center around generating a learning environment in which students are encouraged to develop and communicate an interrelated set of mathematical ideas. These decisions also influence the type of assessment information gathered about each student's mathematical knowledge and dispositions toward mathematics. Intervening variables also capture students' active involvement in learning and applying mathematics as part of everyday classroom events. In the structural model, three intervening variables are identified: pedagogical decisions, classroom events, and student pursuits.

Pedagogical decisions. The pedagogical decisions teachers make prior to and during instruction have a direct impact on student learning. These decisions include deliberate advanced planning such as student grouping for instruction, time allotted to specific aspects of lessons, emphases given during instruction, and other modifications of the intended curriculum. Pedagogical decisions also include teachers' decisions made during instruction, for example, evaluating the correctness of student solutions or resolving critical incidents (related to student understanding of mathematical content or student behavior) during class. If instruction is to increase student involvement in learning mathematics, examine growth in student knowledge over time, and ensure that equitable assessment consideration is given to each student in class, teachers need to make decisions about the organization and documentation of multiple pieces of assessment information for each student (Shafer, 1996). This variable, then, represents the teacher's decisions in defining the actual curriculum.

*Classroom events***.** Pedagogical methods that encourage students to become actively involved in their learning as well as other teacher behaviors during teaching undoubtedly influence student outcomes. Reform-based instruction embodies a shift away from students repeatedly executing routine procedures toward learning mathematics through investigating, reasoning, and communicating their ideas to others. Reform instruction encourages higher order thinking. In order to discover new meanings and understandings, students need to "manipulate information and ideas in ways that transform their meaning and implications . . . combining facts and ideas in order to synthesize, generalize, explain, hypothesize, or arrive at some conclusion or interpretation" (Newmann, Secada, & Wehlage, 1995, pp. 86–87).

Classroom discourse (i.e., expressing thinking, representing mathematical ideas, agreeing and disagreeing) is central to learning mathematics as a domain of inquiry (NCTM, 1991). Encouraging students to look for relationships among mathematical ideas, extend and apply mathematics in novel situations, reflect on their thinking and articulate it to others, and make mathematics their own are significant elements in teaching and learning mathematics with understanding (Carpenter & Lehrer, 1999). These critical elements of classroom interaction should be interwoven as commonplace events during class time.

Student pursuits. In reform recommendations, student involvement is characterized by verbs such as "explore, justify, represent, solve, construct, discuss, use, investigate, describe, develop, and predict" (NCTM, 1989, p. 17). In reform-based classrooms, expectations for students to express their thinking, discuss interpretations of problem situations, consider different levels and qualities of solution strategies shared in the group, and answer questions from others about their reasoning are evident and valued.

Outcome and Consequent Variables

The expected outcomes of mathematics instruction are that students will acquire knowledge of concepts in various mathematical domains and proficiency with mathematical skills, and the ability to apply mathematics in various situations, in addition to developing favorable attitudes toward mathematics and its usefulness. The three outcome variables we include in the structural model are knowledge and understanding, application, and attitudes. We expect as a consequence of these student outcomes that students will continue to study mathematics. The consequent variable, therefore, refers to students' transition into high-school mathematics and the number and type of courses in which they intend to enroll in high school.

Knowledge and understanding. Because the reform approach to school mathematics takes a different perspective of what it means to know mathematics and has different expectations for students' learning of mathematics, reform assessments must provide attention to comprehensive mathematical content appropriate for particular grade levels; relationships among concepts or between concepts and procedures; problem-solving approaches; ways of representing situations; levels of reasoning applied in solving problems; and communication of reasoning. Multiple-choice tests alone will not suffice for assessing these aspects of students' knowledge of mathematics. As appropriate assessments are developed and used, analyses of student responses necessarily move beyond merely checking for correct responses. Student responses are scaled with scoring rubrics (allowing for analysis of partial solutions), and problem-solving strategies are coded (permitting discrimination among problem-solving approaches). Assessment responses can then be aggregated by mathematical content, solution strategies, errors made, and other dimensions such as level of reasoning applied in solving problems.

Application. The second outcome variable, application, relates to the expectation that students will apply the knowledge they acquire. We assume that the more familiar students become with mathematical concepts and procedures and the more coherent and interconnected the mathematical ideas they acquire, the more readily they will be able to solve novel problems. Word problems that require students solely to use specific procedures and concepts learned directly during instruction are not sufficient to assess multidimensional aspects of mathematical knowledge. What are needed are more complex problems appropriate for particular grade levels, in which students apply their knowledge of mathematics both in realistic situations and in mathematical contexts. Such problems also need to provide students opportunities to make connections among mathematical ideas and to use elements of mathematical analysis (e.g., mathematizing problem situations, selecting appropriate mathematical tools and representations, forming generalizations, and writing mathematical arguments).

Student attitudes. Development of favorable student attitudes toward mathematics and the usefulness of mathematics are the third of the outcome variables. The attitude scale examines student attitudes to mathematics and includes: interest in and excitement about mathematics, effort to succeed confidence level in mathematics, and attribution of success or failure in mathematics. We expect that as students learn mathematics in more comprehensive and engaging programs, their experiences will improve their attitudes toward mathematics.

Further pursuits. As students begin Grade 9, they undoubtedly have expectations (and perhaps reservations) about their future success in mathematics courses. The transition into high-school mathematics as well as the number and type of courses they intend to take in high school are affected by the extent of success they had during their middle-school years. As students learn broader mathematics content with greater involvement in their learning, we expect students will continue to study mathematics, exceeding the minimum requirements for high-school graduation.

There were no changes in the specification of variables in this model from the original proposal to the revised proposal in the negotiations with NSF.

Instrumentation

In the original proposal to NSF we proposed to develop for each variable in the model one or more indices or indicators. The methods of gathering evidence were to include for the *prior variables*:

- Fixed characteristics of the students (gender, age, socioeconomic status, ethnic heritage, etc.) gathered via a brief survey.
- Student prior mathematical knowledge and level of reasoning gathered by administering a standardized test, and the *Mathematical Problem Solving Profiles* (Collis & Romberg, 1992).
- School context (demographic location, average socioeconomic status of the students, the percentage of limited-English-speaking students, etc.) gathered via a survey.
- Fixed characteristics of the teachers (gender, age, socioeconomic status, ethnic heritage, mathematics preparation, teaching experience, familiarity with NCTM's *Standards*, etc.) gathered via a survey.

Information about the *independent variables* was to include:

- Descriptions of the administrative support gathered via structured interviews.
- General descriptions of both the available curriculum in each class (what is stated in guidelines, texts, tests, etc.) and the actual curriculum (what is covered and emphasized in classrooms). This data was to be gathered via periodic interviews.
- Teacher self-reports about changes in beliefs and knowledge (as a consequence of workshops, experience of teaching MiC, etc.) gathered via periodic open interviews.
- Teacher self-reports of professional activities (work with other teachers, involvement in NCTM affiliates, conference participation, etc.) gathered via periodic open interviews.

Information about all three of the *intervening variables* was to be derived from classroom logs teachers will be asked to keep and follow-up open interviews. The data was to include:

- A listing of instructional decisions made by teachers (time allocated to topics, adaptations of activities, etc.) and the reasons given.

- A confirmation of the activities students worked.
- How students worked (whole class, groups, independently) and their understandings or problems with an activity.

Information about the *outcome variables* included:

- Student knowledge and understanding of mathematical concepts and procedures to be gathered via three sets of instruments: (1) a set of grade-level tests developed specifically for MiC, (2) four strand tests (given only at the end of Grade 8), and (3) the norm-referenced standardized test currently used in the district.
- Student ability to use what they know to solve problems will be determined using sections of the grade-level tests developed specifically for MiC.
- Student attitudes will be assessed using an adaptation of an instrument that was used in the field-testing of MiC.
- Information about the *consequent variables*, further enrollment and transition to high school, will be gathered via interviews when students enter Grade 9.

Several changes in what was originally proposed were made during negotiations with NSF. First, the instrumentation plans needed to be simplified and instruments tailored to the study. There was too much reliance on unstructured observations and interviews. For example, rather that relying on an existing attitude inventory we needed to develop a survey targeted to the reform notions of mathematics. Thus, in the revised proposal we decided to spend the 1996-1997 school year modifying or developing instruments and scales which was pilot tested and then used. Second, rather than administering a new standardized achievement test we agreed to use data from the tests that schools were already using. Third, in the revised proposal we decided to develop the External Assessment System (EAS; Romberg & Webb, 1997-1998) using items from the National Assessment of Educational Progress (NAEP; National Center for Education Statistics, 1990, 1992) and the Third International Mathematics and Science Study (TIMSS; International Association for the Evaluation of Educational Achievement, 1996) to measure *knowledge and understanding* in the model. This was done to include information related to U. S. student performance on those tests. Finally, we decided to contract with the Freudenthal Institute to create the Problem Solving Assessment System (PSA; Dekker et al., 1997-1998) as a measure of "application" in the model for use in the study.

Proposed Sites

Originally we proposed that the information would be collected from schools in six sites: three urban school districts with an NSF-funded Urban Systemic Initiative grant; and three districts in states with an NSF-funded State Systemic Initiative grant. The sample of schools at each site was to be deliberately chosen. At each site four middle schools (with their fifth-grade feeder schools) were to be selected for the study. Two of the schools at each site were to use the MiC curriculum. Then two schools, matched with the MiC schools, were to be chosen if they planned to continue to use a traditional curriculum.

The cost of doing this work in so many schools was considered to be too high. Thus, through negotiations with NSF we reduced the longitudinal aspect of the study from four to three years. Furthermore, we reduced the number of sites from 6 to 4 (2 urban and 2 SSI sites) and the number of schools at each site from 4 to 3. In the latter case one school was to be traditional, one a new MiC school, and one a MiC field-test school.

Design for Gathering Data

The longitudinal study features a multiple-cohort prospective panel design. Data were gathered at multiple distinct periods for the same duration on the same cohorts of individuals in relation to the same variables (Menard, 1991). In the original proposal we were to begin collecting data in the 1996-1997 school year. Data were to be gathered on three cohorts of students (one cohort beginning the study in Grade 5 and followed for four years, one beginning in Grade 6 and followed for three years, and one beginning in Grade 7 and followed for two years). The overall study was to be thought of as a series of grade-level-by-year studies with classrooms as the unit of analysis. The longitudinal part of the study was to involve tracking the performance of the three cohorts of students over the four years of the study, and the cross-sectional part was to include cross-grade comparisons by year and cross-year comparisons by grade. Such comparisons were to provide a basis for answering the research questions.

Again, in the negotiations with NSF several changes were made. First, as stated the duration of the study was reduced from four to three years. Second, data was to be gathered starting in the 1997-1998 school year rather than in 1996-1997. Consequently, the revised proposal includes eight grade-level-by-year studies, three cross-grade comparisons, three cross-year comparisons, and longitudinal data for three groups of students over the three years of the study (see Figure 2-2).

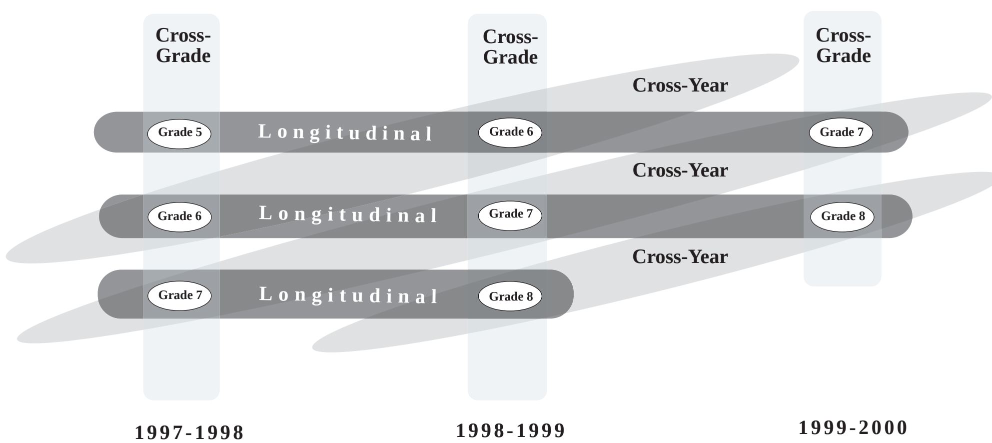

Figure 2-2. Structure of the multiple cohort study design.

Analysis Procedures

From an analytic point of view the project involves examining a large set of separate studies (17 in the revised proposal including 8 grade-level-by-year studies, 6 cross-sectional studies, and 3 longitudinal studies). To answer Questions 1 and 3 in both the original proposal and the revised proposal we expected to use structural equations to estimate the path coefficients in the complete structural model with respect to all seventeen studies. Each individual linkage implicitly represents a hypothesis that can be tested by estimating the magnitude of the relationship. However, in the revised proposal we were aware that the actual estimation of the linkages might or might not be possible, given the limited sample of classrooms at any grade level and the potential of collinearity across indices.

To answer Question 2, we planned to compare treatment groups via multivariate analysis of covariance with measures of the outcome variables as the dependent measures and prior standardized test scores as the covariate. In the revised proposal scores on EAS and PSA were to be the dependent variables for this analysis.

Because collinearity across indices posed a serious interpretation problem, a simplified model was designed during 1998-1999 when Professor Romberg was a Fellow at the Center for Advanced Study in the Behavioral Sciences at Stanford. The statistics working group at the Center, headed by Lincoln Moses, examined the indices and the analysis plan for the study and suggested that four composite variables be created. This later became five composite variables during the examination of the information gathered in the study. The simplified model describes the relationship between variations in classroom achievement (CA), aggregated by strand, or total performance can be attributed to variations in opportunity to learn with understanding (OTL_u), preceding achievement (PA), method of instruction (I), and school context (SC). This relationship can be expressed as—

$$CA = OTL_u + PA + I + SC$$

These composite indices, based on one or more sub-indices created for each variable in the original model, were then to be created. The sub-indices were intended to capture the variability across classes (and schools) in relation to each variable in the structural model. (The details of how these composite indices were created from the sub-indices are discussed in Monograph 3, and in Chapter 1 of Monograph 4)

Based on this simplified model the dependent variable used to answer each of the three questions is an index of Classroom Achievement (CA) derived from data from student responses to eight tests (four EAS and the four PSA). One of each test was given to students at each of the Grades 5, 6, 7, and 8. The CA index was derived by the research team at the Australian Council for Educational Research using “Rasch measurement models” (Masters & Forster, 1996) that allow one to combine information from different tests administered at different times onto a single scale.

To answer Question 1, three analyses were planned. First, the CA index was to be used to compare performance of all students at each grade level (the eight grade-level-by-year descriptions of performance by MiC students). Second, comparisons of performance on CA then were to be derived for the different sites, classes, teachers, gender of students, race of students, etc. Third, from the item parameters used to derive CA new indices were to be created for performances on EAS, PSA, and the items in each content strand (number, algebra, geometry, probability/statistics). Then the same types of comparisons were to be done for these sub-indices as was done for CA for each of these eight grade-level-by-year studies. Next, from these eight studies the three cross-grade and the three cross-year cross-sectional studies indicated in Figure 2-2 were to be done. These comparisons potentially should yield pictures of general student progress in performance as a consequence of instruction using MiC. Finally, for the three longitudinal studies shown in Figure 1 we planned to use changes in CA to describe the growth of students over the years.

To answer Question 2, the MiC data and similar data from students in classes using conventional instructional materials were to be used to make comparisons. Using the composite indices CA is the single dependent variable and PA the covariate to be examined via analysis of covariance. Again, the initial comparisons were to be made with respect to the eight grade-level-by-year studies, followed by the cross-grade, cross-year, and longitudinal comparisons.

To answer Question 3, the relationship between CA and the other composite indices initially were to be examined via linear regression for each of the grade-level-by-year studies followed by the cross-grade, cross-year, and longitudinal studies. Interesting relationships would then be examined in depth by analyses of the sub-indices for variables in the complete model.

It also should be noted that for the three longitudinal comparisons the unit of analysis was school, or district rather than classrooms because students do not stay in the same classes over the two or three years of the study. To carry out this analysis, we planned to classify students on some of the prior, independent, and interim indices via cluster analysis, and compare CA for the different clusters of students.

What We Expected to Learn From the Study

The purpose of this study was to document the impact on student achievement of using the reform-based middle-school mathematics curriculum, *Mathematics in Context* and to compare that student achievement with students in conventional classes. To do that we raised three research questions that we believed we would be able to answer. The importance of our answers for policy makers is critical. However, for the National Science Foundation and educational researchers we believed that the most important outcome of the study was to demonstrate the viability of the methodology we chose to use to in carrying out the study. It is nearly impossible for educational research projects that involve teachers and students (i.e., the units of treatment and analysis are classrooms) to satisfy the assumptions that underlie experimental studies. In real school situations, many other factors that might easily affect outcomes cannot be controlled. This does not mean quality research cannot be done, but rather indicates the difficulty involved in arranging student and school settings to control potential sources of possible variation. As in other social sciences one can build formal models identifying the key variables and to locate causal paths between the variables. This process of structural modeling together with the development of composite scales based on such a model has proven an effective way to carry out research in school settings and should be of interest to educational researchers and potential funding agents.

A second aspect of the study we think is important concerns the construction of composite scales from several sources of data. For example, the composite variable Classroom Achievement (assessment of classroom achievement and growth of knowledge) involves the development of a single index from the data gathered for the two outcome variables (knowledge and understanding and application) in the structural model in order to provide a frame of reference for monitoring growth. This procedure of creating CA to make operational such assessment permits a perspective on mathematical understanding not traditionally available. Portraying where the students are in the developmental corridor from informal to formal thinking and should be of interest to researchers.

A final aspect of the study of probable interest to policymakers will be the detailed information about the several sources of variation documented in this study and about attempts to isolate treatment effects from other possible determiners of outcomes in schools. In fact, in Chapter 4 we have tried to identify many of these sources, and the rationale given by participants.

Although the purpose of this study is to respond to the many calls for evidence showing that a new standards-based curriculum has an impact on student achievement, the study should also provide researchers and policymakers information about a method of validly documenting such evidence.

CHAPTER 3: INSTRUMENTATION, SAMPLING, AND OPERATIONAL PLAN

Mary C. Shafer

During the 1996–1997 school year, a plan for data collection was devised, data collection instruments were developed, and subcontracts with participating school districts were negotiated. Data collection began the following year and was conducted for three school years. In this chapter, the details of study instruments, the selection criteria and negotiation of contracts with participating districts, and the operational plan of the study are described.

Instrumentation

During the year prior to data collection, twenty-two instruments were designed to address the prior, independent, intervening, outcome, and consequent variables in the original research model. Profiles, questionnaires, the teaching log, the classroom observation instrument, and the student attitude inventory used some portions from instruments developed for prior studies of systemic change in school districts, content taught in mathematics and science classes, and student dispositions toward mathematics. Revisions of these study instruments were based on pilot-testing in nonstudy schools and comments from district site coordinators and principals who participated in a meeting in Madison in May 1997. Interview protocols and one set of assessments were developed by project staff. A second set of assessments were designed by Dutch mathematics educators from the Freudenthal Institute.

Prior Variables

Similarities and differences among students, teachers, and school contexts were identified through district and school profiles and teacher and student questionnaires. The purpose of the *District Profile* (Shafer, 1997a) was to gather demographic information about the teacher and student populations: total student enrollment in the district; the number of schools using MiC; ethnicity and gender of students and teachers; percent of students for whom English was a second language; and percent of students participating in government-funded lunch programs. In addition, more detailed information about the teachers was requested: the number of working days for teachers, the number of days provided for general professional development activities; professional development days devoted to mathematics teaching in particular; and description of the state-mandated and district-mandated requirements for preliminary teacher certification and practicing teachers as part of their continuing education. Similarly, the *School Profile* (Shafer, 1997b) gathered information about school demographics: student enrollment by grade level; the number of classes using MiC in the school; ethnicity and gender of students and teachers; percent of students for whom English was a second language; percent of students participating in government-funded lunch programs; and the number of certified mathematics teachers in the school. Information about the school mathematics program and mathematics programs in which parents participated was also gathered through the use of this instrument. District and school profiles were completed in the fall of each study year (see Shafer, 1997a, 1997b).

Two questionnaires were designed to gather information about the background of study teachers. *Teacher Questionnaire: Background and Experience* (Shafer, Davis, & Wager, 1997a) contained items related to the teacher's gender and ethnicity; educational background; experience teaching; experience teaching mathematics; and experience teaching at his/her current school. *Teacher Questionnaire: Experience Teaching Mathematics in Context* (Shafer, 1997c) collected data on the teacher's level of experience teaching MiC and the MiC units previously taught. In the *Student Questionnaire* (Shafer, 1997d), students identified their gender, ethnicity, and whether English was their primary language. Other questions gathered information about what the subjects enjoyed the most; frequency with which the student discussed mathematics with friends and relatives; things enjoyed the most and least about mathematics class; and how knowledge of mathematics helped in other classes such as science and social studies. Teacher questionnaires were completed at the beginning of each teacher's participation in the study. Student questionnaires were completed each fall.

In addition to the instruments specifically designed for the prior variables, each district provided mathematics test scores of the study students from the norm-referenced standardized tests they administered the year prior to the study and for each study year. Because the standardized tests varied by district, students completed the Collis-Romberg Mathematical Problem Solving Profiles (Collis & Romberg, 1992). Standardized test scores and profiles were used to compare students in both MiC and conventional classes in both cross-sectional and longitudinal analyses. The Collis-Romberg Mathematical Problem Solving Profiles were also used as a pre-test (Form A) and post-test (Form B) in the study. (See Chapter 1 of Monograph 2 for details on this instrument).

Independent Variables

Similarities and differences among support environments, teacher conceptions about mathematics teaching and learning, and professional opportunities for teachers were identified through interviews with principals and teachers and through two teacher questionnaires. Principal and teacher interviews were conducted in the fall of each study year. *Principal Interview: School Context* (Shafer, Davis, & Wagner, 1997b) focused on the vision for student learning in the school and for learning mathematics in particular; communication of that vision to faculty and staff; best ways for students to learn mathematics; and the types of support given to teachers. *Teacher Interview: Teaching and Learning Mathematics* (Shafer, Davis, & Wagner, 1997c) focused on teaching and learning mathematics: best ways for students to learn mathematics; support received from school administration; professional-development opportunities; and the influence of district and state mathematics frameworks and standardized testing. Teachers received an honorarium for each interview. Interviews were transcribed for analysis.

Two additional teacher questionnaires were used in the study. *Teacher Questionnaire: School Context* (Shafer, Davis, & Wagner, 1997d) gathered information specifically about the school context: the school as a workplace; the support environment of the school; beliefs about mathematics teaching and learning; assessment of student learning; and the professional-development climate in the school. *Teacher Questionnaire: Professional Opportunities* (Shafer, Davis, & Wagner, 1997e) addressed professional development opportunities for teachers: professional literature read by the teacher; within-school opportunities for professional development; types of discussions at formal within-school meetings; types of professional development activities engaged in; collaboration with other mathematics teachers; and the teacher's personal support of reform of the school mathematics program. Both questionnaires were completed at the beginning of each teacher's participation in the study and in May of each study year. Teachers in Districts 1 and 2 completed the questionnaires during the professional development institute provided by the project for study teachers during August of each study year. Teachers in Districts 3 and 4, and teachers who did not attend the professional development institutes in Districts 1 and 2, received an honorarium for completing the questionnaires upon receipt of the questionnaires at the research center.

Intervening Variables

To characterize instruction in MiC classrooms and to identify differences in instruction between classes using MiC and classes using conventional curricula, two instruments were developed for this study: a classroom observation instrument and daily teaching log which included journal questions. The *Classroom Observation Scale* (Davis, Wagner, & Shafer, 1997) was designed to measure one independent variable (curricular content and materials—the actual curriculum) and the three intervening variables: pedagogical decisions, classroom events, and student pursuits. The observation instrument is composed of seven sections. In the first section, the observer recorded pertinent information related to the teacher and students: the name of the teacher, the school, and the grade level of the students in the class. The observer also recorded information pertinent to the particular lesson: the date of the observation, times the lesson began and ended, text used, unit/chapter taught, and the page numbers taught during the lesson. In the second section of the observation instrument, the observer conducted and recorded notes from a brief pre-observation interview of the teacher during which the teacher was asked to identify the mathematical content to be explored or conveyed in the lesson and the location of the lesson with respect to the development of concepts in the instructional unit/chapter. In the third section, the observer listed lesson activities along with the time allotted to each.

The next two sections of the observation instrument were collectively composed of 12 indices for various dimensions of instruction, which addressed the three intervening variables in the research model for the study. Nine of these indices focused on classroom events; the remaining three indices focused on student pursuits. Pedagogical decisions, although not presented in a separate section of the observation instrument, were central to both classroom events and student pursuits. In general, a rating of 1 on a particular index indicated that the dimension was rarely or never seen in the lesson; the highest rating indicated that the dimension received major emphasis during the lesson. In practice, high ratings were rarely attained on every item during one observation. Ratings also varied in different observations of the same teacher. The observer also provided written evidence to support each rating.

In the sixth section of the observation instrument, the observer conducted and recorded notes from a brief postobservation interview of the teacher during which the teacher was asked to rate and comment on the degree to which he/she felt the lesson achieved the purpose noted in the pre-observation interview. The teacher was also asked whether any incidents occurred during the lesson which revealed student misunderstanding or provided opportunities to facilitated student understanding in any way. In this way, teachers had an opportunity to describe and explain modifications made during the lesson. In the final section of the observation instrument, the observer recorded any additional comments about the lesson. The observation instrument was pilot-tested by project staff numerous times both in classes using MiC and classes using conventional textbooks in order to define and clarify descriptors for each item and to determine ways to achieve interrater reliability.

The teaching log (Shafer, Davis, & Wagner, 1997f), with accompanying journal questions, was designed to record information about daily instruction in study classes. Information from the teaching logs was used in the analysis of the content of the actual curriculum, the use and modification of curricular materials, lesson planning, mathematical interaction during instruction, and classroom assessment.

The teaching log consisted of Introductory Information, Daily Logs, and Journal Questions. The purpose of the Introductory Information was to document the unit/chapter taught, changes in class rosters, information about grouping for instruction, and the physical arrangement of the classroom. Because the study was longitudinal, teachers noted names of students who were added or dropped from the class. Teachers indicated whether they intended for students to work in small groups or pairs during the teaching of the particular unit/chapter, and they described criteria for grouping students. Finally, teachers sketched the physical arrangement of the classroom. The Introductory Information was completed once a month.

One portion of the log was designed to document content taught, forms of instruction, and student activities. After noting the date, unit/chapter and pages taught on a particular day, teachers indicated if the lesson was a continuation of the previous lesson and checked whether all students in the class covered the same content. Teachers also identified instructional activities that were used during the class period (warm-up activity, review of previous material, teacher presentation of material, whole-class discussion, small-group or pair work, independent practice, or another activity specified by the teacher) and circled an emphasis code for each one that ranged from 1 (used for 15% or less of the class period) to 4 (used for more than 75% of the lesson). Similarly, teachers noted the types of activities students engaged in during the class period (listened to teacher or took notes, investigated problems, discussed answers and solution strategies, participated in whole-class discussion, practiced computation, took a quiz or test, reflected on or summarized lesson content, began homework, or another activity specified by the teacher) and circled an emphasis code for each one. In the second portion of the log, teachers provided information about supplemental materials used during that class period, their classroom assessment practice, homework assignments, and formal assessments. The journal questions were designed to document lesson content that was emphasized or modified and notable classroom events. Each journal question was accompanied by a list of suggestions for reflection. Teachers had the option of commenting on other instructional issues of importance to them.

Teachers were instructed to complete a daily log sheet for each day of instruction as soon as possible after the lesson and complete at least one set of journal entries per week for the entire school year. If teachers taught multiple mathematics classes, they were asked to complete the log for the class that was observed monthly by the on-site observer. In this way, the information gathered through the log would add the teacher's perspective on the particular lessons for which observation reports were completed, thereby adding a means of triangulating data from observations and teaching logs. Each month the teaching log was a different color for ease in documenting the receipt of teacher logs. A binder was given to each teacher at the beginning of the school year. This binder contained the one-page Introductory Information, daily log sheets and journal questions for each instructional day for one month, a pocket folder for holding supplementary resources, quizzes, and formal assessments used by the teacher during instruction, and a postage-paid envelope for sending the log to the research team.

Instructions for completing the teaching log and models of completed logs were reviewed with the teachers each August during the professional development institute sponsored by the project for study teachers. District site coordinators reviewed the instructions with teachers who were unable to attend the institutes. Subsequent teaching logs with postage-paid envelopes were sent to each teacher monthly. Logs were sent to a contact teacher at each school for distribution. The contact teacher was given an honorarium each semester for distributing all study instruments to teachers on a timely basis. Teachers received an honorarium for each teaching log upon receipt of the log at the research center. The teaching log was pilot-tested with nonstudy teachers during the spring semester prior to the study. Based on feedback from pilot teachers and input from district personnel involved in the study, the log was refined to clarify items and make the format easier for teachers to complete in five to ten minutes daily.

In addition to the classroom observation and teaching logs, *Teacher Interview: Instruction Planning and Classroom Interaction* (Shafer, Davis, & Wagner, 1998) was conducted with each study teacher in the spring of each study year that focused on instructional planning and classroom interaction, specifically planning for units or chapters and individual lessons, types of student discussion during class, and the nature of acceptable answers. Teachers using MiC were asked additional questions regarding the differences between MiC and conventional curricula when planning to teach and in class discussions and the advantages and disadvantages of teaching MiC. Interviews with teachers in Districts 1 and 2 were conducted by the on-site observer in each district. Teachers in Districts 3 and 4 were interviewed via telephone by the project director or a project assistant at the research center. Teachers received an honorarium for each interview. Interviews were transcribed for analysis. On-site observers were compensated an amount per interview as part of their subcontracts.

Outcome Variables

In order to determine the mathematical knowledge, understanding, and attitudes of students using MiC, and to identify the differences in mathematical performance and attitudes among classes using MiC and those using conventional curricula, three sets of instruments were developed for this study: the Student Attitude Inventory (SAI; Shafer, Wagner, & Davis, 1997); the External Assessment System (EAS; Romberg & Webb, 1997-1998; Chapter 1 in Monograph 4), a set of grade-specific assessments based on national and international tests of educational progress; and the Problem Solving Assessment System (PSA; Dekker et al., 1997-1998; Chapter 1 in Monograph 4), a set of grade-specific assessments that emphasized more complex problem-solving and reasoning in mathematics in addition to the basic skills.

The SAI was developed to elicit information related to student beliefs about the nature of mathematics, the role of communication in learning mathematics, and the importance of and appropriate use of technology. The belief statements were categorized into seven subscales: effort to learn mathematics; interest in and excitement about mathematics; communication of mathematical ideas; attribution for success or difficulty in mathematics; usefulness of mathematics in class and in daily living; confidence in mathematical abilities; and general perceptions of the nature of mathematics. The SAI also contained open-ended questions related to what students think of when they hear the word "mathematics;" jobs and careers that use mathematics; and ways in which mathematics is used by students. The SAI was pilot-tested with classes in which MiC or conventional textbooks were used. Study students completed the SAI at the beginning of their participation and in late April of each study year. (See Chapter 1 of Monograph 2 for details about this instrument).

The EAS is a set of grade-specific assessments that were designed to measure student performance on tasks that were used by previous national and international samples of middle school students. The items were selected primarily from National Assessment of Educational Progress (NAEP; National Center for Education Statistics, 1990, 1992) and the Third International Mathematics and Science Study (TIMSS; International Association for the Evaluation of Educational Achievement, 1996). In each assessment, the number of test items was balanced across the four strands (number, geometry, algebra, probability and statistics). In order to examine growth over time, 20 items of moderate difficulty, called anchor items, were repeated on each grade-specific assessment. The remaining non-anchor items increased in relative difficulty from fifth to eighth grade. Eighty percent of the items were multiple-choice; the remaining items were constructed response. Assessments for Grades 5–7 were pilot-tested in 1997 in classes studying MiC or conventional curricula. A second round of pilot-testing for Grades 5–7 was conducted in 1998 to assess the extent of success by students in large urban settings. The assessment for Grade 8 was pilot-tested in 1998. Study students completed the assessment appropriate for their grade level during one class period (two class periods for Grade 8) in May of each study year. The assessments were scored during scoring institutes in Madison each summer. Responses were scored with the same rubrics used by NAEP or TIMSS (see Chapter 1 in Monograph 4).

The PSA is a set of grade-level assessments that were developed by mathematics educators at the Freudenthal Institute to address three levels of reasoning across four content strands (number, geometry, algebra, probability and statistics): conceptual and procedural knowledge; making connections, finding patterns and relationships; and mathematical modeling, analysis and generalization. The PSAs contained items that all students, regardless of the curriculum they studied, should be able to solve successfully. Moreover, items were designed for accessibility on a variety of levels so that students who rely on direct modeling or drawing strategies and students who use more sophisticated strategies could solve the same problems. To the extent that a paper-and-pencil assessment can be used as an indicator of student thinking, the PSAs were designed elicit students' thinking processes in order to track their progress from informal methods to more formal symbolic notation over time. The general directions for the test and directions for individual items explicitly requested that students demonstrate their reasoning. All items were set in contexts and were constructed response. Items related to one algebra context, which became increasingly difficult, were included over the four grade-level assessments. Item-specific scoring rubrics were created to assign point values to particular responses. Full or partial credit was awarded to responses based upon accuracy of response and thoroughness of explanation. In addition, student solution strategies were coded in order to monitor changes in strategies over time. Assessments for Grades 5–7 were pilot-tested in 1997 in classes studying MiC or conventional curricula. The assessment for Grade 8 was pilot-tested in 1998. Study students completed the assessment appropriate for their grade level during two class periods in May of each study year. Scoring procedures were designed and pilot-tested in Madison during the spring semester of 1998. The assessments were scored during on-site scoring institutes with study teachers in May and in Madison each summer. (See Chapter 1 of Monograph 4 for a description of the scoring procedures and interrater reliability.)

Consequent Variable

An additional student questionnaire was designed in 1999 in order to examine similarities and differences in students' transition into high school mathematics. In January 2000, the *Transition Questionnaire* (Shafer, 1999) was sent to ninth-grade students who had participated in the study as seventh- and eighth-grade students. The questionnaire collected data about the nature of students' experiences in both middle-school and high-school mathematics classes; the types of mathematics courses students intended to study in high school; and students' conceptions about learning mathematics. Student questionnaires were distributed by counselors in the high schools attended by study participants in Districts 1–3. Questionnaires were not sent to students in District 4 because of the dispersion of students throughout numerous high schools in the district. Students were asked to complete the questionnaires under the supervision of their parents and return the questionnaires to the research center in postage-paid envelopes. The response rate was very low. Consequently, transition questionnaires were not distributed during the final year of the study.

Evaluation research of high quality can be conducted in classroom settings when methods of data collection, measurement, and statistical inference are designed to take into consideration the complexities and variation encountered in such settings. Collectively, the instruments developed for this study permitted in-depth data collection on a broad set of complex variables and provided data that were sensitive to differences in school contexts, opportunity to learn with understanding, and instruction whether using MiC, a newly developed mathematics curriculum that emphasized student reasoning, or using conventional mathematics curricula (see Table 3-1). The data collection provided a foundation for triangulating data from multiple sources and permitted the use of both quantitative and qualitative methodologies in examining the culture in which student learning was situated, the instruction students experienced, and students' opportunity to learn comprehensive mathematics content in depth and with understanding.

Table 3-1.
Data Collection Instruments Designed for the Mathematics in Context Longitudinal/Cross-Sectional Study

District Profile
Information about the superintendent
General information about schools
Characteristics of students
Characteristics of teachers
State-mandated mathematics requirements for teachers
District-mandated mathematics requirements for teachers

Teacher Questionnaire: Background and Experience
Gender, ethnicity
Educational background
Experience teaching
Experience teaching mathematics
Experience teaching at current school

Teacher Questionnaire: Experience Teaching MiC
Number of semesters, years
Units taught

Teacher Interview Protocol: Teaching and Learning Mathematics
Best ways for students to learn mathematics
Support of administration
Professional-development opportunities
Influence of district and state frameworks and testing

Teacher Interview Protocol: Instructional Planning and Classroom Interaction
Planning for units or chapters and individual lessons
Types of student discussion
Nature of acceptable answers
For MiC teachers: Difference between MiC and conventional curricula when
- planning to teach
- using class discussions

Advantages and disadvantages of teaching MiC

Student Questionnaire
Gender, ethnicity, primary language
Favorite subjects
Frequency the student talks about
- the mathematics learned in class
- homework assignments
- uses of mathematics

Three things liked about mathematics class
Three things disliked about mathematics class
Influence of knowledge of mathematics and ways to learn mathematics in other subjects

Student Attitude Inventory
Interest and excitement about mathematics
Effort
Usefulness of mathematics
Confidence in mathematical abilities
Communication of mathematical ideas
Attribution
General perceptions
Open-ended questions:
- what comes to mind when student hears “mathematics”
- jobs besides teaching that require mathematics
- personal use of mathematics outside of class
- other ways people use mathematics

Student Questionnaire: Transition into High School
Experiences in middle- and high-school mathematics classes
Number and type of high school mathematics courses
Conceptions about learning mathematics

School Profile
Information about the principal
General information about the school
Characteristics of the students
Characteristics of the teachers
Information about the school mathematics program
Mathematics programs in which parents participate

Teacher Questionnaire: Professional Opportunities
Professional literature
Within-school opportunities for professional development
Types of discussions at formal within-school meetings
Types of professional-development activities engaged in
Collaboration with other mathematics teachers
Personal support of reform in mathematics program

Teacher Questionnaire: Support Environment
School as a workplace
Support environment
Mathematics teaching and learning
Assessment of student learning
Professional-development climate

Principal Interview Protocol: School Context
Vision for learning in the school
Vision for learning mathematics
Communication of vision to staff
Best way for students to learn mathematics
Support of teachers

Classroom Observation Instrument
Placement of current lesson within unit
Teacher’s goals for the lesson and whether they were met
Instruction formats used and time appropriated to each
Classroom events, pedagogical decisions, pupil pursuits

Teaching Log
Pages taught
Whether students studied the same content
Instructional formats used
Physical classroom changes needed to accommodate activities
Student activities used
Additional materials used during the lesson
Type of informal assessment used if any
Type of information gathered during informal assessment
Effect information gathered during informal assessment had on instruction
Type of homework assignment
Type of formal assessment used if any
Journal questions
- emphasis given during instruction or modifications made in lessons
- notable classroom events

Problem-Solving Assessment System
Grade-specific assessments
Constructed-response items
Concepts in algebra, geometry, number, probability and statistics
Three levels of reasoning
Scoring rubrics and strategy codes

External Assessment System
Use of NAEP and TIMSS items
Balance of items across four content strands
Set of anchor problems on each grade-specific assessment

Sampling

Selection Criteria for Districts

The process of finding research sites began in the fall of 1996. Several criteria were used in the selection of school districts for participation in the study. First, NSF requested that participating districts were involved in systemic initiatives for reforming mathematics curriculum and instruction. Districts involved in Urban Systemic Initiatives (USI) or State Systemic Initiative (SSI) and university faculty who participated in mathematics education reform efforts across the United States were contacted. Also, school districts that had participated in the pilot- and field-test phases of the development of MiC were contacted about participation in the study. As a result, four systemic initiatives, five university faculty, and eight pilot- and field-test sites were contacted. Negotiations were actively pursued with nine school districts.

A second set of selection criteria was used to ensure a variety of research conditions. Considerations were given to the size, location (rural, suburban, and urban), and demographics of potential districts; number of mathematics classes taught by fifth-grade and middle-school teachers; typical class size and length of class periods; mobility rate of students; and availability of mathematics books for student use.

Because the initial research design of the study included teachers who had various levels of experience teaching MiC, a third set of criteria for screening potential districts involved the types of experiences teachers had using MiC during years prior to the study. Specifically, we sought districts in which teachers had previously used MiC or were going to use MiC for the first time during the study. Information was gathered on the extent of teachers' implementation, such as whether they used MiC field-test or prepublication units as replacement units and the number and titles of field-test or prepublication units previously taught. The possibility of some teachers gaining experience teaching two or three MiC units in the spring prior to the study was also explored. Subsequently, a workshop was conducted in one district in early May 1997 focusing on teaching one MiC unit before the end of the school year and a discussion of implementation issues such as teaching MiC in block schedules, integration of MiC units into established problem-solving classes, and parent involvement sessions.

Finally, districts were selected on the basis of their willingness to participate in comparative design of the initial research model. The project planned to involve six districts. In each district, two-thirds of the study teachers and students were to use MiC, and the remaining teachers and students were to use whatever conventional textbooks were already in place in the district. Three issues arose relative to the comparative research design. First, the major issue encountered in discussions with districts was their unwillingness to conduct a comparative study with students. Many districts expressed the concern that they might be seen as withholding opportunities for learning mathematics in more powerful ways from students who continued to study conventional mathematics curricula. Also complicating the search for research sites was the request that selected districts were involved in systemic initiatives. Control schools were difficult to find because principals did not want to be perceived as disregarding efforts to reform mathematics curriculum and instruction. In light of these concerns, the research design was modified to exclude the comparative element. In the modified research design, MiC was the only curriculum used by teachers and students. Data were gathered only on prior, independent, outcome and consequent variables, thereby excluding data collection on classroom interaction.

A second issue related to selecting districts willing to participate in the initial research design became apparent in other districts. Many elementary schools fed into each middle school, making it difficult to select the elementary schools that might yield the number of students potentially still participating in the study in the second and third years of data collection. In some cases, districts were unable to identify feeder patterns in which administrators from both elementary and middle schools were interested in participating in the study. The ability to follow fifth-grade students longitudinally was severely compromised in these districts. By dropping the requirement of a fifth-grade cohort of students, discussions were opened with two additional districts. Negotiations with one site were finalized in time for participation in study. The other site delayed the decision to sign negotiated subcontracts, and the offer was rescinded in September 1997.

Another potential research site presented a unique situation. Teachers from a school in Puerto Rico had expressed great interest in participating in the study. At that time, however, Spanish versions of commercial MiC units were unavailable. Although efforts by researchers in Puerto Rico to translate fifth-grade units into Spanish had already begun, translation of sixth- and seventh-grade units was scheduled for the 1997–1998 school year. In light of these difficulties, an alternative research plan was designed for following one cohort of fifth-grade students beginning in the fall of 1997 through the seventh grade. Contract negotiations with this school, however, were not finalized in time for participation in the study during the fall of 1997, and further negotiations were not pursued.

Negotiations with Districts

After initial interviews, two districts were identified as strong candidates for participating in the study. A sample subcontract between each district and the University of Wisconsin–Madison was sent to each of these districts. The scope of work for district administrators, principals, teachers, students, on-site observers, and project staff was thoroughly described as part of the subcontract.

In March 1997, meetings were held with key district administrative personnel in each district to discuss further details about study participation. During these meetings, discussions included: the rationale for MiC; the involvement of Dutch mathematics educators in the development of MiC; importance of conducting longitudinal research; evidence of student achievement from MiC field-test sites; details of the study; and the subcontract. Additional issues were discussed on the basis of prior conversations with both districts: districts' desired involvement of more teachers than had been requested by the project staff, particularly during the proposed summer professional development institutes; and potential study teachers using conventional curricula who might want to teach MiC. Details about the cost of MiC materials were shared. Through an agreement with Encyclopaedia Britannica, special prices were provided to study districts for teacher guides, student books, and ancillary materials used in study classes. Both districts also discussed teacher compensation, which was later decided to be direct payment to teachers for interviews and participation in summer professional development institutes and direct payment to the district for substitutes needed to release teachers for the spring professional development institutes. (See section of professional development later in this chapter for details about these institutes.)

During the March meeting with the first district, administrators asked how the study would benefit their students and questions related to heterogeneous student grouping (e.g., Was MiC challenging enough or could modifications be made for bright students? Were ancillary materials available to help students with lower abilities?). They also talked about the recent state mandate for all eighth-grade students to take algebra and asked how they might approach placement of students in high school algebra given that algebra was the primary focus of many MiC units, particularly at the eighth-grade level. These administrators felt that a parent letter prepared by the district (in addition to the parental consent letter from the study) should include what they deemed to be three major benefits of student participation in the study: use of a mathematics curriculum that addressed state and national standards in mathematics; inclusion in a study funded by NSF; and use of a mathematics curriculum published by Encyclopaedia Britannica.

During the March meeting with the second district, we were informed that a new system of standardized testing would be in place during all years of the study, which included the Stanford 9 (Harcourt Educational Measurement, 1996) at each grade level and a newly developed performance-based assessment for fifth- and eighth-grade students. The baseline administration of the performance-based assessment was scheduled to take place during the spring semester of the first year of data collection. The confidential nature of student standardized test score data was discussed, and the district data administrator sought permission for use of student data reported by groups of students. In subsequent conversations with the second district, the teachers' union asked whether teacher education credits for certificate renewal would be available for the proposed summer professional development institutes (an item that had to be presented by district administrators to the school board) and whether a union representative who was a mathematics teacher could participate in future study meetings (a request that was accommodated). District administrators also informed us of the necessity to have the school board approve the list of schools they selected for participation in the study.

In May 1997, these two districts had agreed to participate in the comparative research design. In these districts, data were gathered on all research variables, which included data from classroom observations and daily teaching logs. A retired mathematics teacher in each district was selected with district input to conduct classroom observations and principal and teacher interviews. Separate subcontracts were negotiated with each on-site observer.

In May 1997, a meeting was held in Madison for district site coordinators and on-site observers. Three uncommitted school districts also attended this meeting. The purpose of the meeting was to provide opportunities for personnel in selected and potential districts to meet and to provide suggestions for final forms of study instruments. After the meeting, two districts decided to participate in the study in a modified research design in which MiC was the only mathematics curriculum used and data on classroom interaction were not gathered. By the fall of 1997, four school districts accepted offers to participate in the study: Districts 1 and 2 participated in the comparative research design; Districts 3 and 4 participated in the modified research design. In District 4, only middle-school classes participated in the study because of the large number of elementary schools that fed into the selected middle school and the dispersion of incoming fifth-grade students among sixth-grade mathematics classes.

District Description

District 1 is located in an urban region in the eastern part of the country. The district had a 45% minority student population with 30% African American students and 12% Hispanic students. Approximately 30–40% of the students in the district were eligible for government-funded lunch programs. Professional development to acquaint teachers of mathematics with reform-based curricula was offered in District 1, and monthly meetings were provided for teachers who were implementing such programs. For preliminary teacher certification, 24 credit hours were recommended for fifth- and sixth-grade teachers; 24 credit hours were required for seventh- and eighth-grade teachers. No specific mathematics requirements were necessary as part of continuing education. District requirements were the same as the state requirements.

District 2, located in a large urban area in the southeastern United States, contained 251 elementary and middle schools and numerous high schools. The district student population was predominantly minority, with 33% African American students and 52% Hispanic students. Over 50% of the students in the district were eligible for government-funded lunch programs. District 2 provided numerous possibilities for professional development. Each school was given six early-release days for general professional development. In addition, each school received 10 substitute days for professional development in mathematics and/or science, 12–18 days of in-service days in mathematics provided by (USI or Eisenhower) government funding (each involving 2–6 teachers), and 3–5 days of districtwide mathematics in-service. Teachers also had opportunities to participate in five days of paid in-service for mathematics during the summer. District requirements for preparation of mathematics teachers were the same as state requirements. For elementary teachers, preliminary teacher requirements mandated the study of arithmetic for the elementary school. For middle-grade mathematics certification (Grades 5–9), 18 semester hours in mathematics were required; for certification in mathematics for Grades 6–12, 30 semester hours in mathematics were required. Continuing certification required the completion of six semester hours in mathematics or 120 district in-service credits in mathematics every 5 years.

District 3 is located in a suburban area of a large western state and is composed of four schools, each specializing in 3–4 grade levels. Study participants included all fifth- through seventh-grade mathematics classes in the district. The six self-contained fifth-grade study classes were in a school for Grades 3–5; Grades 6-8 were in a middle school. One sixth-grade class was self-contained. All other middle-school classes had several subject-matter teachers. The district student population was predominately White. School administrators provided paid monthly meetings for mathematics teachers who were implementing reform curricula for the first time, and teachers often met weekly without pay to prepare mathematics lessons. For preliminary teacher certification, the state mandated single-subject credentials for Grades 7–8. K–6 teachers were required to complete a multiple-subjects credential including several mathematics courses. Although mathematics courses and staff development were provided by the district, the district did not require additional certification and courses for experienced teachers.

District 4 is one of many districts located in a large urban area in the eastern part of the country. Grades 6–8 are contained in middle schools in which students have several subject-matter teachers. Study participants were from one middle school in this district. The district student population was predominantly minority with 50% African American and 37% Hispanic students. Over 50% of the students are eligible for government-funded lunch programs. For new mathematics teachers, 36 credits in mathematics or a mathematics major were required by the state, but no specific mathematics requirements were necessary as part of continuing education. District requirements were the same as the state requirements. Professional development opportunities were provided to all mathematics teachers at both district and school levels, including personalized discussions with the assistant principal for mathematics instruction. These discussions focused on reform recommendations in curriculum, instruction, and assessment; research in mathematics education; and applications of research in classroom practice.

Selection of Schools and Teachers

District administrators and on-site coordinators were asked to select schools that were representative of the district population, rather than selecting schools with extremely low- or high-achieving student populations. Principals of the selected schools chose the study teachers. Teachers, in turn, were asked to select classes of students with average mathematical abilities rather than classes of low ability or classes in honors programs. Teachers who taught more than one class selected the classes of students. Parental consent was requested for all students in study classes.

In the first year of data collection, the sample consisted of 53 fifth-, sixth- and seventh-grade teachers who taught a combined total of 94 classes involving 2222 students in 17 diverse schools in Districts 1–4. The numbers of girls and boys was about the same. Nearly half of study students were children of color, and 10% of the students identified English as their second language. Teachers in Districts 1 and 2 used either MiC (commercial version, first available in summer 1997) or the conventional curricula already in place in their schools. Teachers in Districts 3 and 4 used MiC. (See Monograph 2 for detailed information on teacher and student samples.)

During the spring semester of 1998, teachers in one middle school in District 2 decided to withdraw from the study. Agreements were negotiated to allow participating students to complete study assessments at the same time as other study students in May. Because the teachers withdrew from the study, however, we lost the ability to follow their students and students from the corresponding elementary feeder school during the second and third years of data collection. The district selected another feeder pattern as a replacement, and additional student cohorts were initiated in the fall of 1998: two fifth-grade classes, four sixth-grade classes, and four seventh-grade classes. Also, in the fall of 1998, participating students in all districts were identified in new study classes. Additional students in new study classes were also invited to participate in the study, and all study instruments related to student background were administered to new students whose parents granted consent. During the 1998-1999 school year, the sample included 48 sixth-, and seventh-, and eighth-grade teachers who taught a combined total of 115 classes involving 2351 students in 10 diverse schools in Districts 1–4. The numbers of girls and boys was about the same. Sixty-five percent of study students were children of color, and 16% of the students identified English as their second language.

During the fall of 1999, it became apparent that the new cohort of fifth-grade students in District 2 were dispersed among both study and nonstudy middle-school classes, leaving less than five study students in each class. A decision was therefore made not to follow these students for the remainder of the study. In contrast to data collection in the previous year, new students in study classes in all districts were not invited to participate in the study. During the 1999-2000 school year, the sample was composed of 27 seventh- and eighth-grade teachers who taught a combined total of 67 classes involving 815 students in 8 diverse schools in Districts 1–4. The numbers of girls and boys was about the same. Fifty-five percent of study students were children of color, and 11% of the students identified English as their second language. The great difference in the numbers of study students during the final year of data collection is attributable to several trends. First, many students in Districts 1, 2, and 4 were lost due to dispersion of study students among nonstudy classes. Second, in District 1, parents had the option every year of enrolling their children in different schools. Consequently, study students became enrolled in schools outside the selected feeder patterns. Third, fewer teachers who taught conventional curricula were willing to participate in the study. As a result, the ability to follow study students who used conventional curricula was compromised. Finally, all four districts had instituted initiatives for students to study algebra in the eighth-grade. Numerous study students, particularly in Districts 3 and 4, met established criteria for studying algebra and were consequently no longer followed in the study. (See Chapter 4 in this monograph for details that explore mitigating factors in conducting research in actual school situations.)

The Operational Plan

The operational plan of the study included professional development institutes for study teachers, data collection, interrater reliability between the observers, ongoing development of study assessments, design of the project database, and the analysis strategy.

Professional Development

As part of the project's ongoing work with study teachers, three types of professional development were provided. First, four-day on-site summer institutes were held in Districts 1 and 2 during the August before each year of data collection. During the summer institutes in 1997, the first day was devoted to information about the study and study instruments. For the next 2½ days, separate workshops were held with teachers who used conventional curricula and teachers who were going to use MiC. Topics for teachers who used conventional curricula included alignment of curriculum, instruction, and assessment in relation to the state mathematics framework; assessment of student understanding using nonroutine tasks and student responses to such tasks; classroom assessment practice; intersections among scoring, evaluation, feedback, and grading; multiple-day and long-term assessment projects; and the creation of multi-faceted assessment tasks. MiC teachers experienced in-depth reviews of the MiC philosophy, conceptual development in each content strand, teacher support materials, and homework ideas. For all teachers, the final afternoon of the institutes was devoted to study timelines, scheduling observations, and review of teacher compensation.

During summer institutes in 1998, all teachers reviewed some of the work from student assessments scored in the previous year. Two concurrent institutes were then held. Teachers who used conventional curricula discussed authentic instruction, tasks, and assessments and research on teaching and learning mathematics with understanding. Grade-specific teachers using MiC traced development of particular content for each strand. This was followed by teacher presentations of findings for their grade level to the larger group. Teachers also discussed parent materials, homework, and classroom issues. On the final afternoon, classroom observations and teaching logs were reviewed with all teachers.

In preparation for the summer institutes in 1999, teachers were asked about the content they wanted to pursue and the institutes were planned around those themes. For teachers using MiC, guidance was provided for preparing to teach units that were unfamiliar to them. This involved solving unit problems, determining multiple strategies for problems, selection of problems appropriate for classroom assessment and for homework, and discussion of end-of-unit assessment tasks. Discussions for other teachers who used MiC or conventional curricula focused on classroom assessment. In particular, teachers explored what to assess and how to assess; assessments that attended to the three levels in the Dutch assessment model; grading such assessments; designing assessment tasks; and developing scoring rubrics for those tasks. Thus, the character of the summer institutes changed over time in response to teachers' particular interests and needs.

Another professional development opportunity for teachers was scoring study assessments. Institutes were held in May (or June) of each study year in Districts 1, 2, and 4, in May of the first two years of study participation in District 3, and at the site that pilot-tested eighth-grade assessments during the summer of 1998. The focus of these two-day scoring institutes was for teachers, on-site coordinators, on-site observers, and principals to learn about the construction and scoring of study problem solving assessments. Student assessments that were not scored at these institutes were scored during scoring institutes in Madison.

Finally, scoring institutes were also held in Madison during the summer and fall of 1998 and the summers of 1999 and 2000. Study teachers were invited to travel to Madison for scoring assessments for one of the scoring institutes in Madison in 1998; one teacher from each of Districts 1, 2, and 3 participated. The remaining raters at the Madison scoring institutes were elementary- and middle-school teachers from Madison-area schools. Lists of potential raters were collated from input by local district mathematics coordinators, local MiC pilot- and field-test teachers, and teachers who were participating in other projects at the Wisconsin Center for Education Research. The scoring institutes provided a significant professional development opportunity for teacher-raters who commented that they made changes in their pedagogy to emphasize mathematical communication, included lessons that promoted more complex reasoning, and integrated various types of problems designed to elicit student thinking at more complex levels in their classroom assessment practice. (See Chapter 1 of Monograph 4 for detailed information about these scoring institutes and studies of interrater reliability.)

Data Collection

Data collection for teachers and students began in summer 1997 (see Table 3-2). Throughout the summer, lists of schools and teachers were finalized by participating districts. Teacher consent forms were mailed with information about the study, teacher compensation, and teacher responsibilities in the study. School districts were also asked to send the standardized test scores for study

Table 3-2.
Involvement of Teachers and Students, by Grade Level and Year

	Grade 5	Grade 6	Grade 7	Grade 8
1997–1998	**Teachers** Summer institute; surveys; daily log; two interviews; observations by on-site person; administration of student surveys and assessments; scoring end-of-grade assessments **Students** Surveys; Collis-Romberg Profiles; two end-of-grade assessments; student work selected by teachers	**Teachers** Summer institute; surveys; daily log; two interviews; observations by on-site person; administration of student surveys and assessments; scoring end-of-grade assessments **Students** Surveys; Collis-Romberg Profiles; two end-of-grade assessments; student work selected by teachers	**Teachers** Summer institute; surveys; daily log; two interviews; observations by on-site person; administration of student surveys and assessments; scoring end-of-grade assessments **Students** Surveys; Collis-Romberg Profiles; two end-of-grade assessments; student work selected by teachers	—
1998–1999	**Teachers** Summer institute; surveys; daily log; two interviews; observations by on-site person; administration of student surveys and assessments; scoring end-of-grade assessments **Students** Surveys; Collis-Romberg Profiles; two end-of-grade assessments; student work selected by teachers	**Teachers** Surveys; daily log; two interviews; observations by on-site person; administration of student surveys and assessments; scoring end-of-grade assessments **Students** Surveys; two end-of-grade assessments; student work selected by teachers	**Teachers** Surveys; daily log; two interviews; observations by on-site person; administration of student surveys and assessments; scoring end-of-grade assessments **Students** Surveys; two end-of-grade assessments; student work selected by teachers	**Teachers** Summer institute; surveys; daily log; two interviews; observations by on-site person; administration of student surveys and assessments; scoring end-of-grade assessments **Students** Surveys; two end-of-grade assessments; student work selected by teachers; survey in Grade 9
1999–2000	—	—	**Teachers** Surveys; daily log; two interviews; observations by on-site person; administration of student surveys and assessments; scoring end-of-grade assessments **Students** Surveys; two end-of-grade assessments; student work selected by teachers	**Teachers** Surveys; daily log; two interviews; observations by on-site person; administration of student surveys and assessments; scoring end-of-grade assessments **Students** Surveys; two end-of-grade assessments; student work selected by teachers

classes. In August 1997, professional development institutes were conducted in Districts 1 and 2. As part of the institute teachers completed all four teacher questionnaires, the full set of teacher consent forms was collected, and initial teacher logs and class sets of parental consent forms and student questionnaires were distributed. In September, the set of four teacher questionnaires was sent to teachers in Districts 3 and 4 and to teachers in Districts 1 and 2 who were unable to attend the summer institute. Parental consent letters and student questionnaires were sent to the contact teacher at each school, who then distributed them to study teachers. Student attitude inventories and the Collis-Romberg Mathematical Problem Solving Profiles (Form A) were sent to all four districts. Each class set of materials was packaged in a postage-paid envelope for mailing completed forms back to the research center. As materials were received, they were logged in and entered into the study database, which is described in this section. As students were added to classes, a complete set of student materials was sent to the teacher for new students to complete. In this way, students were added to the study during the first and second years of data collection. Profiles were sent to each district and school to gather demographic and teacher-related information.

In September and October, principals and teachers were interviewed about their views of teaching and learning mathematics and about the school context. Teacher logs were sent to teachers in Districts 1 and 2 monthly. Each was packaged with a postage-paid envelope for returning to the research center. Initial classroom observations were completed in September in District 1 and in October for District 2. Fewer observations were conducted in District 2 due to differences in school schedules, procedures for assigning students to classes, and preparation for district and state standardized testing. Completed observation reports were sent to the research center electronically every month.

In the spring semester of each year of data collection, teachers were interviewed regarding instructional planning and classroom interaction. Two teacher questionnaires (related to school context and professional opportunities) were completed during May of every school year in which the teacher participated in the study. Standardized test scores were requested from the district each spring. Also, every spring, students completed the student attitude inventory and grade-specific External and Problem Solving assessments. In the final spring of each student's participation, the Collis-Romberg Mathematical Problem Solving Profiles (Form B) were also administered.

Data collection in the second and third years of data collection followed the same operational plan with one additional element. In January 2000, a final student transition questionnaire was sent to ninth-grade students who had participated in the study as seventh- and eighth-grade students. These questionnaires were distributed by counselors in the high schools attended by study participants in Districts 1–3. Because the response rate was very low, these questionnaires were not distributed during the fall semester 2000.

Observers' Interrater Reliability

In the August prior to the study, each observer viewed two videotaped lessons with a graduate project assistant who developed the observation instrument and rated the lessons using the instrument. (See Davis, Wagner, & Shafer [1997] for details.) During these meetings discussions centered on consistency of ratings and descriptions of the types of conjectures observed, the nature of student–

student conversation, and instances in which teachers used student inquiries to shape the lesson. In the fall of 1997, the observer at each site and a project assistant visited five classes in District 1 and nine classes in District 2. During the first few observations at each site, the project assistant's and observer's ratings of several items differed by one point. By the last observation, however, this disagreement had subsided considerably. After this training, the observers began observing each study teacher once a month and completed a report for each observation. Completed reports were sent electronically to the research center for analysis. Each observer was compensated an amount per observation as part of a subcontract between the observer and the University of Wisconsin. The amount of payment varied according to the length of the class period observed.

Assessment

During the fall semester 1997 and early spring semester 1998, the Problem Solving Assessment System (PSA) for Grades 5–7 were revised, and scoring rubrics and strategy codes were finalized. Scoring procedures were designed and pilot-tested for both sets of study assessments. During the spring semester 1998, the External Assessment System (EAS) and the Problem Solving Assessment System (PSA) was pilot-tested for a second time in large urban settings, and the eighth-grade EAS and PSA were pilot-tested in May 1998. The PSA was revised during the 1998-99 school year, during which time the scoring rubrics and strategy codes were finalized.

The Research Database

In the fall of 1997, a database was designed to store data generated from the study and allowed many consumers to interact with the database, simultaneously when necessary (Arauco, 1998). The levels of consumers interacting with the database included neophytes and experts in research and analysis of issues in mathematics education either pursuing broad goals of investigation or having a more narrow focus of inquiry. To accommodate such a diverse group of consumers with varying database navigational skills, a modular and extensible database structure and organization was necessary. The modularity of the database was defined by the ease in which additional "modules" or blocks of information can be appended, omitted, or deleted without adversely affecting the integrity of the overall database structure. The modular format of the database enhanced its extensibility. Specifically, modules of information recorded in multiple software applications were added to this database design enhancing its functionality. For example, accurate distribution of specific assessments to study participants was achieved by printing adhesive labels with critical information such as student name, student identification number, school name, and teacher name. This process was quickly and efficiently made possible by a software application allowing the custom design of adhesive labels to our specifications by gathering the critical data directly from the database.

The database is considered relational because columns within the tables or queries are related to one another in one or more ways. The columns not only stored alphanumeric data, but also stored individual words, sentences, and paragraphs. The dynamic design of the database was conducive to the digital modification and implementation of the research model that undergirds the longitudinal/cross-sectional study. The developing ideas from various researchers such as study designers, part and full-time researchers, and their assistants proved a refining force in the design of the database. Data were entered, stored in tables, compiled, and warehoused in an Access database for further processing. The raw data were then applied to the research model for thematic and semblance identification. Upon the request of a specific consumer, the database developer generated a database query subsequently sorting and filtering data applicable to the specific consumer's request. Tables were shown with fields, records, and subsequent compiled data. Both researcher and developer strategized further query design accordingly, repeating the process as new research model linkages and relationships arose. The structure and organization of the study, in conjunction with the input of consumers, were the impetus of the database design.

Each district, school, teacher, and student was assigned a unique number. Tables containing data from each instrument were coded by year and by item. Schools and teachers were also assigned pseudonyms for use in qualitative analyses. Each table in the database had at least one primary key assigned to it, and information in the database was not duplicated. In other words, unique information can be found in one table only. For example, the student information such as student identification number, last name, first name, gender, and ethnicity was stored in the student table. A view or query was used to extract and combine information needed for one or more existing tables. Each student's unique identification number was the key linkage between all tables and queries in the database. Data from completed student instruments were entered and catalogued by the student identification number in every table. The same procedure was used to compile the baseline information regarding all variables in the research model.

Data Analysis

Initially, data analysis was conducted in reference to each variable in the research model. As described in Chapter 2, because collinearity across variables posed a serious interpretation problem, a simplified model was designed as a result of discussions when Professor Romberg was a Fellow at the Center for Advanced Study in Behavioral Sciences at Stanford. In the simplified model, variation in classroom achievement can be attributed to variations in opportunity to learn with understanding, preceding achievement, method of instruction, and school context. Analyses of data gathered for each variable in the initial research model, therefore, were reconceptualized in light of the simplified research model. Each composite variable in the simplified model was analyzed with respect to numerous indices created by project staff, which were intended to capture the variation across classes and schools. The analyses of each composite variable are described in detail in Monograph 3.

Conclusion

In this chapter, the details of study instruments, the selection criteria and negotiation of contracts with participating districts, and the operational plan of the study were described. Although the operational plan was clearly set out in district contracts and in day-to-day operations, changes to the plan occurred for a variety of reasons that arose from conducting quasi-experimental research in actual school conditions with a complex set of personalities, institutional policies, and school contexts. In the next chapter, we describe the extenuating factors and challenges faced in conducting this research in actual school situations.

CHAPTER 4: CONDUCT OF THE STUDY

Mary C. Shafer

Our work with school districts was both rewarding and challenging. Teachers appreciated our continuing efforts to support their teaching of mathematics to diverse students, and we were able to support requests from on-site coordinators for meetings with principals, teachers, and parent or community organizations. Although an assortment of challenges developed during data collection, we feel that our extensive attempts to collect quality data were successful despite the challenges. In this chapter, we describe what we had planned to happen and what actually occurred as we conducted comparative longitudinal research in the reality of everyday school life.

School District Context

In school districts, institutional policies, school contexts, and personalities combined to create unique situations for the research team. As outlined in Chapter 3, districts were selected on the basis of systemic initiatives for reforming mathematics curriculum and instruction, various research conditions, and willingness to participate in a study of this nature. We negotiated a subcontract with each district that outlined the work involved in the study. The significance of these agreements was in the detailed listing of responsibilities for all study participants (including the research staff), summary of the longitudinal data collection, and consent forms for both teachers and students. Renewed each year of data collection, the agreements served as resources for the district administrators to review the scope of work negotiated between the research team and the district.

Selection of Study Schools

District administrators selected study schools. Because one of the criteria for selecting districts was that they were involved in systemic reform initiatives, we anticipated that principals of MiC schools[2] would have reform-oriented agendas. Most principals of MiC schools were supportive of efforts to reform mathematics curriculum and instruction. For example, the assistant principal for mathematics and science in District 4 stated:

> We want children to be problem solvers, inference makers, thinkers, and doers, and be able to access knowledge as we move into the new century, as we go into a technological age. . . Because of that, our vision is to develop in our students this inherent ability to strategize and look at math and numbers in a way that will make more sense to them so they have a clear understanding. . . . Students learn best by sharing their ideas, by doing, manipulating, and seeing why math is being used in a certain situation. I don't think that today we really want math to be taught in a procedural way. I

---

[2] Schools in which *Mathematics in Context* was used by study participants are referred to as MiC schools. Study teachers and students who used MiC are referred to as MiC teachers and MiC students, respectively.

> think we want math to be taught in an understanding, problematic, conceptual way. . . . Problem solving is central. As we tackle our world, regardless of where we are, we need to be problem solvers. . . . If we give students strategies, then those strategies can be carried over so they can have lifelong skills that will enable their lives to be much fuller. (Principal, Lederman, Interview 10/8/98)[3]

Other principals also emphasized the need for a broader view of mathematical competencies than just basic skills, as described by a principal in District 2:

> Mathematical competencies are one of the essential competencies for a fulfilling life in the culture we live in. Math as a problem solving methodology is the underlying vision of mathematics instruction. In part, the fundamental competencies of math computation are important so students can balance their checkbook correctly. But beneath that the concept of mathematics is a way of thinking and a problem solving tool. . . . I think it is an ineffective paradigm to see basic skills first, problem solving second. They overlap. You're not going to be able to solve problems efficiently without the basic skills as tools. On the other hand, there's no particular purpose for mastering the basic skills unless you're trying to use them to solve problems as you go. They go hand in hand. They are simultaneous and not sequenced. (Principal, Burnett, Interview 10/7/99)

Other principals emphasized that change was warranted because students lacked the ability to apply what they learned in mathematics, as noted by the principal in District 3:

> We've had a real problem in the district's history of having students who have memorized things and didn't really have any understanding of how to use them. They left our program unable to apply their learning. We're trying to eliminate that problem. . . . Our feeling is that if we can present concepts and lay out situations that guide students to come up with algorithms on their own, they hang onto the algorithms much better than if we present algorithms early and then teach students how to apply them. . . . I don't think there's any purpose in learning basic skills if you're not applying those to problem solving. . . . We try to present problems that require basic skills to solve problems, and then in the context of solving the problems, we make sure that students have mastered their basic skills. (Principal, Adler, Interview 9/29/99)

Change in mathematics curriculum and instruction was also motivated by attention to performance standards, as described by a principal in District 1:

> Right now the overriding factor is the state and district push for performance standards. And I think performance is probably the key word: What can students do with the skills and the knowledge that they have? How do they apply that to solve problems? And I think that's the shift that at least we're trying to make by going from teaching skills to taking it a step further and finding ways for kids to apply those skills to solve problems, to work on real life kind of situations that are meaningful to them. (Principal, Cohen, Interview 10/28/99)

---

[3] Names of schools, principals, teachers, observers, and students are pseudonyms.

In contrast to principals with reform-oriented agendas, two principals of MiC schools were not convinced that mathematics education reform should wholeheartedly be supported. Their interpretations of state standards led to different approaches for reaching expectations for schools to show improved student achievement. For example, a principal in District 2 expressed reluctance for his teachers to use MiC and strongly encouraged teachers to withdraw from participation in the study. District funding for his school was directly tied to student results on the SAT9 (Harcourt Educational Measurement, 1996), not the state assessment that emphasized performance standards, and he perceived that students would not perform well on the SAT9 if they studied MiC (Principal, Stone, Personal communication 2/27/98). In another situation, the principal of a MiC school in District 1 began to use a program that consisted of portions of MiC units and chapters from conventional texts. The perception was that the combination would better address both SAT9 and state performance indicators. As a result, our offer for this school to participate in the study was rescinded.

District selection of schools using conventional curricula[4] was based on the willingness of principals to participate as control sites in an era of reform. Principals did not want to be perceived as minimizing efforts to reform mathematics curriculum and instruction, and there was no guarantee that these principals did not have reform-oriented goals, as a District 1 principal stated:

> I am hopeful; I guess it would be fair to state at this point in September, that we won't be a particularly good control group because I really do not want to wait another year for some of the changes that we need to make. So, hopefully we won't be a good control group if what happens is what I intend to happen. (Principal, Darrow, Interview 9/16/97)

Furthermore, principals questioned incentives for their teachers to participate as teachers using conventional curricula. They did not perceive as adequate compensation the professional development opportunities for teachers provided by the study. In Districts 1 and 2, in which the comparative study took place, schools in higher socioeconomic locations were selected as sites using conventional curricula, rather than schools that were representative of the district populations. Teachers expressed negative opinions about participating in groups using conventional curricula, which is poignantly reflected in comments from a District 1 teacher, "We are the lab rats" (Fifth grade, conventional curriculum, Fulton, Personal communication 8/11/97). In District 1, some teachers using conventional curricula who taught multiple classes wanted to be perceived as reform-oriented. As a result, they used conventional curricula with study classes and MiC with other classes. Also, in the third year of data collection, seventh-grade conventional teacher St. James was asked by the district coordinator to teach one class using MiC in order to follow students using MiC. These factors presented challenges in obtaining and maintaining both MiC and conventional groups.

---

[4] In Grade 5, students using conventional curricula studied Harcourt Brace *Mathematics Plus* (District 1) and Scott Foresman/Addison Wesley, *Mathematics* (District 2). One teacher in District 1 used materials from a variety of resources in lieu of a textbook. In Grades 6 and 7, students used Collins et al., *Mathematics: Applications and Connections* (District 1) and Charles et al, *Middle School Mathematics* (District 2). In Grade 8, both districts used Price, Roth, & Leschensky, *Merrill pre-algebra: A transition to algebra.*. As the curricula already used in the selected schools, they collectively formed the comparison curricula, which in this paper are referred to as conventional curricula.

On-Site District Coordinators

In each district, a coordinator was selected by the district administration to work with the research staff. In Districts 1 and 3, the on-site coordinators were the contact persons we worked with during initial discussions about the study. In District 1, the coordinator was the district mathematics specialist. She worked closely with the district director of curriculum and the district superintendent to identify study schools and select study teachers. Strong relationships between the research staff and the District 1 coordinator enabled us to obtain extensive sets of data for each teacher and class of students. For example, the coordinator regularly contacted teachers to remind them to send their monthly teacher logs to the research center. She also provided monthly focus meetings for all teachers using reform curricula in the district. During the second and third years of data collection, the District 1 coordinator helped locate study students who were no longer in study classes and asked more teachers to participate in the study in order for us to follow these students longitudinally. In District 3, the on-site coordinator was the principal of the only middle school in the district. He actively supported teachers in their transition to make changes in mathematics curriculum and instruction by informally observing teachers and providing encouraging feedback, handling parental concerns, and providing monthly meetings for teachers to make long-range plans and discuss units in depth before teaching them. In this small district, the coordinator was frequently in contact with the district superintendent and the principal of the only district school that enrolled fifth-grade students.

In contrast, initial contact persons in Districts 2 and 4 were not the on-site coordinators for the study. In District 4, initial discussions transpired through interactions with a university faculty member who regularly worked with one of the multiple school districts in a large urban school system. She was instrumental in working with district administrators to select the sole middle school in the study from District 4. The assistant principal for mathematics and science at the selected middle school became the on-site coordinator in that district. The District 4 coordinator supported teachers' changes in curriculum and instruction through personal monthly meetings with each teacher in which they discussed reform and research-based ideas for mathematics teaching and learning. In District 2, study details were initially discussed with directors of the Urban Systemic Initiative who worked with district administrators to select study schools. The responsibility of coordinating this study was later assigned to the district mathematics specialist in late July, shortly before we met teachers during the first summer institute with study teachers, and details of the study were unfamiliar to her. In contrast with the other districts, a different coordinator was selected in District 2 each year of data collection, partially as a result of restructuring district administrative positions. This meant that each year the coordinator in District 2 was unfamiliar with research goals and the complex issues that had arisen in relation to past data collection.

Districts were asked to provide standardized test scores in mathematics for participating students for the year prior to the study and each year during the study. (As noted in Chapter 3, an additional form of standardized test was not administered explicitly for the study. Districts provided results of the tests they already used.) The type of standardized testing administered to students varied in each district from year to year. Some districts included both SAT9 (Harcourt Educational Measurement, 1996) or CAT (McGraw-Hill, 1992) and a component addressing state performance indicators. Because data from the SAT9 or CAT were used to determine retention and summer school placements, these data were available before results of tests addressing state performance indicators, which became available during the fall semester of the subsequent school year. In District 3, the on-site study coordinator promptly sent standardized test scores to the research center each year. The process of obtaining complete sets of standardized achievement data for study students in the other three districts, however, was an arduous task. In District 1, changes in the software used by the district assessment office led to significant delays in receiving the test score data. The on-site coordinator interceded with the district assessment office to have standardized test scores for study students sent to the research center, which occurred during the final year of data collection. District 2 required the research staff to list both state and district student identification numbers for all students who participated in the study. District ID numbers were included on official class lists. Many teachers, however, provided unofficial class lists, and the research staff had to request official rosters from these teachers and their principals. District student ID numbers for students who entered study classes during the school year were even more difficult to obtain and placed additional burden on study teachers. State ID numbers for all students in District 2 were collated by the on-site coordinator of the study. Results of standardized tests in this district for all three years of the study were received during the spring semester after data collection was completed in the schools. In District 4, data for eighth grade students who were promoted to high school were purged from the school database in June. As a result, scores of the standardized tests for eighth-grade students were collated by the district assessment office and were sent to the research center during the spring semester after data collection concluded.

Cohorts of Teachers and Students

After the schools were chosen, district administrators and principals selected study teachers. By study design, two cohorts of students, one composed of fifth-grade students and one of sixth-grade students were to be followed for three years, and a cohort of seventh-grade students was to be followed for two years. Fifth-grade teachers were, therefore, asked to participate in the study for one year, sixth- and eighth-grade teachers for two years, and seventh-grade teachers for three years. In this way, we envisioned studying the effects of teachers teaching the same MiC units or the same conventional curriculum over multiple years. In Districts 1-3, a total of 19 fifth-grade teachers participated in the study. Fifth-grade teachers in study schools typically taught mathematics as one of many subjects, and many teachers were selected to provide a large cohort of students to follow longitudinally. In District 4, a cohort beginning in fifth grade was not arranged because of the size of this urban district. Many elementary schools were feeder schools into the sole study middle school, and the possibility of scheduling study students into several classes was virtually impossible.

As the data collection progressed, all fifth-grade study teachers, 37% of the sixth-grade teachers, 14% of the seventh-grade teachers, and 46% of the eighth-grade teachers participated in the study as envisioned (see Table 4-1). Although we expected normal occurrences such as changes in personnel due to smaller enrollments or parental leaves, we did not anticipate teachers moving from one grade level to the next with their classes (District 2), which involved 11% of the sixth-grade teachers and 14% of the seventh-grade teachers. Although this “looping” provided opportunities for students to have the same mathematics teacher throughout middle school, it meant that teachers in the MiC group taught different MiC units every year. The increased planning load for these teachers became a serious issue, which the district addressed in the second and third years of data collection by providing one common day of release time per month for teachers to prepare lessons together.

Table 4-1.
Number of Years Expected for Teacher Participation and Actual Length of Participation

Grade *(Number of Teachers)*	Number of Study Years Expected at Grade Level	Teachers Who Met Expectations *(%)*	Other Teachers Assigned to Study Classes* *(%)*	Teachers Who Moved with Their Classes *(%)*	Teachers Who Left the Study** *(%)*
5 *(19)*	1	100	--	--	--
6 *(27)*	2	37	15	11	37
7 *(29)*	3	14	31	14	41
8 *(13)*	2	46	23	--	31

*In order to follow students longitudinally, teachers with classes with many study students were asked to join the study.
**Teachers moved to nonstudy schools, were reassigned to another grade level, accepted an administrative position, were on family leave, or withdrew from the study.

Over the three years of data collection many teachers left the study for various reasons including moving from study schools to nonstudy school (Districts 1 and 2), being on family leave (District 2), being reassigned to another grade level (District 3), accepting an administrative position (Districts 1 and 4), or resigning from participation in the study (District 2). These changes involved 37% of the sixth-grade study teachers, 41% of the seventh-grade teachers, and 31% of the eighth-grade teachers. In most cases, additional teachers at each grade level were asked to join the study in order to follow study students longitudinally, which accounted for 15% of the sixth-grade study teachers, 31% of the seventh-grade teachers, and 23% of the eighth-grade teachers. The resignation of four MiC teachers from one middle school in District 2 (as noted above) had serious ramifications for the study: the loss of the potential to follow eight middle-school study classes and five incoming fifth-grade classes from the feeder elementary school during the second and third years of data collection. Another district feeder pattern joined the study in the second year of data collection at the request of district administrators. MiC teacher and student materials from the teachers who left the study were given to schools in the new feeder pattern. Additional student cohorts were initiated in the fall of 1998: two fifth-grade classes, four sixth-grade classes, and four seventh-grade classes. During the fall of 1999, however, it became apparent that the new cohort of fifth-grade students were dispersed among all sixth-grade mathematics classes, leaving less than five study students in each class. A decision was therefore made not to follow these students for the remainder of the study. Consequently, the two sixth-grade teachers at that middle school no longer participated in the study.

Another factor that affected longitudinal research was initiatives in Districts 3 and 4 for students to take algebra in the eighth grade. Many MiC study students met the criteria established for these programs, eliminating our potential to follow these study students through the second or third years of data collection, depending on their cohort.

In the comparative research design, we had hoped that one-third of the study students would participate as the group using conventional curricula. Districts 1 and 2 (as noted above) had difficulty determining sites using conventional curricula. In the first year of data collection, only four fifth-grade classes studying conventional curricula involving three teachers were in the study. The number of students in conventional groups at other grade levels was also small, and an eighth-grade conventional group in District 2 was not available during the third year of data collection.

Study Assessments

In all research sites, district standardized testing took place at different points during the second semester. Given the importance of standardized tests for teachers and administrators, in three districts study assessments were given to students after district tests. A teacher in District 1 reported, "I have concerns about the two [study] assessments that must be completed early next week. My kids are a bit burned out" (Fifth grade, conventional curriculum, Fulton, Journal entry 5/8/98). In District 3, a sixth-grade teacher reported that students completed study assessments after three weeks of practice for and completion of standardized testing, and students "were burned out by the time MiC testing came" (Sixth grade, MiC, Solomon, Personal communication 5/14/98). Seventh-grade students also felt pressured: "Kids were ''over-tested' and stressed, and we felt rushed" (Seventh-grade, MiC, Perry, Personal communication 5/14/98). Eighth-grade students endured 20 days of standardized tests before taking study assessments (Eighth grade, MiC, Wells, Interview 12/3/98). In District 4, study assessments were sandwiched between two sets of district standardized testing. In District 1 and 2 conventional classes, teachers reported that students were discontent with taking additional tests. Priority given to standardized testing is understandable. Study assessments, however, did not hold the same high-stakes nature to teachers and students.

While most teachers followed the systematic procedures developed for administration of study assessments, results in some MiC classes in District 2 were compromised by changes in these procedures. In one case, extensive amounts of time were devoted to instructions for completing fifth-grade assessments, which gave students insufficient time to complete the items. In the second year of data collection, two teachers (one seventh, one eighth) in the MiC group administered multiple study assessments on the same day, one immediately after the other. An eighth-grade teacher told students that she showed the Problem Solving Assessment (PSA) to a group of non-teaching professionals who were surprised that students were asked to complete items designed to elicit complex reasoning in multiple content strands. The student who relayed this incident in her assessment booklet commented that she subsequently did not attempt to solve assessment items. Other students in this teacher's classes also left many items blank. Another teacher complained that her "at risk" eighth-grade students were unable to concentrate long enough to take study assessments, and students protested having to take assessments that were not counted as part of their course grades. Students subsequently wrote comments such as “I don’t know” or “I don’t understand” instead of reasoned responses. These statements were evidenced in 80% of the assessments taken by students in this teacher’s classes, with students completing at most 4 of the 21 items on the PSA. In the spring of the second year of data collection in Districts 1 and 4, class periods were interrupted by multiple bomb scares (in the aftermath of the school shootings in Colorado that spring), and some teachers did not allow extra time for students to complete the assessments. Additional challenges developed with respect to teachers in the conventional group who did not administer study assessments because of special school-related programs in April and May. When possible, the research staff returned assessments for completion before scoring institutes were held. Thus, many students took study assessments in an array of less than ideal conditions. Furthermore, many assessments were incomplete. The results on these assessments were not necessarily reflective of students’ understanding and application of mathematics.

Instructional Context

The longitudinal study began when MiC was initially published in its commercial forms with comprehensive teacher guides and ancillary materials. Although 17% of the MiC teachers had taught at least one MiC prepublication or field-test unit prior to the study (see Table 4-2), the units were used as supplementary materials to their primary mathematics curricula. Teachers generally received MiC units during district professional development opportunities, and the units they received may not have been specific to the teachers’ particular grade levels. When the study began, MiC teachers implemented MiC during the entire school year for the first time, and they were asked to teach units designed for the grade level they were teaching.

Table 4-2.
Experience Teaching Mathematics in Context Prior to the Study

Grade *(No. of Teachers)*	Years Experience Teaching MiC Prior to Study* *(Number of Teachers)*			
	None	< 1 semester	1 year	> 1 year
1997-1998				
5 *(16)*	5	3	8	0
6 *(15)*	4	4	3	4
7 *(11)*	4	2	0	5
*1998-1999***				
6 *(7)*	5	1	0	1
7 *(7)*	6	0	1	0
8 *(6)*	2	2	1	2
*1999-2000***				
7 *(6)*	4	0	1	1
8 *(2)*	2	0	0	0

*Experience teaching MiC field-test or prepublication units

**Includes only teachers new to the study

Interview Data

As part of their participation in the study, each principal was interviewed once during the school year, and each study teacher was interviewed twice each year of data collection. Interviews were conducted with all study principals and teachers during the fall semester about the school context. Among the questions asked to both principals and teachers were a set of questions about the best ways for students to learn mathematics, opportunities for professional development, and administrative support for improvement of teaching. The consistent set of questions enabled the research staff to look for similarities and differences in ideas among the principal and teachers from the same school and provided information for descriptions of the context in which students' learning took place. The following descriptions from a District 2 middle school illustrate confluence or inconsistencies in visions for students' learning of mathematics:

> Burnett, the principal at Guggenheim Middle School, believed the role of his school was to prepare students for the technologically-based job market that they would enter in the twenty-first century. He stressed the importance of providing students with the problem solving and analytic skills necessary to compete for higher paying jobs and that would enable them to be flexible life-long learners. Burnett defined mathematics as problem solving and underscored the need for mathematics instruction that capitalized on the connections between mathematics and real life applications. Although Burnett indicated that students have different learning styles and capabilities, he believed that, overall, engaging them in problem solving experiences best develops mathematical ideas and enables students to apply and connect knowledge in new situations. Not surprisingly, Burnett noted the predominance of calculators in the real world and supported their integration into mathematics instruction (Principal, Burnett, Interview 9/18/97).
>
> Keeton, a seventh-grade teacher at Guggenheim, felt that students learned best when they used manipulatives and when they applied mathematical ideas to real world contexts. She felt student engagement was a critical component of instruction and that students benefited from participating in exploration of ideas rather than listening to lectures. Keeton viewed mathematics as a system used for problem solving, and she valued students' ability to problem solve over the mastery of basic facts and skills (Seventh grade, MiC, Keeton, Interview 9/18/97).
>
> Teague, another seventh-grade teacher at Guggenheim, identified repetitive drill as an important instructional mode. Although she felt that students could engage in problem solving without having mastered basic skills, she indicated that non-mathematical factors, such as reading comprehension and discipline, often mediated against successful problem solving. Lamenting students' lack of basic fact recall, Teague had resigned herself to allowing calculator use in the classroom to support problem solving. She noted, however, that this policy had resulted in parental resistance (Seventh grade, MiC, Teague, Interview 9/18/97).

As these descriptions illustrate, Burnett expressed strong reform-minded beliefs about the goals for mathematics instruction and the pedagogy that would best engender these goals. Keeton's views seemed to align closely with her principal's although they were far less emphatic. On the other hand, Teague's beliefs about mathematics instruction were more traditional. She tended to bemoan perceived student weaknesses rather than focus on instructional issues that might address student needs. Burnett underscored the importance of preparing students to be productive citizens, yet Teague did not mention connecting mathematics with the students' world as an important component of instruction. Burnett's vision for learning and for mathematics instruction had not yet filtered through to his entire staff.

The interviews conducted yearly provided substantial insight into the instructional context students experienced. The interviews also provided information that supplemented questionnaire data regarding school capacity and students' opportunity to learn mathematics with understanding, which are described in Monograph 3.

In the first interviews of each year of data collection, teachers were also asked about the influence of standardized tests on instruction. Responses varied from changing the content and sequence of instruction based on the tests to limited change because of alignment between the curriculum and the testing program. For example, Keeton stated that the state standardized tests influenced the order of content taught and the emphasis on the types of skills necessary to be successful with the tests (Seventh grade, MiC, Keeton, Interview 9/18/97). Teague agreed that the state standardized tests influenced the order of content taught and the emphasis she placed on computation skill drills (Seventh grade, MiC, Teague, Interview 9/18/97). In contrast, Wells, an eighth-grade teacher in District 3, felt that the district standards, which were well aligned with MiC, were more rigorous than the state standards for mathematics, and as such, did not change instructional emphases in preparation for standardized testing (Eighth grade, MiC, Wells, Interview 12/3/98).

In the second and third years of data collection, teachers were also asked to describe how their experiences teaching MiC would influence their teaching in the current year. Teachers reported that they were becoming more facilitative during instruction and that instructional pacing was much improved (Interviews Broughton, 10/20/98; Carlson, 9/13/99; Downer, 12/16/98; Perry, 12/18/98; Schlueter 11/9/98). Teachers noted differences in student engagement during instruction because they had studied mathematics with MiC during the previous school year (Interviews Perry, 12/18/98; Schlueter, 11/9/98). Teachers had developed ways to allow students to do more mathematical thinking in class, as described by Keeton:

> Before we open the book, I read the questions and ask them if there any questions on each lesson question. Do you understand what it's asking for? What does that mean? How are you going to do that? And we'll go through all of the pages that I want them to do for that day. Then they actually start their assignment. I found that to be much more successful. (Eighth grade, MiC, Keeton, Interview 10/23/98)

Teachers also helped students with expressing their ideas in writing, as illustrated by Downer:

> [Students] are having the most difficulty with explanation. . . . They'll give me the answer—and the answer is right—but they won't explain or sometimes they can't explain to me how they came up with that answer. So my approach is what did you say to yourself first? Talk it out to me. Now please write that down. Forget the grammar for a minute. Just write your thoughts, as they are in your head, on paper. And that's how I've been getting them to write. (Sixth grade, MiC, Downer, Interview 12/16/98)

Teachers also cut back on supplemental activities such as problems of the week (Eighth grade, MiC, Keeton, Interview 10/23/98) and instituted different methods for grading such as partial-credit scoring (Sixth grade, MiC, Broughton, Interview 10/20/98) and checking more work during class time (Eighth grade, MiC, Keeton, Interview 10/23/98). Furthermore, teachers felt that they better understood the goals MiC curriculum, as noted by Heath and Gallardo:

> I feel like I'm more organized this year. Last year I kept going along, but didn't see really see the whole picture. . . . Now sometimes I know what is expected and I can lead toward that direction. I feel like I can better pick out which things I might want to spend more time on, less time on, which things I might say is a good homework question, as opposed to using it because the book said to make it homework. (Seventh grade, MiC, Heath, Interview 10/26/98)
>
> The biggest thing I'm doing differently this year is that I'm focusing in on the math concepts more. . . . I want to get to the math and not at the interesting context that is being used to teach these math concepts. I'm a lot more familiar with how to present it, how to teach it, and how to evaluate it. (Eighth grade, MiC, Gallardo, Interview 9/20/99).

Teachers also commented that they varied grouping for instruction (i.e., independent work, work in pairs, work in groups, work with the whole class) based on the mathematical content, which was emphatically stated by Gallardo: "I decide on what's the best way to teach the class" (Eighth grade, MiC, Gallardo, Interview 9/20/99). Thus, teachers' comfort with the curriculum and changes in pedagogy resulted from their experience with MiC. These changes were substantial and reflected personal growth.

Teachers also completed an interview late in the spring semester of each year of data collection; only a few teachers did not complete the interview because they resigned from the study (5%), were replaced by other teachers during the school year (3%), or chose not to complete the second interview (2%). This interview focused on instructional planning and classroom interaction, specifically planning for teaching units or chapters and individual lessons, types of student discussion during class, and the nature of acceptable answers. These interviews were important in documenting and understanding variation in several dimensions of instruction, which are described in detail in Monograph 3.

Teaching Logs

As part of the comparative research design, teachers in Districts 1 and 2 compiled daily teaching logs and wrote weekly journal entries. Information from the teaching logs was used in the analysis of the content of the actual curriculum, the use and modification of curricular materials, lesson planning, instructional activities, and classroom assessment (see Chapter 3 for detailed description of the teaching log). The following two excerpts illustrate the differences between the forms of instruction and types of student activities used throughout the school year by two teachers, eighth-grade MiC teacher Reichers and seventh-grade teacher Hodge who used a conventional curriculum. The teachers' self-reported data were triangulated with the data observers provided regarding the time allotted to various lesson activities.

> Reichers frequently used whole-class discussion and small group work (on nearly half of the 166 reported days), and they were generally given at least equal emphasis with other forms of instruction. Although Reichers reported review of previous material on 50% of the reported days, this received less emphasis than discussion and group work.

Furthermore, discussing answers and solution strategies (on 64% of the reported days) and participating in whole-class discussions (51%) were important student activities in Reichers' instruction. Students listened to their teacher or took notes on 52% of the reported days and reflected on or summarized lesson concepts on 42% of the reported days. These student activities, however, were given less class time (Reichers, Teacher Log 1999-2000). In general, observation reports completed during 1999-2000 class periods supported the information Reichers reported in her teacher logs. The lesson observed on 5/25/00, for example, included: large-group review, p. 16, problem 8 (11 minutes); teacher presentation, p. 17 (3 minutes); small-group work, pp. 17–18, problems 1 and 2 (19 minutes); large-group discussion of solutions to problem 1 and procedures for solving problem 2 (4 minutes); small-group work, continued assignment (14 minutes); large-group, teacher-led discussion of problem solutions (10 minutes); and large-group homework assignment (3 minutes; Diver, Observer notation of time allotted during instruction, eighth grade, MiC, Reichers, Observation 5/25/00).

Hodge usually began each class period with a warm-up activity (on 93% of the 75 reported days), but gave it less than 15% of class time. He frequently used review of previous material (77%), teacher presentation (73%), and whole-class discussion (68%). Teacher presentation and whole-class discussion were most frequently used at least half of the class period, and review of previous material was often given equal emphasis with other forms of instruction. Three student activities were important elements in Hodge's instruction: discussing answers and solution strategies (72%), listening to the teacher or taking notes (71%), and participating in whole-class discussions (67%). Each student activity was given a significant amount of class time (Hodge, Teacher Log 1999-2000). In general, observation reports completed during 1999-2000 class periods supported the information Hodge reported in his teaching logs. The lesson observed on 5/2/00, for example, included: warm-up activity (14 minutes); large-group, correct warm-up problem and homework (18 minutes); large-group, teacher-led review of integers, coordinate system, and rules for operations with integers (29 minutes); and large-group closure (2 minutes; Diver, Observer notation of time allotted during instruction, seventh grade, conventional curriculum, Hodge, Observation 5/24/00).

As these data illustrate, differences in classroom instruction were evident in these study classes. Although both teachers devoted large amounts of time to discussion of strategies, Reichers tended to allow the students to do more of the mathematical work before such discussion, which in the selected class period was precursor to working on the second problem in the lesson. Hodge, on the other hand, tended to present the mathematical content directly during whole-class discussion. In the selected class period, half of the class period was devoted to the warm-up activity and review and half the class on teacher-led, large-group instruction on operations with integers.

Teachers were asked to respond to journal questions once a week. The questions were designed to document lesson content that was emphasized or modified and to describe notable classroom events. Journal entries provided glimpses of classroom interaction from the teachers' perspectives and insight into teachers' decision making, both prior to instruction and during instruction. Variation was also documented in the instructional methods used. In many cases, teachers' journal entries indicated that MiC was taught in ways more in line with the intended philosophy of the curriculum. For example, Keeton noted that she added questions to help students understand mathematical content:

> As we reviewed past assignments, students' verbal responses were slow, which led me to believe I needed to place more emphasis on factors and multiples. Students began to respond quickly to my questions for examples of factors and multiples. Then I went ahead with today's assignments. (Eighth grade, MiC, Keeton, Journal entry 11/17/98)

Teachers described helping students make connections among representations of situations. For example, Heath noted in a journal entry:

> When we were going over p. 27, I had a student try to make a graph on the chalkboard as I showed my car (a piece of chalk) driving around a track. That way the class saw simultaneously what happened to the graph as the car moved. (Seventh grade, MiC, Heath, Journal entry 10/23/97)

Teachers also noted how they integrated students' suggestions into instruction. When teaching geometry lessons on two-dimensional representations of three-dimensional objects, for example, Piccolo emphasized that "what we see is determined by our perspective" and described that "we gave examples and changed positions as we looked at models" (Fifth grade, MiC, Piccolo, Journal entry 9/16/97). Piccolo allowed students to introduce different strategies that provided access to problem solutions, noting:

> Using blocks [to model objects in a diagram], one student suggested to put the model on the diagram in the book and lift the book to eye level at each arrow [camera position] to determine the correct matching [of side views to corresponding camera positions]. Some students lifted the book and others got on their knees in front of their desks. It worked! The majority of the students were successful after using this strategy. (Fifth grade, MiC, Piccolo, Journal entry 9/23/97)

In other classes, MiC was taught with conventional teaching methods. St. James, who taught in very conventional ways before teaching one class using MiC, found it difficult to teach without, for example, "expounding all I know about fractions" (Seventh grade, MiC, St. James, Journal entry 2/8/00), especially when students did not understand lesson content. In the middle of the spring semester, he modified his instructional approach:

> I have adjusted the presentation a little. The class seems a little more into lesson and not talking so much. At least I'm getting more direct feedback from them. What I did was present it a little more. Students responded better. (Seventh grade, MiC, St. James, Journal entry 3/29/00)

Later in the school year, he made further changes:

> We are doing a lot of teacher-to-student lessons, and I'm reversing styles of teaching, traditional mixed with the MiC style. The class seems to be doing better right now. They learned a little about linear equations. (Seventh grade, MiC, St. James, Journal entry 4/18/00)

In some cases, MiC units were supplemented with traditional materials to the extent that the units were subsumed by the supplementary materials. A fifth-grade MiC teacher of low ability students in District 1 stated:

> My students are incredibly weak. As they progress through the units, they need to understand the “real” value of numbers. I intend to give them that basic foundation as they continue to succeed in this program. (Fifth grade, MiC, Linne, Journal entry 9/7/97)

A sixth-grade MiC teacher of low ability students in District 2 reported in journal entries for one month that she taught a sequence of lessons on fraction operations with varying success. These lessons included use of specific algorithms such as “cross canceling” in multiplication of fractions, steps to change a mixed number to an improper fraction, division of fractions by changing the division sign to multiplication and flipping the second fraction, and changing the final answer from an improper fraction to a mixed number (Sixth grade, MiC, Broughton, Journal entries 4/28/98, 5/4/98, 5/6/98, 5/13/98, 5/18/98, 5/20/98). A seventh-grade MiC teacher in District 1 regularly reported in journal entries that his lessons were supplemented with daily worksheets. In one series of journal entries, for example, Donnely wrote that these worksheets were necessary to “maintain students’ arithmetic skills” and “reinforce work” or “refresh students’ memories about concepts that were covered in previous lessons” (Seventh grade, MiC, Donnely, Journal entries 1/7/98, 1/8/98, and 1/6/98). On occasion, students completed assignments for extra credit when they “wrote definitions 20 times for bonus points” (Seventh grade, MiC, Donnely, Journal entry 1/6/98). Parental concerns also influenced teachers’ use of supplementary materials:

> I have had many parents express concerns that their children aren’t getting enough “math” instruction. For these parents, math means computation only. What I have decided to do is use some chapter post-tests from a textbook that correlate to what we have been doing in MiC. (Fifth grade, MiC, LaSalle, Journal entry 3/2/98)

When teachers were absent due to illness or participation in workshops, they felt that it was inappropriate to ask substitute teachers to continue MiC lessons. Instead, they prepared lessons that reinforced computational skills.

Another factor that influenced the use of supplementary materials was the significant pressures teachers felt for their students to achieve high scores on district and state standardized tests, as illustrated by Carlson:

> I’m sure you guys aren’t too happy with me only doing MiC one day a week, but the state test is hanging over us like death. The kids are nervous and we feel a huge amount of our worth as teachers is based on the outcomes of these tests. I really feel that the kids need practice on test issues. (Eighth grade, MiC, Carlson, Journal entry 11/8/99)

Preparation took different forms, particularly in District 2 in which teachers taught basic skills for several weeks through pre- and post-tests on targeted content, vocabulary used on middle-school and high-school standardized tests, worksheets on operations with fractions, commercially published test preparation workbooks, and special programs in which students practiced particular mathematics procedures for 30 minutes in every subject matter class.

In their journal entries, teachers wrote about planning for instruction, teaching new content, learning new instructional approaches, and developing pedagogical content knowledge for new instructional topics. The planning involved in teaching standards-based curricula requires more time than generally perceived necessary for teaching using conventional methods (Battista, 1999; Bay, Reys, & Reys, 1999), which was also true for MiC teachers. For example, when planning to teach individual MiC lessons, Heath worked through all of the problems, then read the teacher guide to check whether she had given the intended meaning to the problems and whether she understood the purpose and direction of the lesson. She included hints from the teacher guide during instruction. She also planned to emphasize portions of the lesson that would clarify its purpose and direction. Heath planned daily for individual lessons, taking into consideration students' performance on the previous lesson, group dynamics, and the conceptual complexity of the lesson. The form of instruction she chose, whether whole-class instruction, group work, or independent work, was determined by the lesson and accompanying activities (Seventh grade, MiC, Heath, Interview 5/11/99).

The content of MiC units was sometimes a challenge for MiC teachers. In MiC, mathematical content is introduced in a different instructional sequence than in conventional middle-school curricula. For example, concepts related to percent are introduced in fifth- and sixth-grade MiC units, rather than more conventionally in eighth grade, and content traditionally reserved for high-school students such as concepts in algebra and geometry is introduced in fifth- through eighth-grade units. Teachers also had to learn how to introduce and work with new mathematical tools that supported students' thinking such as the ratio table, which is illustrated in this teacher's request:

> Please tell me how to explain p. 52 Section B The Ratio Table in *Number Tools* [a MiC ancillary resource]. The students and I understand doubling and adding. However, we want to know which columns to add if they are not in bold type. (Fifth grade, MiC, Murphy, Journal entry 1/6/98)

In addition to the challenges of teaching new mathematical content and new ways to support student thinking about mathematics, few MiC teachers had experience teaching mathematics that emphasized the development of conceptual understanding and student reasoning rather than algorithms and procedures, which is illustrated in this comment:

> I am personally surprised at the difficulty in teaching the interrelated concepts of ratio, fractions, decimals, and percent. Since I personally learned all of these concepts by algorithms, it is difficult for me to rethink how to present this to students. (Sixth grade, MiC, Gollen, Journal entry 2/11/99)

During instruction, teachers draw on pedagogical content knowledge, which includes difficulties students might encounter as they learn new topics, typical sequences students go through as they learn particular content, and potential ways of helping students overcome difficulties (Bransford, Brown, & Cocking, 1999). Experimental teachers frequently talked about their lack of pedagogical content knowledge for new content, for example:

> I am often unable to anticipate the students' success or difficulties. For instance, I felt that students would be able to easily change from ratios to fractions to decimals to percent, but they had much difficulty. On the other hand, I was surprised that the students read the pie chart meters very well. (Sixth grade, MiC, Gollen, Journal entry 2/22/99)

Teachers expressed concern about how to help students when they could not find their own solutions, as one teacher explained:

> I am concerned about my students' reluctance to answer questions. I am not sure whether it is the unknown and unfamiliar that is preventing spontaneous behavior or something else. (Fifth grade, MiC, Murphy, Journal entry 1/22/98)

Teachers worked with guiding students to complete mathematical tasks and found that they needed to develop ways to provide time for students to think about instructional tasks, reason out strategies, and determine solutions: "It is so hard to get the kids to think through things, to take enough time to relax with the information, so they can see patterns emerge" (Eighth grade, MiC, Reichers, Journal entry 1/6/99). Teachers also struggled with providing time for students to discuss instructional tasks and gradually developed ways for students to do the mathematical work *and* discuss various strategies, as one teacher noted: "I learned to introduce the lesson, ask students to work on a few problems, reconvene to discuss the content with the whole class, and repeat that during the lesson" (Sixth-grade, MiC, Dillard, Personal communication 2/26/98). Teachers also worked at improving the quality of group work:

> I had been struggling with group work. I felt the students were not working, and it was a waste of time. Today, I was able to see how it can work. I changed class procedure a little. We discussed as a class what we were going to do and how we can answer some of the questions. Then students worked on their own. This material is not easy for these children, and I was pleased to see what an exceptional job they had done and how well they worked in their groups and helped each other understand. (Sixth-grade, MiC, Weatherspoon, Journal entry 1/28/98)

Teachers' own understanding of the mathematics, ways the mathematics is presented, pedagogical content knowledge, and instructional strategies, all of which are central to effective instruction, were being developed as teachers taught MiC for the first time for the whole school year.

Different instructional methods and supplementary materials also affected classes using conventional curricula. Some teachers taught in more reform-oriented ways and used curricular materials that emphasized mathematical problem solving, reasoning, and communication. One teacher using a conventional curriculum in District 1, for example, taught for conceptual understanding:

> Sometimes I'm at a loss to explain some of the "why do we do that" when we're discussing decimals operations. Addition and subtraction I can usually explain, but multiplication and division sometimes throw me. I pride myself on making sure my students understand all the "whys" in math, so I'll have to do more research on decimal concepts. (Fifth grade, conventional curriculum, Fulton, Journal entry 4/3/98)

Fulton also modified instruction based on students' inquiries, as he noted in this example:

> Somehow the fraction 222/391 ended up on my board and someone asked if it could be reduced. I assured them I didn't know! So I asked how we could find out. Someone said, "Get a calculator." I gave calculators to all students and said, "Find out." It was not without silence and stunned looks! *No one* knew exactly how this calculator was going to help us decide if 222/391 could be reduced. We finally decided to try dividing numbers into each numerator and denominator and found factors, but none were common. I thought this was a good lesson in terms of finding out the calculators are not necessarily the "saviors of the world" that most kids think they are! (Fifth grade, conventional curriculum, Fulton, Journal entry 12/5/97)

In this situation, Fulton launched a mathematical investigation in response to a student's inquiry. During this exploration he encouraged students to determine approaches to solving the problem on their own, and with his guidance, students were able to solve the problem. Another teacher using a conventional curriculum focused on the mathematical processes students used in solving problems, as illustrated in this description of classroom interaction in a seventh-grade class in District 2:

> Students were continually asked to justify and explain their answers and conjectures. They were asked questions about the lesson's logic problem. Students made statements about who was taking or not taking particular language courses. They had to give the reasoning for their statements . . . and the rest of the class either agreed or disagreed. (Seventh grade, conventional curriculum, Cunningham, Journal entry 3/16/98)

Teachers who used conventional curricula also modified textbook materials. For example, one teacher reported that review of fraction operations was necessary:

> I have opted to go back to Chapters 4 and 5 and teach fractions again, as my students seem to have a limited understanding of fractions and fractional operations. I am moving topics around from Chapters 4 and 5 to better meet the needs of my students. (Seventh grade, conventional curriculum, McLaughlin, Journal entry 4/1/98)

A sixth-grade teacher commented that she disliked the textbook explanation of the distributive property and skipped the lesson until she found a different lesson (Sixth grade, conventional curriculum, Tallackson, Journal entry 2/23/98). Other teachers added favorite projects, as illustrated by Fulton:

> I always enjoy these couple of weeks during the running of the Iditarod. It provides many high interest activities for the students while exposing them to something that I love—sled dog racing— and something very few of them ever heard about—the race itself. (Fifth grade, conventional curriculum, Fulton, Journal entry 3/13/98)

Other teachers rewarded students with special lessons, as McLaughlin described:

> Students earned a day of "fun" math for good behavior. Therefore, I added these two worksheets to this unit. Students could work alone or with a partner. I counted these worksheets as project grades. (Seventh grade, conventional curriculum, McLaughlin, Journal entry 11/7/97)

Teaching logs and journal entries added significantly to our understanding of what transpired on a daily basis in study classes. The numbers of teachers in Districts 1 and 2 who sent logs and journal entries to the research center monthly varied greatly (see Table 4-3), despite our extensive efforts to collect a full set of teaching logs from each teacher. During the first and second years of data collection, reminders were sent to teachers from the research staff, and graduate research assistants encouraged teachers to continue completing this important source of data through personalized letters of interest in the teachers' work, as illustrated in excerpts from such letters:

> It's wonderful to hear that students are enjoying the activities and are even surpassing your expectations for understanding the concepts! In the checklist portion of the log, it is evident that informal assessment is an important vehicle for you to gauge student understanding of a variety of concepts. In the journal, you might include examples of the types of questions you ask as you observe the students working as well as examples of student comments that are typical, or even atypical, during their small group or whole class discussions. Your extensive concluding entry was marvelous; we are always eager to see examples of student strategies during problem solving! (Graduate research assistant, Personal communication to fifth-grade MiC teacher Mitchell 11/14/97).

Thanks so much for the interesting anecdotes about the students' and parents' impressions of MiC. In your journal entries, please continue to keep us posted on your students' difficulties and successes with MiC, the decisions you make in response to student difficulties, and the reasons you make these decisions (Graduate research assistant, Personal communication to fifth-grade MiC teacher Fiske 11/24/97).

Table 4-3.
Number of Teaching Logs Received, by Grade and Year

Grade *(No. of Teachers*)*	Number of Teaching Logs Per Teacher	Percent of Teachers Submitting Teaching Logs			
		0-2 Logs	3-6 Logs	7-8 Logs	9 Logs
1997-1998					
5 *(13)*	0-9	23	15	23	38
6 *(12)*	0-9	33	42	8	17
7 *(10)*	0-9	10	20	30	40
1998-1999					
6 *(12)*	0-9	33	25	0	42
7 *(12)*	0-9	50	17	8	25
8 *(10)*	0-9	30	20	0	50
1999-2000					
7 *(9)*	0-9	44	22	0	33
8 *(9)*	0-9	22	11	11	56

*Includes teachers who taught portions of the school year

District on-site coordinators interceded in asking teachers to send in the logs, and principals were contacted to remind teachers of this commitment. Prior to the second year of data collection, teachers negotiated a substantial increase in the amount compensated for each teaching log received by the research center. This increase in compensation, however, did not increase the number of teachers who regularly compiled teaching logs, nor did the quality of teacher journal entries change substantively.

Classroom Observations

Journal entries provided glimpses of classroom interaction from the teachers' perspectives. By their nature, the data gathered through journal entries varied by teacher. Classroom observations, on the other hand, provided a picture of the mathematical interaction that transpired in study classes based on a consistent set of variables considered during each observation. Observations, which were conducted in Districts 1 and 2, provided a means of comparison of instruction experienced by study students in each treatment as well as between treatments. Furthermore, information gathered during observations was used to triangulate information provided by teachers in teaching logs with respect to content covered and time allotted to various lesson formats and student activities, in journal entries related to classroom events and modifications of the curriculum, and on questionnaires regarding the number and titles of MiC units or textbook chapters taught. (See Chapter 3 for a detailed explanation of the classroom observation instrument used to gather this information.)

Among the items included in the observation instrument were nine indices that provided information related to the nature of the mathematical interaction that occurred during instruction. Each index was designed for the observer to record a rating that captured the essence of the interaction that occurred. Additionally, observers were asked to write evidence that supported each rating. This evidence provided substantial insight into the instruction students experienced. To illustrate the descriptions written by observers, the following excerpts exemplify the development of conceptual understanding of lesson content for MiC students and students using conventional curricula:

> Students created their own meaning for rules governing the multiplication property of inequalities. Ms. Pimm worked with concrete examples to foster conceptual understanding. (Diver, Observer evidence, eighth grade, conventional curriculum, Pimm, Observation 2/9/00)
>
> Students were creating meaning for themselves of many ideas: angle measure, number of faces, edges and vertices of polyhedra, and construction of polyhedra. Ms. Lawton did a very good job of connecting procedural and conceptual knowledge. (Diver, Observer evidence, seventh grade, MiC, Lawton, Observation 9/20/99)
>
> The entire quiz promoted conceptual understanding of number, arithmetic operations, and reasonableness of mathematics statements. The focus of the quiz and of the problem of the day was on conceptual understanding of the concepts, not procedural knowledge. (Diver, Observer evidence, seventh grade, conventional curriculum, Hodge, Observation 9/22/99)
>
> Continuous concept development was present. Mr. Gallardo used past lessons, such as p. 29 in the prior section, and expanded on them to develop today's lesson on optimization. (Gibbs, Observer evidence, eighth grade, MiC, Gallardo, Observation 4/12/00)

Observers also noted occasions in which conceptual development of mathematics concepts was compromised in some way, for example:

> Although the students worked on conceptual understanding with the problems in the unit, Ms. Reichers tended to stop them before they really created their own meaning, and she then carefully directed their answers and conclusions. For 10a, she told them that if the angle was 45 degrees, the height and distance were the same. The students were not given enough time to make discoveries on their own. (Diver, Observer evidence, eighth grade, MiC, Reichers, Observation 11/15/99)
>
> Mr. St. James gave rules for positive and negative number operations and did not relate this to the lesson with Veda patterns at all. He gave a rule to the students for determining if a pattern ended where it started, using addition of signed

> numbers. There was no conceptual development of any mathematical ideas. Mr. St. James altered the lesson in the unit so that was not possible. (Diver, Observer evidence, seventh grade, MiC, St. James, Observation 10/18/99)
>
> When Ms. Carlson reduced the fraction, 15/27 = 5/9, she drew a cube root sign around the 15/27. For 4/12, she drew a fourth root sign around 4/12. She must have meant to "divide out" the numerators and denominators by factors of 3 or of 4, but the symbols she used were root signs. After the warm-up problem was explained by the a/b + c/d = (ad + bc)/bd method, I asked a student what was meant by the vertical method. He said that it was a completely different method and showed me examples from his notebook. These examples were addition/subtraction problems where the fractions were written vertically and equivalent fractions with the same denominator (LCD) were found. That student, and maybe many others, had not seen the commonality of the two methods. During the large group lesson, the focus was on an algorithm for solving rate problems using a given formula. Regarding of the Practice Quiz (for the state performance assessment), the ratio table in problem 3 was not viewed as a ratio table. In fact, the table was not discussed. Students were having trouble with it. Many left the spot above $145.00 blank. (Gibbs, Observer evidence, eighth grade, MiC, Carlson, Observation 10/7/99)

Observers were also given the opportunity to write additional comments about the lesson. Ms. Diver, the observer in District 1, applauded excellent teaching when it occurred, for example: "Wow! I only wish I could see more lessons like this one. It was wonderful, even if it could have been better!" (Diver, Observer comments, eighth grade, MiC, Reichers, Observation 1/13/00). She frequently included detailed descriptions of lessons, which is illustrated in this excerpt:

> In the homework for today, students were asked to draw two pyramids, but Ms. Burton never went over this. She talked about the base of a pyramid and its relationship to the number of edges in the pyramid. She did not look at the relationship between the number of edges in the base and its relationship to the number of vertices. Students said to multiply by 2. They never looked at the figures or discussed why this would be true. Then she asked students if they could build a pyramid with 9 edges. Someone sidetracked her, and she never answered or referred to it again. There was no continuity. She was not building toward a conclusion. She completely ignored the fact that triangles add to the stability of a figure. Even though the students had made bar models yesterday, Ms. Burton was the only person in the room to touch one today. This made it very difficult for the students to make conjectures about the mathematics. (Diver, Observer comments, seventh grade, MiC, Burton, Observation 2/23/99)

Ms. Gibbs, the observer in District 2, also celebrated excellence in teaching by specifically noting particular events for the research coordinator, such as:

> Note: I loved today's MiC lesson. I'd seen it done before. But this one was exciting because of the teacher's questions, the students' responses, and the teacher's handling of responses. There was a lot of background discussion at the beginning so that every student should have understood what the situations were about and what was being asked of them. (Gibbs, Observer comments, seventh grade, MiC, Broughton, Observation 1/18/00)

In contrast to Diver's lengthy descriptions of lessons, Gibbs tended to provide great detail in the evidence for each rating (as demonstrated above for eighth-grade MiC teacher Carlson on 10/7/99) and thorough detail of the time allotted to instructional activities. For example:

> Pair work, then large-group discussion of the Problem of the Day. Students were to write everyday items in "simplest form." Examples: 50 pennies (50¢); 24 eggs (2 dozen); 8 ounces of butter (1/2 pound or 1 cup; 7 minutes); Large-group spin-off discussion of metric prefixes, including a demonstration with a meter stick (9 minutes); Large-group discussion of MiC pp. 11–12 (arithmetic trees) and pp. 13–15 (finding prime numbers; 9 minutes); Individual or group work. Today students had 5 assignments: (1) Six students at a time, in 3 rotations, worked on a computer skills test until they obtained a set proficiency; (2) Supplementary worksheet 1.2 on place value; (3) MiC ancillary worksheet on arrow language from MiC *Number Tools, Volume 1*, p. 6; (4) Two pages of Hands-on-Equations solved using manipulatives; (5) The MiC assignment was problems 1-13, pp. 11–15. Today students were allowed to work on the assignments in any order they wished (1 hour 19 minutes; Gibbs, Observer notation of time allotted during instruction, eighth grade, MiC, Carlson, Observation 9/15/99).

The detailed descriptions of instruction experienced and time allotted during instruction provided the research staff with information that would not have been available only by writing supporting evidence for the rating on each index. For example, in the last excerpt of time allotted during instruction, little class time was actually devoted to the MiC lesson assigned for that day, and most of it was completed independently as part of an eclectic assortment of instructional activities during the class period.

At the end of each observation, the observer also conducted a brief postobservation interview of the teacher during which the teacher was asked whether any incidents occurred during the lesson that revealed student misunderstanding or provided opportunities to facilitate student understanding in any way. By including this brief interview, teachers had an opportunity immediately after teaching the class to describe and explain modifications made during the lesson. In some cases, teachers noted specific actions undertaken to clarify explanations or address misunderstandings, for example:

> I needed to ask students to come to the overhead projector to explain how they were working certain problems so the rest of the class could understand their verbal explanations. Just discussing the solutions was not enough. (Sixth grade, MiC, Gollen, Postobservation interview 10/1/98)
>
> Ms. Carlson said that students were having trouble figuring out the height, length, and width of the cereal boxes until she took a blank sheet of paper and physically cut out the square corners. (Seventh grade, MiC, Carlson, Postobservation interview 12/11/98)

On other occasions, teachers mentioned actions that promoted student participation:

> I felt I had to really guide and pull answers out of students. There was a lot of repetition. I had to go back and repeat how we arrived at our answers. (Sixth grade, MiC, Ferguson, Postobservation interview 2/22/99)
>
> I changed my direction and allowed more student participation, through open discussion (probing questions) and board work. (Eighth grade, conventional curriculum, Cunningham, Postobservation interview 3/10/99)

After other observations, teachers noted their responses to student difficulties with respect to performing particular procedures:

> Ms. Friedman said she noticed that some students had some trouble with today's lesson. She did not modify the lesson, yet circulated around to help individual students with what they were doing wrong. (Sixth grade, conventional curriculum, Friedman, Postobservation interview 1/12/99)
>
> Ms. Renlund used the example in which a student asked if you changed improper fractions to mixed numbers or if you reduced a number before going through the process of adding given fractions. Ms. Renlund reiterated that she had told the student not to use either one of those methods and said to wait until the fractions had been added before reducing and/or changing to mixed numbers. (Sixth grade, conventional curriculum, Renlund, Postobservation interview 2/18/99)

The brief postobservation interview provided insight into various pedagogical decisions that transpired during instruction that may not have been otherwise observed.

Classroom observations provided information about the content of instructional materials, the ways lessons were presented and developed, the nature of the mathematical inquiry during instruction, teachers' interactive decisions, and students' involvement in the lesson. In many cases, observations relayed situations in which MiC was taught in ways more in line with the intended philosophy of the curriculum, for example, in this description of a lesson presented by Piccolo (Observation, 5/8/98) from a class using the MiC fifth-grade geometry unit *Figuring All the Angles* (de Lange, van Reeuwijk, Feijs, Middleton, & Pligge, 1997). In the lesson (pp. 15–17), students investigated the distortion caused by representing curved surfaces with flat maps. They looked at flattened curved sections of grapefruit peel and lines of latitude and longitude on flat maps and globes. (The terms latitude and longitude were not used formally in the unit.) In previous lessons in this unit, students used flat maps to locate places using distances and directions. The warm-up questions designed by Piccolo were informal assessments of what the students thought the new section in the unit was about and of prior student knowledge about the shape of the Earth and history of theories about that shape. The information Piccolo learned through this informal assessment was used throughout the class period, as reported by the observer: "All discussions built on prior knowledge and deducing new ideas." Students then completed question 1 on p. 15, "Describe the location of the United States [on pictures of the Earth in the unit]" and shared their answers with the class. Piccolo had placed large laminated maps of the United States on their desks for reference and had a globe in the classroom. She led them in a discussion of the map in the text and the text that described the location of St. Louis, Missouri as 38ºN and 90ºW. Students then began work in their small groups on questions 2-4. Piccolo valued students' statements about mathematics and used them to work toward shared understanding for the class. The observer noted:

> Particularly with respect to airplanes traveling north or west (questions 3 and 4), Ms. Piccolo used students' statements about whether or not the planes would meet to encourage discussions and to lead into following plane flights on a globe.

In the postobservation interview, Piccolo stated:

> Students had been convinced that airplanes flying north would never meet. Using the globe and letting students use their fingers to trace northerly flights helped them reconstruct their ideas. Having a globe for each table might have facilitated that part of the lesson.

In this situation, Piccolo introduced a strategy in a way that did not reduce the mathematical work required of the students and that opened opportunities for students to think about and explore the mathematical ideas. She also noted ways that she could improve instruction the next time she taught the unit. Continuing the lesson, Piccolo asked the students what would happen to the map if it were pressed onto a large grapefruit that she had brought to class. She flattened grapefruit peels on the overhead projector. The demonstration provided a way for students to begin thinking about unit questions related to the distortion caused by creating flat maps of curved surfaces. Piccolo provided a visual support for students' thinking, one that students could refer back to when solving problems. In doing so, she provided a conceptual basis for the mathematics content. Students completed questions 5-7 in their groups and shared their answers in subsequent whole-class discussion. In this lesson, the mathematical content was explored in enough detail for students to think about relationships. The lesson promoted connections among mathematical ideas (flat vs. spherical maps, number lines, latitude and longitude, geometry, and measurement), and connections between mathematics and students' life experiences (maps, globes, and compass directions) were discussed.

In other cases, it was apparent that MiC was taught in ways consistent with conventional pedagogy, which is illustrated in a lesson presented by Harvey (Observation, 3/16/98) using the sixth-grade MiC number unit *Fraction Times* (Keijzer, van Galen, Gravemeijer, Shew, Cole, & Brendefur, 1998). The task asked students to describe the amount each child should receive when groups of children divide a given amount of money and to express this amount as a fraction and a decimal. The lesson as presented provided limited attention to conceptual understanding. The observer noted:

> Some students (not many) understood conceptually that division and fair share were involved. Mr. Harvey gave the class his methods for solving the problem: "I have $2.00 to split 8 ways, so [each gets] 1/4 [of a dollar] or 25 cents," and "To split $9.00 among 5 people, I used $1.75. I'll keep the extra 25 cents."

Students did not have a conceptual basis for dividing monetary amounts into equal parts. Most students relied on trial-and-error strategies. Harvey told students that a fraction can be converted to a decimal by dividing the "top number" by the "bottom number," but he did not provide a way for students to make sense of this procedure other than to say that it worked.

Some teachers using conventional curriculum taught in more reform-oriented ways and used curricular materials that emphasized mathematical problem solving, reasoning, and communication, which is illustrated in a lesson presented by fifth-grade teacher Fulton (Observation, 4/23/98) on operations with decimals and relationships between fractions, decimals and percent. The observer noted:

> Mr. Fulton asked the class where to put the decimal point in the answer when multiplying decimals. One student answered, "Count the places to the right of the decimal in the problem and come in from the right in the answer that many places." Mr. Fulton asked why. Another student answered, "When you multiply tenths times tenths you get hundredths." A discussion followed. This was the first time I have ever heard a student explain why you put the

> decimal where you do in [the answer to] a multiplication problem—and from a fifth grader! Mr. Fulton returned a quiz on decimals at the beginning of the period. Students gave answers and explanations for the procedures they used, such as “You line up the decimals to add them so you keep the whole numbers and the fractions separate. It’s easier to read that way.” Then each student was given a hundreds chart and seven different colored crayons. They were told to use the seven colors to color the blocks and that all spaces of one color must be connected. They were then asked to list the number of spaces for each color and add the numbers to make sure they added up to 100. [The next step was to] make a table, write [the number of spaces for each color] as a fraction and a decimal, and total the fractions and decimals. The class then had a discussion about the equivalence of the fractions and the decimals and their sums. Mr. Fulton introduced the word percent and asked if any one knew what it meant. [Their responses] ranged from “piece” to “average” to “cents” to “piece of a whole” to “part of a hundred” during several minutes of large-group discussion. Mr. Fulton also talked about batting averages and how the newspaper misrepresents them as decimals and percents. When the period was over he gave homework, to name each part of the chart as a percent.

In this lesson, Fulton worked toward conceptual understanding of decimal multiplication, asked students to communicate conceptual underpinnings for procedures, and provided situations in which students could use fractions, decimals, and percents to represent quantities and make connections among these representations. Connections were also made between the mathematics and students' life experiences. Decimals were discussed in relation to money and baseball averages, and percents were discussed in relation to statistics from the sports page.

Through classroom observations, the instruction students experienced was documented for each teacher during the school year. The number of observations per teacher varied in each district (see Table 4-4). Most teachers in District 1 were observed once a month for a total of nine observations per teacher. During the first year of data collection, one teacher in District 1 accepted an administrative position in December; consequently, she was observed three times, and the newly assigned teacher was observed five times. During the second and third years of data collection, one eighth-grade class using a conventional curriculum had three teachers, and two seventh-grade MiC classes had two teachers over the course of the school year. As a result, each teacher was observed only a few times. Teachers in District 2 were observed a total of two to nine times each. Fewer observations were conducted in District 2 due to differences in school schedules, procedures for assigning students to classes, and preparation for district and state standardized testing. In addition, four teachers from one school in District 2 withdrew from participation in the study during the spring semester of the first year of data collection; consequently, they were observed only three times. During the third year of data collection, two seventh-grade MiC classes were observed twice because the teacher had been on parental leave.

Despite having made advanced appointments with teachers for classroom observations, on a few occasions, observers were unable to observe teaching situations. Some teachers compromised observations by administering tests on days of scheduled observations. Other teachers neglected to cancel planned observations when school schedules changed or when they were absent due district professional development opportunities or to illness. These few occasions, however, did not interfere with the breadth and depth of data gathered by the two on-site observers, who acted as classroom research associates by documenting in rich detail the instruction experienced by study students.

Table 4-4.
Number of Observations Conducted, by Grade and Year

Grade *(No. of Teachers*)*	Number of Observations Per Teacher	Percent of Teachers Observed 1-3 Times	4-6 Times	7-8 Times	9 Times
1997-1998					
5 *(13)*	5-9	0	38	0	62
6 *(12)*	3-9	25	33	8	33
7 *(10)*	3-9	20	40	0	40
1998-1999					
6 *(12)*	5-9	0	33	25	42
7 *(12)*	5-9	0	42	25	33
8 *(10)*	2-9	20	40	30	10
1999-2000					
7 *(9)*	2-9	22	11	11	56
8 *(9)*	8-9	0	0	44	56

*Includes teachers who taught portions of the school year

Treatment Fidelity

The research team did not interfere with the ways in which MiC or conventional curricula were implemented during the study. Rather, we wanted to document the decisions teachers made regarding the curriculum and the factors that influenced these decisions. We examined the number and content of the MiC units taught by MiC teachers, the number and content of chapters taught by teachers using conventional curricula, and the integration of supplemental programs and materials by all teachers. We use the term treatment fidelity to describe the extent to which the use of curricular materials promoted the intended philosophy of the curriculum, whether it is MiC or a conventional textbook.

In MiC the ten units designed for each grade level included the study of concepts in the domains of number, algebra, geometry, probability and statistics. When nonstudy teachers used prepublication units, they rarely were able to teach ten units in one school year. Because none of the teachers in the study taught MiC for an entire school year, we anticipated that teachers would teach 6-8 units from various content strands during one school year. During the study, considerable variation was found in the number of MiC units and the content of units taught by each teacher during each year of data collection (see Table 4-5). Teachers taught from one to nine units during each year of data collection. Some teachers used units from more than one grade level or chose to emphasize one or two content strands, particularly number and algebra. The amount of time devoted to teaching a particular unit also varied from the prescribed 3-4 week period up to 10 weeks per unit. In some cases, units were not taught in the sequence recommended in teacher support materials. District 3 fifth-grade teachers, for example, sequenced units on the basis of themes that were common across all subject areas. District 4 teachers determined the content they wanted to address during a particular school year and taught units that included that content.

Table 4-5.
Number and Content of Units Taught by MiC Teachers, by Grade and Year

Grade *(No. of Teachers)*	Number of Units Taught	Units at Grade Level *(%)*	Content of Units Taught *(%)* Number	Algebra	Geometry	Statistics/Probability
			1997-1998			
5 *(16)*	4-8	100	38	14	30	19
6 *(12)**	3-7	79	32	41	23	5
7 *(9)**	3-7	63	38	30	30	2
			1998-1999			
6 *(16)*	2-7	77	40	37	17	4
7 *(13)*	3-9	62	33	38	22	7
8 *(8)*	4-6	63	18	40	34	8
			1999-2000			
7 *(11)*	1-8	57	35	39	27	0
8 *(11)*	3-7	61	14	49	31	6

*Excludes teachers who withdrew from the study

The variation in the number of units taught in a given year was influenced by several factors. In District 1, teachers regularly integrated programs that encouraged students to use mathematics. Competitions included the 24 Game (Heath, Observation 5/6/98; Hodge, Observation 3/24/00; Kipling, Journal entry 11/20/97; McLaughlin, Journal entry 3/20/98; Mitchell, Observations 10/20/97, 4/8/98, 5/26/98; Parsons, Observation 4/16/99), the state mathematics league (Fulton, Journal entry 9/27/97; Krittendon, Journal entry 2/2/98; Lovell, Journal entry 3/5/98; Parsons, Observation 2/7/98), and a game based on the stock market (LaSalle, Observation 3/17/98; Lovell, Journal entry 3/5/99; St. James, Journal entry 1/16/99). In District 2, mandated school initiatives, such as 30 minutes of class time devoted to silent reading in every class period in all middle schools and class periods devoted to using mathematics in the workplace (Hirsch Metro Middle School), affected the amount of class time devoted to MiC. In both districts, teachers added projects because they received grant awards for particular programs such as constructing kites with anti-tobacco messages (Sixth-grade, MiC, Redling, Observation 4/13/00). Other MiC teachers used computer-assisted drill-and-practice programs **(**Brown, Observations 9/22/98, 11/9/98, 2/22/99; Broughton, Observations 11/2/99, 12/6/99; Carlson, Observation 9/15/99**).** A teacher using a conventional curriculum used brain teaser puzzles and games to reinforce basic skills at the end of each grading period (Sixth grade, conventional curriculum, Krittendon, Journal entry 2/6/98). In other situations, teachers added a project that coordinated with the content in MiC units, as one teacher's comments illustrate:

> I added this project because I thought it would give students the opportunity to work on a related project that got them out of the unit for a day or two. I really feel that they are doing a very good job in this unit but feel that if we go page by page every day they get restless. I wanted to add something that would give them the feeling of a change but still focusing on graphs and weather. (Sixth grade, MiC, Weatherspoon, Journal entry 4/3/98)

Students in seventh-grade classes constructed mobiles of three-dimensional solids before studying about them in the geometry unit *Packages and Polygons* (Kindt., Spence, Brinker, & Burrill, 1998; Seventh grade, MiC, Heath, Journal entries 9/4/97, 9/18/99).

Despite four eighth-grade MiC units devoted specifically to algebra, District 2 eighth-grade teachers stopped using MiC or used MiC in combination with a conventional high school algebra textbook midway through the second semester, as illustrated in comments from one teacher:

> The Algebra I text will be emphasized more for the remainder of the year. The senior high school that most of my students will be going to next year is very traditional, and I want to give them a head start on next year's work. (Eighth grade, MiC, Dillard, Journal entry 4/14/00)

Another teacher was concerned that students would not be prepared for high school algebra if she did not teach from a conventional text:

> I'm doing algebra not out of MiC, as the kids from previous years have come back to teachers here and said when they got to high school they were lost. Their teachers use only traditional textbooks, and they had trouble keeping up. I feel if they go into ninth grade with a foundation of algebra skills, they will be ready or above the other kids. (Eighth grade, MiC, Carlson, Journal entry 3/15/00)

In Districts 1 and 2, eighth-grade MiC teachers chose to integrate a resource that provided manipulatives for solving equations (Carlson, Observations 9/15/99, 4/10/00, Journal entry 4/12/00; Dillard, Journal entries 5/28/99, 6/2/99; Reichers, Journal entries 12/12/98, 12/16/98).

Another factor that influenced variation in implementation of MiC was fifth- and sixth-grade teachers' use of traditional algorithms for fraction operations before teaching *Some of the Parts* (van Galen, Wijers, Burrill, & Spence, 1997) and *Fraction Times* (Keijzer, van Galen, Gravemeijer, Shew, Cole, & Brendefur, 1998) which were designed to build on students' informal knowledge of fractions and to develop conceptual understanding of operations with fractions. The developers of MiC expected that students would come into fifth grade with a basic recognition and understanding of benchmark fractions such as ½ and ¼. Yet teachers felt that students needed to understand procedures for fraction computation before MiC introductory number units, as reflections from District 1 teachers illustrate:

Students were given a brief introduction to adding fractions with algorithms. Then, adding using fraction strips in *Some of the Parts* was emphasized because some students had some difficulty at first. After modeling and discussing different combinations of fractions, the students seemed to understand much better. (Fifth grade, MiC, Mitchell, Journal entry 11/14/97)

> My students have very little background knowledge when it comes to fractions. I added a lesson when teaching *Fraction Times* on how to find the fraction of a number, which was very lengthy. Then I gave a practice sheet for homework. I felt that this additional lesson was essential to students' understanding of what they were being told to do in the unit. (Sixth grade, MiC, Weatherspoon, Journal entry 1/14/98)

In District 3, fifth-grade teachers taught chapters on fraction operations from a conventional textbook before teaching *Some of the Parts*.

In sixth- and seventh-grade classes in District 1, teachers used a conventional curriculum when students were unruly, as noted by a seventh-grade teacher:

> I began using a textbook. I'm having disciplinary problems during classes in which we do so many group activities. I feel that students need a mix of the textbook and MiC. (Seventh grade, MiC, Lawton, Journal entry 9/14/99)

Thus, considerable variation in teachers' fidelity to the two treatments was documented with respect to the use of MiC units, integration of supplemental programs and materials, and pedagogy. In some cases, MiC students used conventional instructional materials and experienced instruction more aligned with conventional pedagogy. On the other hand, some teachers using a conventional curriculum supplemented conventional texts with problem solving and developed instruction that promoted conceptual understanding. In other words, some students using conventional curricula experienced mathematics curriculum and instruction that were more reform-oriented. The variation in treatment fidelity underscores the need to examine more than student achievement scores in studying the impact of MiC and other standards-based curricula. Because teachers implement curricula in different ways, descriptions of the alignment of teachers' use of MiC and other study teachers' use of conventional curricula with respective curricular philosophies are critical in comparative research.

Professional Development

Two types of professional development were provided as part of our ongoing work with teachers: summer institutes and scoring institutes. In addition, teachers had access to a toll-free telephone number for the research center and professional development opportunities provided by the districts that were open to all teachers of mathematics.

Summer Institutes

Four-day on-site summer institutes were held in Districts 1 and 2 during August prior to each year of data collection. These institutes were conducted for several reasons. First, the institutes allowed the research team to establish rapport with teachers and district personnel. Second, the institutes provided occasions to explain details about the study and provided opportunities for teachers to review all study instruments. Most importantly, the institutes provided incentives for districts and support for study teachers. The faculty for the 1997 summer institutes were the principal investigator (Tom Romberg), the research coordinator (Mary Shafer), the senior researcher on the project (Norman Webb), and three graduate project assistants (Jon Davis, Lesley Wagner, and David Webb). Two researchers from the Freudenthal Institute were also among the faculty (Nanda Querelle, an author of MiC units, and Truus Dekker), both of whom were involved in writing study assessments. A member of the MiC development team (Beth Cole) assisted in the design and implementation of the institutes; support for this work was provided by Encyclopaedia Britannica (the publisher of MiC). Most study teachers, on-site coordinators, district-level contacts, and the on-site observer in each district participated in the summer institutes.

The first day was devoted to the importance of conducting longitudinal/cross-sectional studies and information about the study design and data collection instruments. On the remaining days, concurrent workshops were held with teachers who used conventional curricula and teachers who were using MiC. All teachers reconvened as one group on the final afternoon of the institute, which was devoted to timelines for administration of initial study instruments, scheduling observations and interviews with on-site observers, and review of the teacher compensation package.

Institutes for teachers who used conventional curricula focused on alignment of curriculum, instruction, and assessment. The institutes began with a discussion of why the NCTM *Standards* were developed, the importance of all students developing mathematical power, and two working assumptions about the changes envisioned in the *Standards*: that teachers are the key figures in changing the ways in which mathematics is taught and learned in schools and that these changes require that teachers have long-term support and adequate resources to promote professional growth. The alignment of curriculum, instruction, and assessment was presented through a discussion of the effects of the state mathematics framework on the content they taught and on their instruction. Another focus of these institutes was classroom assessment practice. Teachers viewed two videos of classroom interaction. After the first video, teachers discussed opportunities for students to understand and engage in proofs, the importance of lesson summaries, classroom interaction that attends to individual needs, and evaluation of teaching. After the second video, teachers talked about using examples with physical materials, ways to generate meaningful class discussions, and procedural vs. conceptual understanding. Teachers also learned about assessment of student understanding through nonroutine tasks and student responses to such tasks. They discussed intersections among scoring, feedback, grading, and evaluation of teaching and the possibilities for assessment through the use of multiple-day and long-term assessment projects. The institutes closed with discussion of the important elements to consider when selecting or creating multidimensional assessment tasks.

For teachers using MiC, the research team selected major themes or substrands as the focus of the selected activities for each content strand. Teachers worked through student lessons in cooperative groups, learned about methods for generating classroom discussions, and talked about linkages among the lessons. The philosophy of MiC was illustrated with the algebra strand. Lessons from three substrands were used in these sessions—patterns and regularities [through *Patterns & Symbols* (Roodhardt, Kindt, Burrill, & Spence, 1997), *Building Formulas* (Wijers, Roodhardt, van Reeuwijk, Burrill, Cole, & Pligge, 1998) and *Patterns & Figures* (Kindt, Roodhardt, Spence, Simon, & Pligge, 1998)], restrictions[through *Comparing Quantities* (Kindt, Abels, Meyer, & Pligge, 1998)], and graphing [through *Ups & Downs* (Abels, de Jong, Meyer, Shew, Burrill, & Simon, 1998)]. Work in the number strand focused on developing rational number concepts [through *Some of the Parts* (van Galen, Wijers, Burrill, & Spence, 1997), *Measure for Measure* (Gravemeijer, Boswinkel, Meyer, & Shew, 1997), *Fraction Times* (Keijzer, van Galen, Gravemeijer, Shew, Cole, & Brendefur, 1998), and *Cereal Numbers* (Abels, Gravemeijer, Cole, Pligge, & Meyer, 1998)]. Teachers also discussed MiC ancillary materials [*Number Tools* (Van Galen, van den Heuven-Panhuisen, & Pligge, 1998)] to support students' thinking in lessons related to rational number. Teachers learned about MiC end-of-unit assessments through scoring and discussing student responses on the assessment for *More or Less (*Keijzer, van den Heuvel-Panhuizen, Wijers, Shew, Brinker, Pligge, Shafer, & Brendefur, 1998). Geometry sessions used lessons from both substrands—orientation and navigation [through *Side Seeing* (Jonker, Querelle, Clarke, & Cole, 1997), *Figuring All the Angles* (de Lange, van Reeuwijk, Feijs, Middleton, & Pligge, 1997) , and *Looking at an Angle* (Feijs, de Lange, van Reeuwijk, Spence, & Brendefur, 1998)] and shape and construction [through *Reallotment* (Gravemeijer, Pligge, & Clarke, 1998)]. Sessions on the statistics strand emphasized statistics [through *Dealing with Data* (de Jong, Wijers, Middleton, Simon, & Burrill, 1998), *Statistics and the Environment* (Jonker, Querelle, Wijers, de Wild, Spence, Fix, Shafer, & Browne, 1998), and *Insights into Data* (Wijers, de Lange, Shafer, & Burrill, 1998)] and chance and probability [through *Take a Chance* (Jonker, van Galen, Boswinkel, Wijers, Simon, Burrill, & Middleton, 1997) and *Great Expectations* (Roodhardt, Wijers, Cole, & Burrill, 1998)]. Teachers also discussed ideas for homework when teaching MiC. Throughout the institutes, teachers openly used multiple ways to solve problems and talked about explicit and implicit concepts in the selected lessons. Teachers raised and discussed a variety of issues including the conflict between mastery of concepts over time and grading; representations to support reasoning; prerequisite knowledge for using MiC number units; managing students who work at different paces; meaningful feedback during whole-class discussions; and calculator use. During the summer institutes, the research team learned about special conditions for teachers in each district such as daily 90-minute class periods for MiC teachers in Districts 1, and 120-minute class periods that met five times in a two-week period for MiC teachers in District 2. Teachers who had some experience teaching MiC prepublication or field-test units readily offered advice to other teachers, and sample lesson arrangements were described for 120-minute class periods.

The research team also supported district efforts to reform mathematics education. For example, in District 1 the district held a press conference with Professor Romberg. Reporters also interviewed a fifth-grade MiC teacher, who was featured in the local newspapers. Video clips of institute sessions were shown on the evening newscasts. Teachers were excited about the attention they received in their community because of their willingness to change mathematics curriculum and instruction.

In addition to working with teachers during the institutes, project assistants worked with on-site observers on interrater reliability. Using videotaped lessons, the assistant and observer independently used the observation instrument, and discrepancies in ratings were discussed until agreement was reached. Opportunities to work toward rater agreement were continued during numerous observations of study teachers for one week in September.

In 1998, the research coordinator held half-day institutes for principals of participating schools in Districts 1 and 2. The institutes began with an overview of the study goals and research design. Student work from study assessments and findings from the initial review of the student attitude inventory were shown to illustrate some preliminary differences between students who used MiC and students who used conventional curricula. Results from the observation data that were statistically significant in favor of classes using MiC were illustrated with examples (development of conceptual understanding, the nature of student conjectures, connections among mathematical ideas and between mathematics and students' lives, and the nature of student explanation, sharing of multiple strategies, conversation among students, and students' collaborative working relationships). Differences in implementation of MiC were described. For instance, some teachers prepared their instruction thoroughly, allowed students to think about mathematics, realized the importance of students working together in mathematics class, completed 6-7 units, and were less reliant on conventional practices. On the other hand, some MiC teachers showed students how to do unit problems, emphasized independent work, and completed 3-4 units. Conclusions related to professional development for MiC teachers were also shared, including the need for increased planning time for teachers to work through the units alone and with others at school and through sessions supported by the publisher, support staff for assistance and encouragement, and follow-up sessions after teaching MiC for the entire school year. Principal institutes ended with the introduction of the new MiC parent materials and plans for data collection in the upcoming school year.

The faculty for the 1998 summer institutes for teachers included the principal investigator, the research coordinator, and one graduate project assistant (Lesley Wagner). A member of the MiC development team (Beth Cole) assisted in the design and implementation of the institutes; support for her work was provided by Encyclopaedia Britannica, who also provided the services of a teacher consultant for the institute in District 1 and refreshments for both institutes. Most teachers for the second study year, some teachers from the first year, on-site coordinators, district-level contacts, and the on-site observer in each district participated in the summer institutes.

On the first morning of the institute, new study teachers learned about the study goals and research design, and all teachers learned about preliminary study results (the same information shared during the principal institute). Concurrent workshops were then held with teachers who used conventional curricula and teachers who were using MiC. All teachers reconvened as one group on the final afternoon of the institute, during which the observation instrument was reviewed, the importance of consent letters was discussed, and observations and interviews were scheduled by on-site observers.

The themes of the institute for teachers using conventional curricula were authentic instruction, tasks, and assessments (Newmann, Secada, & Wehlage, 1995) and research on teaching and learning mathematics with understanding (Carpenter & Lehrer, 1999). Teachers read and discussed the resources and synthesized the ideas into concise statements to be used with their students (see

Table 4-6). The summary was later enlarged to poster size for displaying in their classrooms. In District 2, only one teacher using a conventional curriculum attended the institute; she chose to participate in the institute for the MiC teachers.

Table 4-6.
Poster Created by Teachers Who Used Conventional Curricula, Summer Institute, District 1, 1998

Expectations for Classroom Interaction in Mathematics

Organization of Information
We will work with complex information, evaluate and interpret it, and use it to solve problems. We will organize the information in ways that enable us to find patterns, relationships, and conclusions.

Consideration of Alternatives
We will write and discuss various ways to look at and solve problems. We will listen to other students' points of view, seek to understand them, raise questions, and make comparisons between various perspectives.

Showing Understanding of Mathematical Ideas
We will learn about and understand important topics in mathematics.

Use of Mathematical Strategies
We will learn different methods to organize and evaluate information. We will use drawings, tables, graphs, and operations with numbers to develop conclusions.

Communication of Mathematical Ideas
We will clearly explain our ideas and conclusions in written and spoken forms so that another person will understand our thinking.

Connections between Mathematics and the World
We will investigate the way mathematics is encountered in various life situations.

Communication of Mathematics beyond the Classroom
We will use mathematics to communicate with others beyond the classroom in an informative or persuasive way.

Commitment to Quality
We will maintain our commitment to high quality, in-depth thinking in mathematics.

The institutes for study teachers using MiC began with a session during which teachers raised and discussed issues related to implementing MiC (see Tables 4-7 and 4-8). Teachers in both districts recognized that planning was imperative in teaching MiC, as opposed to the limited planning when they taught conventional curricula. The issues mentioned by teachers in District 1 were more related to the units themselves such as understanding the mathematics, unit goals, and vocabulary. Relative to pacing and time management, they were concerned with identifying the most important concepts and questions in each lesson, portions that could be omitted, and items for homework. Issues such as cooperative groups, assessment, and working with parents were discussed but plans of action for developing effective strategies were in the beginning stages. The issues raised by teachers in District 2 were more concerned with pedagogy relative to teaching entire lessons, working toward the quality of student responses, effective use of cooperative groups, and homework assignments related to the lesson but from ancillary resources. Teachers had clear ideas for assessment and for work with parents.

The institutes for MiC teachers continued with a demonstration lesson on geometry. In District 1, this lesson was presented by a seventh-grade study teacher. Other sessions included attention to the orientation and navigation geometry substrand, the bar model and its connections to other mathematical models in the number strand, the patterns and regularities and restriction substands in algebra, and the critical attitude theme in the statistics strand. These sessions were different from the 1997 institutes in that teachers at particular grade levels traced development of particular content for each strand, and teachers presented the findings for their grade level to the larger group. Teachers also discussed possibilities for informal classroom assessment, homework, and the newly developed parent materials. At the end of each institute, teachers and consultants worked with feeder-pattern groups to list units taught in the previous school year and to suggest order of teaching MiC units for the current school year. Copies of the list were given to each teacher and on-site coordinators.

In addition to the principals' institutes, the research team supported district efforts to study the impact of changes in mathematics curricula. For example, in District 1 the research coordinator shared preliminary study assessment results during a meeting with the superintendent, the district curriculum supervisor, mathematics specialist, and assessment team. During the meeting, one eighth-grade study teacher talked about the benefits of using MiC and the emphasis on algebra in eighth-grade units.

Table 4-7.
MiC Teachers' Discussion of Implementation, Summer Institute, District 1, 1998

Planning and Time Management

Teachers found that it was important to follow the sequence of units suggested in teacher support materials.

Teachers realized the importance of "sticking to" the units since future concepts were dependent on the concepts developed in them.

Mastery of concepts over time involved trusting that concepts would eventually be formalized.

Teachers realized the importance of working through lessons before teaching them so that they would "know where the unit is headed."

Concepts "sneak up" on students and even teachers.

Some concepts show up earlier in MiC than in traditional textbooks.

Teachers found it helpful to make a checklist of the concepts and vocabulary that was encountered in each unit. They used unit goals and objectives to help in this process.

Teachers felt that time was needed for "basic skills" when students needed to use concepts they had not previously studied.

Teachers needed to determine the important concepts and questions in each lesson and the parts of the lesson they could "cut off."

Organization beforehand assisted teachers with management. Options for homework could be decided if they did not get as far in the lesson as anticipated.

Other Issues and Ideas

Fifth-grade teachers thought it would be better to begin the school year with *Patterns & Symbols* rather than *Side Seeing*.

Cooperative groups could be effective, but teachers must prepare students to work well together.

Parents expressed concern about the absence of "traditional" work. Teachers decided to supplement units, not replace them, with such work.

Teachers helped students organize their work by using math notebooks (3-ring binders) that included student book and labeled responses to unit questions. The notebooks were also used during parent conferences.

Teachers expressed the need for formal assessments replete with scoring guides.

When teaching multiple classes, storage of books and distribution of books were problematic.

For students who were absent, teachers kept a checklist of the lessons on the board.

Teacher logs (for the study) were useful to teachers in reflecting about their lessons, but they were time consuming to complete.

Table 4-8.
MiC Teachers' Discussion of Implementation, Summer Institute, District 2, 1998

Planning
Teachers realized that they must go over lessons in advance. They cannot "wing it."
Teachers felt that every minute of class time was valuable. Each lesson took a full 60 minutes.

Pedagogy
Teachers realized that modeling explanations was crucial. They established guidelines for good answers and graded answers together as a class.
Teachers felt that group work was important, but they helped students work effectively by

- mixing students who spoke different languages;
- having students work independently first before sharing in the group;
- recognizing that group work need not be the sole method of instruction, that using a variety of strategies was important;
- knowing that some questions facilitate "real" discussion;
- knowing that discussion will be useful when students ask particular types of questions; and
- establishing clear expectations for group work.

When students finished their work earlier than others, teachers asked them to assist other students.

Other Ideas
For assessment purposes, teachers used questions in section summaries and Try This!.
Teachers assessed students at the end of each section rather than at the end of each unit.
Teachers felt that *Number Tools* (MiC ancillary materials) was useful for homework.
Teachers noted that being proactive with parents was important. They did this through

- grade-level meetings to show parents MiC units;
- letters to parents; and
- open house nights during which parents completed sample lessons.

Teachers completed teaching logs (for the study) every day before leaving school.

The faculty for the 1999 summer institutes for teachers included the research coordinator and one graduate project assistant (Lesley Wagner). In District 1, a member of the MiC development team and assessment specialist from the Freudenthal Institute (Meike Abels) assisted in the design and implementation of the institutes; support for her work was provided by Encyclopaedia Britannica, who also provided the services of a professional development consultant (Sheldon Fine) for the institute in District 2 and refreshments for both institutes. Teachers from the second and third study years, on-site coordinators, district-level contacts, and the on-site observer in each district participated in the summer institutes.

On the first morning of the institutes, new study teachers learned about the study goals and research design, and returning teachers reviewed some interesting student work from the Problem Solving Assessments. In preparation for the 1999 summer institutes, study teachers were asked about the content they wanted to pursue. Teachers who used MiC requested information on classroom assessment and time for preparing to teach units that were unfamiliar to them. Teachers who used conventional curricula asked to learn more about informal classroom assessment and creating assessment tasks and scoring rubrics. Two concurrent institutes were planned around those themes. One group of teachers participated in discussions about what to assess and how to assess; assessments that attended to the three levels in the Dutch assessment model; grading such assessments; designing assessment tasks; and developing scoring rubrics for those tasks. The other teachers worked together to solve unit problems, determine multiple strategies for problems, select problems appropriate for classroom assessment and homework, and discuss end-of-unit assessment tasks. They generated a list of “time savers” in teaching MiC. For example, teachers felt it was important to read the teacher guide, do all the problems before teaching, and plan a time line for teaching. They talked about avoiding the temptation to delve into the context at the expense of the mathematical content and trying not to tell the students everything they themselves knew about the topic. Teachers decided to give students a set number of problems to do at one time and to establish time limits when they worked in small groups. They also realized that they did not have to discuss every problem during class discussion. They decided to focus on key items that involved mathematical concepts and discussion about various strategies.

Teachers in the group that worked on planning joined the other group for joint sessions on assessment. Teachers worked individually or together to create and refine assessment tasks. For example, teachers in District 1 created tasks using the context of the drought that was occurring in their state (see Table 4-9). Teachers discussed the variety of mathematical content necessary to complete the tasks and the ways that additional information in Items 3 and 4 created distraction from the serious nature of the water loss in the reservoir. They decided to keep the mathematical focus on fair representation of the data and reworked the task with that emphasis. During the institute, teachers brought in three examples of general scoring rubrics to discuss and compare: one strictly focused on mathematics in student responses; one expressed in “kid-friendly” language with attention to mathematical understanding, thinking, and communication; and one attending to the correctness of the solution, the formulation and execution of a solution plan, and the clarity of explanation. Teachers scored a set of twenty student assessments from a task on the end-of-unit assessment for *Building Formulas* (Wijers, Roodhardt, van Reeuwijk, Burrill, Cole, & Pligge, 1998) according to their own scoring principles. The results were compared during an intense discussion of the consistency among scorers, the types of items that should receive more score points, attention to what students were trying to convey in their responses, and translating the scores into grades. All teachers were encouraged to record their thoughts and comments about characteristics of good assessment problems and things to remember when scoring assessments. The list was discussed and distributed to each teacher (see Table 10). All teachers convened as one group on the final afternoon of the institute, during which the teacher log was reviewed, the importance of consent letters was discussed, and observations and interviews were scheduled by on-site observers.

Table 4-9.
Example of Teacher-Created Assessment Tasks, Summer Institute, District 1, 1999

Original Task Based on Consecutive Reports in the Local Newspaper
The Drought of 1999
The graphs below show the amount of water available at the local reservoir on Monday, August 9 and on Tuesday, August 10.

1. Find the percent loss of water in the reservoir from August 1 to August 10. Show your work.
2. Do you think that the graph shows this loss in a fair way? Why or why not?
3. The reservoir lost 10,000,000 gallons of water in just one day. If there are approximately 7.5 gallons of water in one cubic foot, estimate the number of swimming pools that could be filled using that amount of water. Explain and show all work.
4. Each time you take a shower, you use about 25 gallons of water. Assume that you only take on shower a day. Estimate the number of years it would take you to use the amount of water lost in the reservoir from Monday to Tuesday. Show your work and explain your reasoning.

Revised Task
The three graphs below show information about water use and the amount of water that was available on August 11, 1999 during the drought.

1. Which graph represents its data most accurately? Explain your answer.
2. Choose one of the graphs and explain how you would change it so that it would show the information more accurately. Explain why you made the changes. Make a sketch of your new graph.

Table 4-10.
Teacher-Generated List of Characteristics of Good Assessment Problems and Things to Remember When Scoring, Summer Institute, District 1, 1999

Characteristics of Good Assessment Problems
Designed for all students to experience success
Allowance for multiple solution strategies
Items independent of success on other items
Assessment of what was taught (unit/chapter goals)
Reflection on mathematics learned, not just processes
Use of engaging context, if possible
Items balanced in level of reasoning (using the Dutch assessment pyramid model)

Things to Remember When Scoring
Assign most of the score points to items assessing Level 1 reasoning
Scoring items assessing Level 3 reasoning is difficult
Grading takes time and a lot of thought
Scoring should be an open and clear process

Scoring Institutes

The research staff also conducted scoring institutes for study teachers. The purpose of these two-day scoring institutes was for teachers, on-site coordinators, on-site observers, and principals to learn about the construction and scoring of the Problem Solving Assessments, which were administered to study students in the spring of each year of data collection. On-site scoring institutes were held in May or June of each study year in Districts 1, 2, and 4, in May of the first two years of study participation in District 3, and at the site that pilot-tested eighth-grade assessments during the summer of 1998.

Some items from each grade-specific Problem Solving Assessments were scored during the on-site scoring institutes. For each context, teachers solved the problems. The presenter then reviewed the scoring rubric and strategy codes and used student work samples to discuss appropriate scores and coding. In preparation for the institutes, covers were removed from assessment booklets, and all booklets from a particular grade level were shuffled and separated into packets of ten booklets. After each teacher scored a packet, the packets were redistributed to different teachers for a second scoring. Discrepancies in scoring and coding were resolved through adjudication. The entire process was repeated multiple times during the institutes. (See Chapter 1 of Monograph 4 for detailed information about the scoring institutes including studies of interrater reliability.) During scoring, study teachers noticed differences in the ways students expressed their thinking, and they talked about how they might work with their students to develop more complete responses. Teachers commented on the value of this experience, as illustrated in the following excerpts from their feedback on the institutes:

> It was very interesting. It's such an eye opener to see student work. I really enjoyed the opportunity to see problems from all three grade levels. (Eighth grade, MiC, Wells, Personal communication 5/99)
>
> Very beneficial in understanding how students think about solving math problems. I always look forward to it! (Seventh grade, MiC, Perry, Personal communication 5/99)
>
> You should come and do this 3-4 days, not just 1-2 days. (Sixth grade, conventional, Friedman, Personal communication 5/99)

Some teachers also requested opportunities to learn about the process of developing scoring rubrics for MiC end-of-unit assessments (e.g., Sixth grade, MiC, Schlueter, Personal communication 5/99). Teachers in District 3 outlined their goals for using MiC during the next school year and listed their requests for professional development opportunities:

- Opportunities for fifth- and sixth-grade teachers to observe other MiC teachers;
- Development of assessments for classroom use, with accompanying scoring rubrics;
- Homework materials that parents can understand;
- Alignment of MiC with the state curriculum guide;
- Time allotment for teaching each unit;
- From the test results, what are the weaknesses of students?
- When is it appropriate to add to MiC units? To reinforce?;
- Information about working with a range of students in the same class; and
- Time to discuss the transition from fifth grade to sixth grade.

The research team also supported district efforts to change mathematics curriculum and instruction. For example, in District 3 a graduate project assistant answered questions from parents at the May school board meeting. During the scoring institutes in 2000, the research coordinator shared preliminary results from the External Assessment System and the Problem Solving Assessment with district administrators in Districts 1, 2, and 4. One principal in District 1 requested a special meeting with the research coordinator to discuss findings related to instruction and the factors that contributed to the differences noted.

Student assessments that were not scored at on-site institutes were scored during the summer and fall of 1998 and the summers of 1999 and 2000 at the Wisconsin Center for Education Research. Study teachers were invited to travel to Madison for scoring assessments for one of the scoring institutes in Madison in 1998; one teacher from each of Districts 1, 2, and 3 participated. The remaining raters were elementary- and middle-school teachers from Madison-area schools who were recommended by local district mathematics coordinators, local MiC pilot- and field-test teachers, and teachers who were participating in other projects at the research center. The scoring institutes provided a significant professional development opportunity for teacher-raters who commented that they made changes in their pedagogy to emphasize mathematical communication, included lessons that promoted more complex reasoning, and integrated various types of problems designed to elicit student thinking at more complex levels in their classroom assessment practice.

Additional Contacts

Teachers had access to a toll-free phone number to talk with members of the research team or members of the MiC development team. Teachers felt this was a beneficial resource, as illustrated by a fifth-grade teacher:

> [Finding the measures of exterior angles of polygons] was difficult for the students. I must admit, I was a little confused myself in the beginning. It did help talking to [a member of the research team]. (Fifth-grade, MiC, Mitchell, Journal entry 2/27/98)

Other teachers contacted the research team about guidelines for selecting computer software (Fifth grade, MiC, Fiske, Personal communication 2/26/98) and to learn about alternate ways to approach teaching integers (Seventh grade, conventional curriculum, Stark, Personal communication 2/98).

District Professional Development Opportunities

Professional development opportunities varied in scope and frequency in each study district. In District 1 during the first year of data collection, the district mathematics specialist arranged focus group meetings for all teachers who were implementing reform curricula. Each month teachers explored general pedagogical issues including student-centered instruction, assessment, and use of mathematical tools such as the ratio table. These meetings were held after school hours, and teachers were compensated by the district for their participation. During subsequent years, however, the focus meetings were not held, which led to distinct challenges for three MiC teachers who were new to the study. Although these teachers had taught mathematics for many years, this was the first time they taught MiC. Two of these teachers requested support beyond what the study provided in the summer professional development institutes. Because Gollen was the only sixth-grade mathematics teacher in her school, she did not have the opportunity to collaborate with other teachers who were implementing MiC. The on-site observer commented that Gollen had difficulty discerning the alignment of lessons with the unit goals, presenting lesson content, and orchestrating classroom discussions:

> After every class, Ms. Gollen asks me how she could have taught the lesson better. She needs mentoring on a regular basis. She needs someone to go over each unit and lesson with her before she teaches it to her students. Many other teachers that I observe also need the same thing. (Diver, Observer comments for sixth grade, MiC, Gollen 3/18/99)

Gollen reported that she felt support from the observer, even though she did not act as a mentor: “Ms. Diver was here to observe. She gives me support for my program. I feel that we’re on the right track. The students understand and like the work” (Sixth grade, MiC, Gollen, Journal entry 11/19/97). Reichers, an eighth-grade MiC teacher, requested that a consultant teach her class while she observed the lesson, a request that was accommodated by the district mathematics specialist. The teacher found this to be a worthwhile experience that influenced her teaching. Reichers also welcomed visits from the on-site observer:

> I really have trouble evaluating how things are going sometimes. From her comments I realized that the lesson went a whole lot better than I thought. It was so helpful to me and the kids. (Eighth grade, MiC, Reichers, Journal entry 10/1/98)
>
> It seems the students really need to work in pairs. It's worth being in the study just to have her here once a month (Eighth grade, MiC, Reichers, Journal entry 2/4/98).

Another situation occurred in the third year of data collection. St. James, the seventh-grade mathematics teacher in the same school as Gollen, had been using a conventional curriculum during the first two study years. But in order to follow students from Gollen's class, St. James was asked by district personnel to teach one section of MiC. St. James found teaching MiC difficult: "I feel MiC is a good course, and I enjoyed the units. But I'm not as on top of it as in the prior two years when I taught from the textbook" and "This study was enjoyable, even this year which was the hardest for me. I felt sometimes like I was in the middle of an ocean treading water" (Seventh grade, MiC, St. James, Journal entries 12/22/99 and 5/00, respectively). St. James taught MiC in conventional ways (see Monograph 3). Although all three of these teachers participated in the professional development institutes provided by the study, MiC teachers clearly needed ongoing support and sustained professional development opportunities.

In District 2, teachers had numerous possibilities for professional development. Each school was given six early-release days for general professional development. In addition, each school received 10 substitute days for professional development in mathematics and/or science, 12–18 days of in-service days in mathematics provided by (USI or Eisenhower) government funding (each involving 2–6 teachers), and 3–5 days of districtwide mathematics in-service. Teachers also had opportunities to participate in five days of paid in-service for mathematics during the summer. During the second and third years of data collection, MiC teachers were given one day of release time per month in order to collaborate on planning to teach MiC units. As one teacher reported, these collaborative times were effective for changing instruction:

> Interesting fact: The more comfortable I am teaching MiC, the better the kids like it. Having the MiC days to prepare, discuss, and work through the units is very beneficial. (Eighth grade, MiC, Teague, Journal entry 4/29/99)

In District 3 during the summer before the first year of data collection, teachers participated in a district-funded weeklong camp in which they looked for conceptual development across the MiC units at all grade levels, determined the sequence of teaching the units, and discussed instructional approaches for effectively teaching MiC. In addition, fifth-grade teachers met weekly before school without pay to collaborate on teaching MiC. Throughout the school year, school administrators provided paid monthly evening meetings for the teachers during which they discussed implementation issues and continued their in-depth review of specific units in preparation for teaching them. The monthly meetings continued in the second year of data collection. In the third year of data collection, the focus of meetings changed to teachers' understanding of new state requirements for standardized testing.

In District 4, professional development opportunities were provided for all mathematics teachers at both district and school levels. The district sponsored monthly one-day workshops for all teachers implementing MiC for the first time. During these sessions, groups of teachers at each grade level focused on one MiC unit, learning about the presentation of the content and discussing instructional approaches and methods of classroom assessment. In the middle school that participated in the study, the assistant principal for mathematics and science held monthly discussions with each mathematics teacher, which included such topics as reform recommendations in curriculum, instruction, and assessment; research in mathematics education; and applications of research in classroom practice.

As these examples illustrate, the support teachers felt for implementing MiC varied greatly among the districts. In Districts 3 and 4, because MiC was the primary curriculum used, more support was available for teachers as they implemented MiC for the first time. In District 2 during the second and third years of data collection, MiC teachers were given release days to collaborate on teaching MiC. The most varied levels of support were experienced by study teachers in District 1.

Conclusion

The challenges we faced in conducting comparative longitudinal research in the reality of school life seemed daunting at times. As numerous as these challenges were, however, we feel that our extensive attempts to collect quality data were successful in many ways. We were able to follow many students longitudinally in each cohort, collect rich observation data, and record teacher accounts of what transpired in study classes. Our data collection was thorough and diverse, enabling us to identify and scale information that captured variation among study participants. In the finish, teachers appreciated our feedback in response to their journal entries and our continuing efforts to support their teaching of comprehensive mathematics content to diverse student groups. District and school personnel requested and received support from the research staff in forms of presentations at teachers' meetings and meetings of parent or community organizations. These gestures of good will and support enhanced the extensive data collection for the longitudinal study.

In this chapter, various interpretations of commitment, treatment fidelity, and teachers' needs for professional development during the conduct of the longitudinal study were described. These variations draw attention to the need to study the effects of the culture in which student learning is situated when analyzing the impact of standards-based curricula. Controlling potential sources of variation, as is done in laboratory experiments, is more difficult in classroom settings. But this does not mean that comparative research cannot be done in today's schools. Rather, as this study demonstrates, impact studies of high quality can be conducted in classroom settings when data collection and analysis are designed to take into consideration the variations encountered in these settings. In later chapters, we examine in depth the dimensions that framed the analysis of the instruction students experienced, the opportunity students had to learn significant mathematics content with understanding, and the capacity of their schools to support mathematics teaching and learning.

References

Abels, M., de Jong, J. A., Meyer, M. R., Shew, J. A., Burrill, G., & Simon, A. N. (1998). Ups and downs. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Abels, M., Gravemeijer, K., Shew, J., Cole, B. R., Pligge, M. A., & Meyer, M., R. (1998). Cereal numbers. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Adams, L., Tung, K., & Warfield, G. (2000). *Middle-school curricula comparisons of Connected Mathematics Program (CMP), Mathematics in Context (*MiC*), and Singaporean books.* Report to the National Science Foundation.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. Washington, DC: Author.

American Association for the Advancement of Science. (1998). *Benchmarks-based evaluations*. (Available at http://www.project2061.org/matheva/index.htm)

Arauco, T., (1998). *Utilizing a modular database design: Analysis of a reformed middle school mathematics curriculum.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 17). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Barbeau, E. (1989, September). Mathematics for the public. Paper presented at the meeting of the International Commission on Mathematics Instruction, Leeds University, Leeds, England.

Battista, M. T. (1999). The mathematical miseducation of America's youth: Ignoring research and scientific study in education. *Phi Delta Kappan, 80*(6), 425–433.

Bay, J. M., Reys, B. J., & Reys, R. E. (1999). The top 10 elements that must be in place to implement standards-based mathematics curricula. *Phi Delta Kappan, 80*(6), 503–506.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 171–246). Chicago: Rand McNally.

Carpenter, T. P., & Lehrer, R. (1999). Teaching and learning mathematics with understanding. In E. Fennema & T. A. Romberg (Eds.), *Classrooms that promote mathematical understanding* (pp. 19–32). Mahwah, NJ: Erlbaum.

Charles, R. I., et al. (1998). *Middle School Mathematics.* Menlo Park, CA: Scott Foresman/Addison Wesley. Collins, W., et al. (1998).*Mathematics: Applications and Connections*. Lake Forest, IL: Glencoe.

Collins, W., et al. (1998).*Mathematics: Applications and Connections*. Lake Forest, IL: Glencoe.

Collis, K., & Romberg, T. (1992). *Mathematical problem-solving profiles*. Melbourne: Australian Council for Educational Research..

CTB/McGraw-Hill (1992). *California achievement test.* Monterey, CA: Author.

Davis, J., Wagner, L. R., & Shafer, M. C. (1997). *Classroom observation scale.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 6). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Dekker, T., Querelle, N., van Reeuwijk, M., Wijers, M., Fejis, E., de Lange, J., Shafer, M. C., Davis, J., Wagner, L., Webb, D. (1997–1998). *Problem solving assessment system*. Madison, WI: University of Wisconsin.

de Jong, J. A., Wijers, M., Middleton, J. A., Simon, A. N., & Burrill, G. (1998). Dealing with data. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

de Lange, J., van Reeuwijk, M., Feijs, E., Middleton, J. A., & Pligge, M. A. (1997). Figuring all the angles. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Feijs, E., de Lange, J., van Reeuwijk, M., Spence, M. S., & Brendefur, J. (1998) Looking at an angle. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures.* Dordrecht, The Netherlands: D. Reidel.

Freudenthal, H. (1987). Mathematics starting and staying in reality. In I. Wirszup & R. Street (Eds.), *Proceedings of the USCMP conference on mathematics education on development in school mathematics education around the world.* Reston, VA: National Council of Teachers of Mathematics.

Gravemeijer, K. (1994). *Developing realistic mathematics education*. Utrecht, The Netherlands: Vakroep Ondersoek Wiskundeonderwijs en Onderwijscomputercentrum, Rijksuniversiteit.

Gravemeijer, K., Boswinkel, N., Meyer, M., R., & Shew, J. A. (1997). Measure for measure. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Gravemeijer, K., Pligge, M. A., & Clarke, B. (1998). Reallotment. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Harcourt Brace. (1994). *Mathematics Plus.* San Antonio, TX: Author.

Harcourt Educational Measurement. (1996). *Stanford Achievement Test, Ninth Edition*. San Antonio, TX: Harcourt Brace Educational Measurement.

Hoover, H. D., Hieronymus, A. N., Dunbar, S. B., & Frisbie, D. A. (1996). *Iowa Test of Basic Skills*. Itasca, IL: Riverside Publishing Company.

International Association for the Evaluation of Educational Achievement. (1996). *TIMSS mathematics items: Released set for Population 2 (seventh and eighth grades).* The Hague, The Netherlands: Author.

Jonker, V., Querelle, N., Clarke, B., & Cole, B. R. (1997). Side seeing. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Jonker, V., Querelle, N., Wijers, M., de Wilde, A., Spence, M. S., Fix, M. A., Shafer, M. C., & Browne, J. (1998). Statistics and the environment. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Jonker, V., van Galen, F., Boswinkel, N., Wijers, M., Simon, A. N., Burrill, G., & Middleton, J. A. (1997). Take a chance. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Keijzer, R., van den Heuvel-Panhuisen, M., Wijers, M., Shew, J. A., Brinker, L., Pligge, M. A., Shafer, M. C., & Brendefur, J. (1998). More or less. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Keijzer, R., van Galen, F., Gravemeijer, K., Shew, J. A., Cole, B. R., & Brendefur, J. (1998). Fraction times. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Kindt, M., Abels, M., Meyer, M., R., & Pligge, M. A. (1998). Comparing quantities. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Kindt, M., Roodhardt, A., Spence, M. S., Simon, A. N., & Pligge, M. A. (1998). Patterns and figures. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Kindt, M., Spence, M. S., Brinker, L. J., & Burrill, G. (1998). Packages and polygons. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Labaree, D. (1999). The chronic failure of curriculum reform: With recurring waves of rhetoric and galcial change in practice, little that is truly new has trickled down to students. But Why? *Education Week,* May 19, pp. 42-44.

Masters, G., & Forster, M. (1996). *Progress maps*. Melbourne, Australia: Australian Council for Educational Research.

Mathematical Sciences Education Board. (1990). *Reshaping school mathematics*. Washington, DC: National Academy Press.

McGraw - Hill. (1992) *California Achievement Test.* Monterey, CA: Author.

Menard, S. (1991). *Longitudinal research*. Newbury Park, CA: Sage.

National Center for Education Statistics. (1990). *National Assessment of Educational Progress: Grade 8 mathematics assessment.* Washington, DC: Author.

National Center for Education Statistics. (1992). *National Assessment of Educational Progress: Grade 8 mathematics assessment.* Washington, DC: Author.

National Center for Research in Mathematical Sciences Education & Freudenthal Institute. (Eds.). (1997–1998). *Mathematics in context.* Chicago: Encyclopaedia Britannica.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* Washington, DC: U.S. Government Printing Office.

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

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (1995). *Assessment standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

National Science Board Commission on Precollege Education in Mathematics, Science and Technology. (1983). *Educating Americans for the 21st century: A plan of action for improving the mathematics, science and technology education for all American elementary and secondary students so that their achievement is the best in the world by 1995*. Washington, DC: National Science Foundation.

Newmann, F. M. & Associates. (1996). *Authentic achievement: Restructuring schools for intellectual quality.* San Francisco, CA: Jossey-Bass Publishers.

Newmann, F. M., Secada, W. G., & Wehlage, G. G. (1995). *A guide to authentic instruction and assessment: Vision, standards, and scoring*. Madison, WI: Wisconsin Center for Education Research.

Price, J., Roth, J.N., & Leschensky, W. (1992). *Merrill pre-algebra: A transition to algebra.* Lake Forest, IL: Glencoe.

Romberg, T. A. (Ed.). (1985). *Toward effective schooling: The IGE experience.* Washington, DC: University Press of America.

Romberg, T. (1987). A causal model to monitor changes in school mathematics. In T. Romberg & D. Stewart (Eds.), *The monitoring of school mathematics: Background papers*, Vol. 1, (pp. 63–79). Madison, WI: Wisconsin Center for Education Research, University of Wisconsin–Madison.

Romberg, T. A. (1997). Mathematics in context: Impact on teachers. In E. Fennema & B. S. Nelson, *Mathematics teachers in transition.* Mahwah, NJ: Erlbaum.

Romberg, T., & Pedro, J. (1996). *Developing "Mathematics in Context": A research process.* Madison, WI: National Center for Research in Mathematical Sciences Education.

Romberg, T. A., & Webb, D. C. (1997–1998). *External assessment system*. Madison, WI: University of Wisconsin–Madison.

Roodhardt, A., Wijers, M., Cole, B. R., & Burrill, G. (1998). Great expectations. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Roodhardt, A., Kindt,M., Burrill, G., & Spence, M. S. (1997). Patterns & Symbols. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Scott Foresman/Addison Wesley, (1999) *Mathematics.* Menlo Park, CA: Author.

Shafer, M. C. (1996). *Assessment of student growth in a mathematical domain over time.* Unpublished doctoral dissertation, University of Wisconsin–Madison.

Shafer, M.C. (1997a). *District profile*. (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 13). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C. (1997b). *School profile*. (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 14). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C. (1997c). *Teacher questionnaire: Experience teaching Mathematics in Context*. (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 8). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research..

Shafer, M. C. (1997d) *Student questionnaire*. (Mathematics in Context Longitudinal/Cross-Sectional Study Working Paper No. 2). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C. (1999). *Student questionnaire: Transition into high school* (Technical Report No. 55). Madison, WI: University of Wisconsin–Madison.

Shafer, M. C., Davis, J., & Wagner, L. R. (1997a). *Teacher questionnaire: Background and experience*. (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 9). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research..

Shafer, M. C., Davis, J., & Wagner, L. R. (1997b). *Principal interview: School context*. (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 12). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C., Davis, J., & Wagner, L. R. (1997c). *Teacher interview: Teaching and learning mathematics*. (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 4). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C., Davis, J., & Wagner, L. R. (1997d). *Teacher questionnaire: School context*. (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 10). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C., Davis, J., & Wagner, L. R. (1997e). *Teacher questionnaire: Professional opportunities*. (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 11). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C., Davis, J., & Wagner, L. R. (1997f) *Teaching log*. (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 5). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C., Davis, J., & Wagner, L. R. (1998) *Teacher interview: Instructional planning and classroom interaction*. (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 3). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C., Wagner, L. R., & Davis, J. (1997). *Student attitude inventory*. (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 7). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

U. S. Department of Education. (1999). *Reviewer evaluation instrument for mathematics programs.* Washington, DC: Author.

University of Pittsburgh & National Center on Education and the Economy. (1997). *New Standards Reference Examination*. San Antonio, TX: Harcourt Brace Educational Measurement.

van Galen, F., van den Heuvel-Panhuisen, M., & Pligge, M. A. (1998). Number tools. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

van Galen, F., Wijers, M., Burrill, G., & Spence, M. S. (1997). Some of the parts. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Webb, D. C., & Meyer, M. (2001). *Summary report of achievement data for Mathematics in Context: A Connected Curriculum for grades 5-8.* Madison, WI: Wisconsin Center for Education Research.

Webb, N., & Romberg, T. (1992). Implications of the NCTM Standards for mathematics assessment. In T. Romberg (Ed.), *Mathematics assessment and evaluation.* Albany, NY: State University of New York Press.

Weller, M. (1991). Marketing the curriculum: Core versus non-core subjects in one junior high school. Unpublished doctoral dissertation, University of Wisconsin–Madison.

Wijers, M., De Lange, J., Shafer, M. C., & Burrill, G. (1998). Insights into data. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.

Wijers, M., Roodhardt, A., van Reeuwijk, M., Burrill, G., Cole, B., & Pligge, M. A. (1998). Building formulas. In Wisconsin Center for Education Research & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago, IL: Encyclopaedia Britannica.