This paper examines conditions influencing achievement in and attitudes toward mathematics among middle-school Latino students. Observations were carried out in a rural California middle school that served a largely Mexican American student population and that was in the process of implementing a new thematic mathematics curriculum in heterogeneously grouped classrooms. The thematic approach was characterized by a hands-on emphasis, use of small-group learning activities, and the integrated use of a wide range of mathematics skills and concepts. Students were placed in three groups that received thematic instruction in heterogeneous groups, traditional textbook-based instruction in homogeneous groups, or blended instruction in heterogeneous groups. Perhaps because of the dedicated involvement of all teachers in improving mathematics instruction in this school, students in all three groups made significant normal-curve-equivalent gains in mathematics achievement on standardized tests. Achievement gains did not differ between groups, but in all groups girls made larger gains than boys. The three treatment conditions did not differ in effectiveness for low-achieving students, but high-achieving students did best in the homogeneous group. At the level of implementation reached during this first year, the thematic approach appeared to exert little influence on mathematics attitudes or students' motivation-linked self-perceptions. A primary lesson from the year was that real change in instructional approaches comes very slowly, even among teachers committed to reform. Contains 29 references and 10 figures. (SV)
Middle-School Mathematics for Students of Mexican Descent:
A Thematic Approach to Contextualization of Instruction

Ronald W. Henderson

National Center for Research on Cultural Diversity and Second Language Learning
University of California, Santa Cruz

Presented as part of a Symposium entitled Contextualizing Mathematics and Bridging Cultures: Supporting Reform Mathematics Classrooms for Latino Students, at the Annual Meeting of The American Educational Research Association, San Francisco, April 1995. The work reported is a project of the National Center for Research on Cultural Diversity and Second Language Learning. This work was supported, in part, by funding from the Office of Educational Research and Improvement (OERI) of the U.S. Department of Education, under Cooperative Agreement No. R117G10022, and in part by a grant from the California Academic Partnership Program. The findings and opinions expressed here are those of the authors, and do not necessarily reflect the positions or policies of OERI or the California Academic Partnership Program.
Middle-School Mathematics for Students of Mexican Descent: A Thematic Approach to Contextualization of Instruction

The relatively low level of school achievement among Latino students has been a source of concern among educators and policy makers and researchers (Haycock & Navarro, 1988; Henderson & Landesman, in press; MacCorquodale, 1988; Policy Analysis for California Education, 1990, Valencia, 1991). Researchers attempting to identify factors that contribute to educational vulnerability among Latino students have tended to focus on literacy issues, where instruction in reading and writing for these students has been characterized as "intellectually limiting" (Moll, 1992) because of its emphasis on "low-level" literacy. But that critique applies as well to mathematics as to reading and writing instruction. Indeed, the discrepancy between Latino and majority group performance in mathematics is especially striking for tasks that require a strong conceptual grasp of the subject matter, as contrasted with skill in computation (Dossoney, et al., 1988), and that discrepancy may result from an intellectually limiting approach to mathematics instruction for students from minority backgrounds (see Cole & Griffin, 1987), who are disproportionately represented in lower achievement tracks. In many school districts the sorting process begins in earnest at the middle school level, formalizing and hardening unequal access to mathematics curricula that expand students' educational options. Unfortunately, relatively little research has examined factors that contribute to success or failure in middle school mathematics among Latino students. Both the research agenda on mathematics learning and the mathematics reform movement have largely bypassed the Latino student population (see Henderson, Landesman, Nur & St. John, 1995; Secada, 1991).

The present report describes our attempts to identify and understand the influence of factors that may account for relatively low levels of achievement and participation in mathematics among Latino students and to explore instructional practices within a thematic curricular emphasis—thought likely to increase Latino participation and success in mathematics. Our intent was to investigate a range of sociocultural, instructional, and motivational dynamics hypothesized to influence learning outcomes in mathematics in actual classrooms for these students.

The use of a thematic curricular emphasis as the context for our exploration of mathematics learning and motivation is influenced by arguments that the problem of Latino underachievement in schools should be addressed through curricular restructuring to achieve cultural compatibility (Trueba, 1988; Moll & Diaz, 1987). For example, Mexican-American children consistently have been found to exhibit a more field dependent cognitive style than their Anglo-American peers (Garcia, 1983). The causes and implications of the field dependent/independent construct are subject to debate (Henderson, 1980), but the consequences of differences in cognitive style for instruction remains an important issue. The more field dependent style of Mexican American children may reflect what Tharp (1989) calls holistic/visual thought. Since children from cultures that employ holistic/visual patterns of cognition tend to learn in their natural environments, Tharp (1989) suggests greater curricular contextualization as a means of making instruction more culturally compatible.

If instruction is to overcome barriers to learning that result from an inadequate match between the culture of the school and the community backgrounds of cultural minority students, the schools must, as Tharp (1989) suggests, help these students by demonstrating how rules, abstractions, and verbal descriptions are drawn from the everyday world and how they are applied to it. If the fragmented and decontextualized nature of contemporary mathematics instruction poses barriers to
achievement and participation for important segments of the population, then the practices that promote integration and provide a meaningful context for learning offer a potential remedy worth trying.

Cooperative learning and heterogeneous grouping do not, by themselves, guarantee success. But they can be employed to involve students in finding, inventing, and solving problems in familiar contexts, and these procedures enable students and teachers to engage in authentic, subject-based exchanges. Small groups organized around specific tasks make it possible for assisted learning (Tharp & Gallimore, 1988) to take place, leading to the internalization of the metacognitive skills students need to regulate their own learning (Henderson, 1986, Henderson & Cunningham, 1994), and such autonomous learning is considered essential to the development of higher order thinking in mathematics (Fennema and Peterson, 1985). Learning based on concrete activities rather than textbook abstractions should enable learners to make real world sense of their school experiences. Such opportunities are likely to be especially important for students whose experience may not be highly congruent with the abstractions of schooling (Tharp, 1989). Since students in the classrooms with which we have been working do not share a common background of cultural experience, we have chosen a thematic approach to contextualization because it appears to provide a means of mapping the procedural and conceptual knowledge of mathematics onto shared experiences created within the classroom.

Our work has progressed through two discrete phases. The present paper focuses on the second phase, but a brief summary of the earlier work is presented to establish the broader context of the work because insights during Phase I informed a deliberate shift in our university-school relationship. We progressed from an approach that primarily monitored and supported exploratory implementation to one that cultivated reciprocal mentoring in a partnership for change.

During Phase I we focused on students of Mexican descent attending the target school and a neighboring school in a rural agricultural area of central California (see Henderson, Landesman, & Allen, 1994; Henderson, Landesman, Nur, & St. John, 1995). In both schools, the majority of students were of Mexican descent. While these two schools served comparable populations, they differed markedly in their professed philosophies of mathematics instruction. The target (experimental) school's administration espoused a commitment to heterogeneous grouping, although this was not completely implemented since some students were assigned to "accelerated" classes as a result of parental pressure. This administration was also dedicated (with varying degrees of teacher buy-in) to a hands-on approach to mathematics instruction and to implementing the California Mathematics Framework. Thus, the philosophy at the experimental school was consistent with a thematic approach. The comparison school used a contrasting approach that emphasized rote learning of basic facts and operations as a foundation for later mathematics.

Monitoring of actual classrooms in observations indicated that the implementation of thematic instruction varied greatly among the teachers. Between what we envisioned as thematic instruction and what actually transpired in these classrooms, we found an immense discrepancy. For example, thematic units provided opportunities to integrate many mathematics skills and concepts within the context of a set of hands-on and applied activities. This is consistent with the thematic approach that allows students to learn and reinforce mathematical skills and concepts through collaborative activities simulating real-world activities—through learning events within contexts—instead of through drills on isolated skills. However (and in spite of our urging), the common practice was to conduct formal lessons on the skills to be used in the activity (such as linear measurement) prior to permitting the activity. One of the teachers who worked most earnestly on the theme insisted that her students came to seventh grade so deficient in measurement "ruler skills" took several days, in isolation
from the theme.

Classroom observations also revealed that Phase I thematic teachers typically implemented thematic lessons and activities in traditional ways, without altering instructional practices to exploit opportunities for encouraging higher order thinking (comparison, conjecture, and exploration) and without altering classroom discourse to exploit teachable moments. They also generally chose to implement only ready-to-use thematic curriculum without modifying it.

Our findings showed that the Thematic approach neither advantaged nor disadvantaged students with respect to achievement on the kinds of tasks assessed by traditional standardized tests. On criterion tests referenced to the thematic content of instruction, did make gains on criterion tests referenced to the content of the instruction they received, and criterion referenced gains were shown to be related to implementation levels in the classrooms (Henderson, Landesman, & Allen, 1994). Among thematic students we also found that mathematics self-concept and interest in future work at occupations involving mathematics exceeded the comparison group on these variables. Modest though they are, these results amount to good news for teachers who wish to adopt instructional practices that are congruent with mathematics reform recommendations. A familiar refrain to anyone working with teachers who are attempting to align their teaching with NCTM standards is the concern that their efforts will be judged negatively if their students are assessed on measures that are not aligned with the new standards. These and other results (Henderson & Landesman, in press) suggest that students do learn the computation and application skills, as assessed by standardized tests, in the course of participating in activities that subordinate such skills to larger problem-solving situations. But it is also quite clear that a "reformed" curriculum alone does not constitute a sufficient condition to promote major improvements in achievement. Our observations over two years in Phase I thematic classrooms demonstrated quite forcefully that teachers tend to teach new thematic curricular materials via instructional practices that mimic their previous presentation of traditional text-based curriculum (with the possible exception of a few concessions to small group learning and hands-on activities). Thus, we realized that the major task ahead was to improve implementation of thematic curriculum, if we wished to give it a fair test.

The Partnership for Change phase of our work represents an extension of our Phase I work to test the effects of thematic instruction at a public middle school in another community. Although the community was larger than the one in which Phase I was conducted, the setting was similar in that agriculture was the dominant industry, and the school served a population of students of whom the majority were of Mexican descent. Our entry at this middle school was solicited by an established teacher, solidly committed to reform in mathematics instruction and to classroom practices that foster success in students of varying achievement levels and learning styles. Our original agreement was to collaborate with her department and to foster the implementation, development, and dissemination of successful thematic curriculum for heterogeneous classrooms by means of a four-way partnership involving the school, the university, a business partner, and a community college. We have attempted to document the change process, and to assess the mathematics achievement, self-perceptions, and mathematics attitudes of seventh graders assigned to three different instructional conditions at the Partnership school: a thematic approach, a traditional approach, and an approach that blended project activities with instruction organized around the textbook. We also compared the achievement of students in the thematic condition with that of two previous cohorts of seventh graders who had attended the Partnership and two sister middle schools serving similar populations.

Method

Subjects

The student population at the Partnership School is divided into families comprised of 150 students. This arrangement is intended to ease the transition to middle school by avoiding the impersonal character of traditional
departmentalized structures. With minor exceptions, the students in a given family spend their entire schedules with just four teachers. This configuration placed restrictions on the degree to which true random assignment of students to treatment groups could be implemented because, for scheduling purposes, the 70 students assigned to the school's Gifted and Talented Enrichment (GATE) program were all assigned to a single family unit. It was necessary to make a number of adjustments to accommodate students for whom the options for class assignment were restricted because of their participation in groups such as the GATE or band. For example, while there were no special mathematics classes specifically arranged for students who participated in the GATE program, it was necessary to assign students who were selected for this program to classes designated for GATE students in other subjects, thus reducing the "degrees of freedom" available for the assignment of these students to instructional conditions. Within the Homogeneous treatment group assignment to classes was based on achievement level, as had been the custom at this school prior to the Partnership. Most GATE students were assigned to the Homogeneous treatment group, with a small number scattered throughout classes in the Thematic condition. None were assigned to the Blend condition. This resulted in substantial differences in the entry level mathematics achievement of students in the different instructional conditions.

The Thematic instructional approach was under development during this first year of the program. The teachers who were assigned to teach the thematic approach were involved, along with their partners from the University, in selecting and adapting curricular materials deemed appropriate for thematic instruction for heterogeneous achievement levels. This instruction was characterized by a hands-on emphasis, use of small group learning activities, and the integrated use of a wide range of mathematics skills and concepts. Students in the Homogeneous treatment group received instruction that was organized mainly around textbooks, which varied according to teacher choice, based on achievement level. Students in the Blend condition received a combination of project-based and textbook-based instruction in classes comprised of heterogeneous achievement levels.

Measures

**Achievement.** The school district administered selected subscales from the Comprehensive Test of Basic Skills (CTBS) during the spring and fall of each academic year. Data from the Mathematics Concepts and Applications subtest were used to assess overall mathematics achievement. In our analyses we also planned to also use computation as well as Concepts/Applications and Computation subtest scores from spring testing. However, few students received the Computation subtest due to mis-communication from the District administrative offices to individual school sites. In addition to the Mathematics Concepts and Applications scores, Language Total scores were used as a covariate in some of our analyses. A locally designed test was also administered on a pre- and posttest basis to seventh graders at The Partnership Middle School. This test assessed specific skills and concepts that were included in a curriculum unit used in the Thematic treatment group, but that are also standard elements in the school's traditional grade seven mathematics curriculum (e.g., measurement, ratio and proportion, perimeter and area, circles, and percents).

**Self-Perception.** A Self-Perception instrument entitled "How I See Myself" was administered at the beginning and end of the year. This instrument included 5 subscales, three of which (self-perceptions of Academic Competence, Social Acceptance, and Global Self-Worth) were borrowed from Susan Harter's (1985) Adolescent Self-Perception Profile. The remaining two scales, Challenge-Seeking/Persistence and Goal Orientation (Henderson, 1991) assessed variables thought to be involved in the development of mathematics motivation. These measures will be used in on-going studies of motivational development, but for present purposes we hypothesized that the Thematic treatment would have a favorable effect on Goal Orientations and Challenge-Seeking/Persistence because of its emphasis on the integrated application of mathematics skills and concepts to ongoing.
multifaceted problems.

Mathematics Attitudes Questionnaire. The attitudes and beliefs that students hold toward mathematics may play an important role in their mathematics achievement, and in their interest in participating in mathematics (Haladyna, Shaughnessy & Shaughnessy, 1983; McLeod, 1985; Silver, 1985). Some educators have argued that disinterest and relatively poor mathematics achievement among students of Mexican descent is the result of early and repeated failure in school, and that these experiences lead to poor academic self-concept, negative attitudes toward school subjects, and alienation from school. Others have suggested that Hispanic children are directed away from mathematics as the result of negative cultural stereotypes of mathematicians (Cocking and Chipman, 1988). These beliefs have not been widely confirmed, and some research with other samples from the population involved in the present study has found quite positive attitudes toward mathematics and relatively high levels of mathematics self-concept (Henderson & Landesman, in press). Therefore, the present evaluation examined attitudes, with the long range aim of understanding the connection between mathematics attitudes and achievement, and how that relationship might be affected by variations in the approach to mathematics instruction.

Mathematics attitudes were assessed with a questionnaire utilizing a 5-point likert-like format, with responses ranging from strongly agree to strongly disagree. Several items were adapted from Fennema and Sherman's (1976) Mathematics Attitude Scale whereas others were developed specifically for this study. The instrument included six subscales, the items of which were randomly distributed throughout the questionnaire. (Two of these subscales, Mother, and Father Parental Support and Encouragement in Mathematics, are not presented here but will be employed in the longitudinal analysis as the program continues).

The Mathematics Self-Concept scale consisted of 5 items such as "I am good at working with numbers," and "I am good at mathematics." intended to assess students' confidence in their ability to learn mathematics. The Math Attitudes subscale (8 items) employed items such as "Mathematics is boring to me," "Mathematics is interesting," and "I really enjoy math class" to assess students' interest in and enjoyment of mathematics.

The Mathematics Future subscale (5 items) was devised to assess both interest in working at an occupation that uses math and expectations that their work as adults would require mathematics, as indicated by responses to items such as "When I am older, I expect to work in an area that requires mathematics," and "I would like to have a job working with math when I finish school." These items were adapted from items used in the National Assessment of Educational Progress (Dossey, Mullis, Lindquist, & Chambers, 1988).

Three items were adapted from the National Assessment of Educational Progress (Dossey, Mullis, Lindquist, & Chambers, 1988) to assess students' conceptions of the nature of mathematics. These items asked students to respond to statements such as "Learning mathematics is mostly memorizing," and "There is always a rule to follow in mathematics."

The Stereotypes of Mathematicians subscale (6 items) was included to examine the assertion that Mexican-Americans attribute characteristics to mathematicians which are contrary to their cultural values (Cocking & Chipman, 1988), characteristics such as being impersonal and distant from other people, and being unable to engage in normal family life. Items developed for this purpose included statements such as "Mathematicians don't show much interest in people," and "Mathematicians can have a normal family life."

We benefited from our Phase I experience in a number of ways. For instance, in originally proposing that Phase I middle school teachers undertake thematic instruction, we had made certain assumptions. We supposed that teachers who contracted to do so would write and develop thematic curriculum. We supposed that teachers who decided to use prepared thematic curriculum in their classrooms would spontaneously reflect and adjust
their teaching practices appropriately. Given their expressed interest in mathematics reform and the reform oriented philosophy of the schools' principal, we expected that these teachers would coordinate and consult with each other as a matter of course to adapt, extend, and revise thematic curriculum used, with mathematics learning objectives in mind. In Phase I of our work, we were clearly mistaken about all three suppositions. In Phase II we would successfully address two of these three areas.

Thematic teachers in both phases articulated two of the barriers to teaching and learning in a classroom newly committed to thematic instruction as the following: the need for curricular development time, and the need for methods appropriate to thematic instruction. A baseline condition of entry at the Phase II school was that the teachers would help write and develop thematic curriculum, and that we would help too. Here, when we supported curriculum development explicitly with released time and instrumentally with curriculum forging. The process of teachers developing and writing curriculum was slow in coming but did occur. Also at the Phase II school, we established early clear expectations for thematic teachers to reflect and adjust their teaching practices appropriately. With respect to our original third supposition, we repeated our mistake at the Phase II school, assuming that Phase II teachers would consult with their peers to hone thematic curriculum and their teaching of it, because they were clearly committed to the welfare of their students and seemed collegial and compatible among themselves. In retrospect we realize that teachers at the Phase II school never expressed the need to collaborate with peers.

**Results**

How students in the Thematic Treatment Group achieve in relation to students in the comparison conditions within their same school? This question was examined by means of a 2 (trials) x 3 (treatments) x 2 (sex) repeated measures analysis of variance, with pre- and posttest scores from the Mathematics Concepts and Applications subscale of the CTBS as the repeated measure. The analysis revealed significant main effects for Treatment ($F = 12.362, df = 2/362, p < .0005$) and Trials ($F = 54.955, df = 1/362, p < .0005$), as well as a significant Trials by Sex interaction ($F = 5.897, df = 1/362, p < .02$). The trials by sex interaction was accounted for by steeper gains for girls than for boys. The lack of a significant Treatment by Trials interaction suggests that although pretest scores of the Homogeneous group were significantly higher than the Theme, which was higher than the Blend, the three groups made equivalent gains during the year. Figure 1 shows that initial achievement differences between the groups were maintained. The trials by sex interaction is displayed in Figure 2.

**Insert Figures 1 and 2**

We then used an analysis of covariance (ANCOVA), with Language Total and Mathematics Concepts and Applications pretest scores as covariates, to control for initial differences among the groups. With this statistical adjustment for initial differences in language and mathematics achievement at the time of 7th grade entry, the differences among treatment groups did not attain significance at alpha level .05 ($F = 2.58, df = 2/408, p < .08$), providing further evidence that the three treatment conditions did not have a differential effect on mathematics achievement, as assessed by the Concepts and Applications subtest of the CTBS.

Since different instructional approaches may be differentially effective for students of different achievement levels, it is important to consider whether the instructional treatments were equally effective for students differing in their initial achievement levels. We addressed this issue by controlling for initial differences in the intact treatment groups by means of an ANCOVA, using pretest scores from the CTBS Language Total subscale as the covariate. Given differences in the central tendency and distributions of scores in the different treatment groups, an "Achievement Level" variable was created by splitting each treatment group at its own median on scores from the CTBS Mathematics Concepts and Applications subtest. The intent was to determine the effectiveness of each treatment for higher and
lower achieving students within its own mix of achievement levels. The main effect for Treatment was significant \( (F = 6.627, df \ 2/407, p < .001) \). As we would expect, by definition, the main effect for Achievement Level was also significant \( (F = 59.918, df \ 1/407, p < .0005) \). Post hoc comparisons using Fisher's LSD test showed the Treatment effect to be the result of significantly higher scores for the Homogeneous group than for the Blend group. The homogeneous and Thematic treatments did not differ. The effect of major interest in this analysis, the interaction between Achievement Level and Treatment, was also significant \( (F = 5.287, df \ 2/407, p < .005) \). The comparisons of principal interest in this analysis are those that explain the Achievement Level by Treatment interaction. The results suggest that, with statistical adjustments for entry level language and mathematics achievement, the three treatments did not differ in effectiveness for low achieving students. But for the higher achieving students, the Homogeneous treatment was associated with higher achievement than either of the other two treatments. In absolute value, the adjusted posttest scores of Theme students were higher than those of their peers in the Blend treatment, but the difference fell just short of attaining significance (Fisher's LSD test, \( p < .06 \)). Thus, it appears that the treatments do not differ in their effects for lower achieving students, but students assigned to the traditional homogeneous arrangement appears to have an advantage for higher achieving students, at least in terms of performance on traditional achievement tests.

We conducted an additional analysis to determine whether or not the treatments had differential effects on Concepts and Applications achievement for students of different ethnic backgrounds. The small number of students who were not of Euro-American or Latino background were not included in this analysis. Using pretest scores as a covariate to control for initial achievement, the only significant result was the main effect for treatment \( (F = 4.137, df \ 2/361, p < .02) \). Post hoc comparisons employing Fisher’s LSD test showed the main treatment effect to be accounted for by the fact that the adjusted mean for the Homogeneous treatment exceeded that of the Blend group. No other comparisons were significant. The lack of a significant treatment by ethnicity interaction indicated that the treatments were not differentially effective for Latino and Euro-American students.

To compare the achievement of students in the thematic treatment with previous cohorts of students we examined the pretest to posttest achievement of all students from three cohorts of students for whom data were available from school district files. Cohort 1 constituted the seventh grade classes at The Partnership School and two sister middle schools serving similar populations during the 1991 - 92 academic years. Cohorts 2 and 3 attended seventh grade at these schools during 1992 - 93 and 1993 - 94, respectively. A 2 (Trials) by 3 (Cohorts) ANOVA with CTBS Mathematics Concepts and Applications pre- and posttest scores as the repeated measure yielded significant main effects for both Cohort \( (F = 13.145, df \ 2/1266, p < .0005) \) and Trials \( (F = 45.70, df \ 1/1266, p < .0005) \) as well as a significant Cohort by Trials interaction \( (F = 3.715, df \ 2/1266, p < .02) \). A steeper pre- to posttest slope for Cohort 1, compared to other Cohorts, accounted for the interaction.

A comparison of Cohorts 1 and 2 against the PMS Thematic treatment subset of Cohort 3 also produced significant main effects for Cohort \( (F = 3.059, df \ 2/719, p < .05) \) and Trials \( (F = 67.16, df \ 1/719, p < .0005) \) and a significant Cohort by Trials interaction \( (F = 7.261, df \ 2/719, p < .001) \), with the pre-to-posttest slope for the thematic subset of Cohort 3 showing a steeper positive pretest to posttest slope than for the other cohorts.

A major instructional activity for the Theme group during the year was a thematic unit entitled Quincy Market (Regional Math Network, 1987). A criterion referenced test of 30 items was created to assess student learning from this unit. The tasks sampled by the test were all specifically within the scope of the Quincy Market theme, but the entire item set also represented material typically covered in the seventh grade mathematics curriculum at
the target school. The test was administered on a pre- and posttest basis to students in all treatment groups at PMS. The results of a 2 (trials) by 3 (treatment groups) ANOVA showed significant main effects for treatment ($F = 11.785$, $df = 2/393$, $p < .0005$) and for trials ($F = 68.879$, $df = 1/393$, $p < .0005$), and a significant trials by treatment interaction ($F = 15.015$, $df = 2/393$, $p < .0005$). Given the initial differences among the intact groups, the effects for treatment groups were to be expected. Whereas all groups made gains in the absolute value of their scores on the Quincy Market test, the significant interaction is explained by a significantly steeper positive slope for the Thematic group than for either of the other treatment groups. The pattern of outcomes is further clarified by a subsequent analysis in which CTBS Concepts and Applications, and CTBS Language Total were used as covariates in an ANCOVA to adjust for initial achievement differences among the treatment groups. The treatment effect was significant ($F = 14.451$, $df = 2/370$, $p < .0005$), with post hoc tests (Fisher's LSD) showing that the Thematic Group achieved higher scores than the two comparison groups. No other comparisons were significant.

The previous analysis provides information on the relative performance of the treatment groups on the assessment items incorporated in the Quincy Market test. However, this analysis does not reveal whether the seventh grade programs experienced by the different treatment groups may have been differentially effective with students who differed in their incoming mathematics achievement. In order to examine this issue, we conducted a 2 (trials) by 3 (groups) by 2 (levels) ANOVA. As before, Achievement Levels were defined by splitting each group at the median of its own CTBS Concepts and Applications pretest. The main effects for treatment ($F = 15.662$, $df = 2/372$, $p < .0005$), levels ($F = 194.491$, $df = 1/374$, $p < .0005$), and trials ($F = 56.862$, $df = 1/374$, $p < .0005$) were all significant. The between subjects Treatment by Levels interaction was significant ($F = 10.459$, $df = 2/306$, $p < .0005$), as was the trials by treatment ($F = 14.102$, $df = 2/374$, $p < .0005$) interaction. The within subjects Trials by Treatment by Levels 3-way interaction was not significant. The results suggest that pre-to posttest change was associated with treatment, but that change for high and low achieving students was not a function of which treatment they received.

The pattern of results may also be clarified by controlling for differences among the intact groups by using a pretest measure that is correlated with the criterion variable as a covariate. We used the CTBS pretest Language Total subscale scores as the covariate in an ANCOVA, with Quincy Market posttest achievement as the dependent variable. Both the main effect for treatment ($F = 11.668$, $df = 2/372$, $p < .0005$) and the Treatment by Achievement Level interaction ($F = 6.694$, $df = 2/272$, $p < .001$) were significant. The interaction was the outcome of interest in this analysis. Among the higher achieving subsamples, post hoc analysis with Fisher's LSD test identified the following differences in performance on the skills and concepts sampled on the Quincy Market test: Homogeneous > Theme; Homogeneous > Blend, and Theme > Blend. Among the lower achieving subsamples, scores of the Theme group were significantly higher than those of both the Homogeneous and Blend groups. The Homogeneous and Blend groups did not differ. Thus, the interaction appears to be accounted for by the fact that the instructional experience provided in the Homogeneous condition was of greater benefit to high achieving students than was that of either of the other two conditions, whereas the Thematic treatment was of greater benefit to lower achieving students than were the other treatments.

An analysis of achievement by ethnicity, with the Quincy Market pretest performance as the covariate, produced significant main effects for Ethnicity (with Euro-American students scoring higher than Latino students) and Treatment. The Treatment by Ethnicity interaction was not significant. Post hoc comparisons showed that the Treatment effect was explained by an adjusted mean for the thematic group that exceeded the adjusted means of both of the other two groups. The adjusted means of the Blend and Homogeneous treatments did not differ.
Mathematics Attitudes Subscales. The several mathematics attitudes variables were examined by means of a series of 3 (treatments) by 2 (sex) by 2 (achievement level) ANCOVAs, with the pretest score for the relevant subscale as the covariate in each case and the posttest score as the dependent variable. The analysis of the Mathematics Self-Concept variable revealed a significant Achievement Level effect \( (F = 10.595, \text{df} 1/325, p < .001) \), indicating stronger Mathematics Self-Concept for higher achieving students than for their lower achieving peers. There were no other significant effects. The majority of students agreed or strongly agreed with the statements "I usually understand what we are talking about in mathematics." (74 percent), and "I am good at working with numbers" (62 percent). However, slightly fewer than half agreed or agreed strongly with the statement, "I am good at mathematics." Fewer still agreed or strongly agreed that "Math is easy for me" (36 percent) or that "I am good at word problems" (37 percent).

The 3 (Treatment) by 2 (Sex) by 2 (Achievement Level) ANCOVA for the Stereotypes of Mathematicians subscale yielded a significant main effect for Treatment \( (F = 4.143, \text{df} 2/330, p < .02) \). Fisher's LSD test showed that the Homogeneous group held less unfavorable stereotypes of mathematicians than did either the Thematic or Blend groups. The Theme and Blend groups did not differ from one another. There were no other significant effects, but the effect for Sex approached significance \( (F = 3.44, \text{df} 1/330, p < .065) \), with a trend toward less unfavorable stereotypes among girls than among boys. There was an approximately even split between those who agreed (agree + strongly agree) and those who disagreed (disagree + strongly disagree) with the statements "Mathematicians are too interested in their work to enjoy life," and "Mathematicians seem to be in a world of their own." On the other hand, only a small minority expressed unfavorable stereotypes in response to the items "Mathematicians are less friendly than other people," and "Mathematicians can have a normal family life." Students appear to see mathematicians as people who are very absorbed in their work, but who nevertheless have rather normal relations with other people, including living a normal family life.

The ANCOVA on the Mathematics Future variable revealed a significant main effect for the Achievement Level by Treatment interaction \( (F = 3.001, \text{df} 2/327, p < .05) \). There were no other significant effects. Students who entered seventh grade with different levels of achievement in mathematics were differentially affected by the Thematic and Homogeneous treatments. Post hoc analyses using Fisher's LSD identified the specific differences accounting for the interaction as those between the high and low achieving homogeneous groups \((p < .03)\) and between the high homogeneous and the high theme group \((p < .02)\) accounted for the interaction.

With data from the Mathematics Future subscale pooled across the treatment, minority of students expressed interest in working at a job involving mathematics or expected to be working at a job requiring mathematics after finishing school. For example, 36 percent agreed or strongly agreed that "I would like to work at a job using mathematics," 39 percent agreed or strongly agreed that "When I am older, I expect to work in an area that requires mathematics," and 38 percent agreed or strongly agreed with the statement "I would like to have a job working with math when I finish school." Just under half (48 percent) agreed or strongly agreed that "A job working with mathematics would be an interesting way to earn a living." The pattern of responses to these four items is especially interesting in relation to the item that stated, "You need a good math background to get ahead in most jobs these days." Students overwhelmingly agreed that preparation in mathematics is important to "getting ahead," yet a fair number of those same students showed little interest in working at jobs involving mathematics. These
results are depicted in Figure 8.

The ANCOVAs conducted on the Attitudes Toward Mathematics and the Nature of Mathematics variables yielded no significant results. A minority of students thought mathematics was boring. An impressively high proportion of students agreed or strongly agreed with the statements "I am willing to work hard to do well in mathematics," and "I feel good when I solve a mathematics problem by myself." Over 40 percent of the students agreed with the statements "I would like to take more mathematics," and Mathematics is one of the most interesting school subjects," whereas over 60 percent agreed or agreed strongly that "Mathematics is interesting." The pattern of responses is shown in Figure 9.

There were no significant treatment effects on beliefs about the Nature of Mathematics (see Figure 10). A small minority of students expressed disagreement with statements presenting the nature of mathematics as a subject primarily consisting of the rote application of memorized rules. For data pooled across the three treatment groups, only 17 percent disagreed or strongly disagreed with the statement that "Mathematics is mostly memorizing." A mere 4 percent disagreed or strongly disagreed with the statement "There is always a rule to follow in mathematics," and only 22 percent disagreed or strongly disagreed with the statement that "Mathematics is made up of unrelated topics."

Self-Perceptions. Data for the three treatment groups at PMS were pooled for initial analyses of pre- to posttest changes in students' self-perceptions. Repeated measures ANOVAs were used to compare pre and posttest scores for the self-perception variables Challenge-Seeking/Persistence, Scholastic Competence, Goal Orientation, Social Acceptance, and Global Self-Worth. Significant declines were found for the variables Challenge-Seeking/Persistence ($F = 23.595, df/1/392, p < .0005$), Scholastic Competence ($F = 10.072, df/1/392, p < .002$), and Global Self-Worth ($F = 9.226, df/1/392, p < .003$) Reports of Goal Orientations did not change over the course of the year, while there was a marginally significant increase in self-perceptions of Social Acceptance ($F = 3.683, df/1/392, p < .056$). Follow-up ANCOVAs tested for treatment and gender effects, controlling for initial differences by using pretest scores for the relevant subscale in each analysis. The only significant effect identified in this analysis was a significant main effect for sex on the Social Acceptance variable, on which girls had higher scores than boys ($F = 4.139, df/1/333, p < .05$).

Predictors of Mathematics Achievement. Using pooled data, stepwise multiple regressions were carried out to explore the efficacy of self-perception and attitudinal variables as predictors of mathematics achievement. The five self-perception variables were entered into the first analysis, with posttest scores on the CTBS Concepts and Applications subtest as the criterion variable Challenge-Seeking/Persistence (with a standardized coefficient of .135) and Scholastic Competence (with a standardized coefficient of .383) met the criterion for entry (15), together accounting for approximately 22 percent of the variance in mathematics achievement ($R = .47, F = 58.292, df/2/414, p < .0005$).

The second stepwise regression included the attitudinal variables of Mathematics Self-Concept, Attitude Toward Mathematics, Mathematics Future, and Nature of Mathematics as predictors, with CTBS Concepts and Applications as the criterion variable. Stereotypes of Mathematicians (standardized coefficient = .104) and Attitude Toward Mathematics (standardized coefficient = -.139) met the criterion for entry, accounting for a small ($R^2 = .017$) but significant portion of the variance in mathematics achievement ($R = .13, F = 3.584, df/2/414, p < .03$).

Conclusions and Postscript

This report has focused on student achievement in mathematics and on attitudes
and self-perceptions that are thought to be instrumental in students' motivation to achieve and to participate in the study of mathematics. The evaluation of achievement has relied heavily on conventional standardized measures, specifically, the Concepts and Applications subscale of the CTBS. It may seem paradoxical that we have relied heavily on this test, considering that we regard it to be poorly suited as a gauge of student learning from a thematic approach that emphasized problem solving, higher order thinking skills, and communication. It is reasonable, therefore, to question the prominent role played by this traditional standardized measure in the first year's evaluation. The rationale for giving so much attention to traditional standardized measures must be understood within the context in which this project was situated.

One of the events that led to the planning for the present project was a decision by the central administration of the school district to adopt heterogeneous grouping as the organizational mode for all instruction. For several reasons, this decision raised concerns among teachers of mathematics in the target school. For one thing, teachers lacked access to curricular materials appropriate to support instruction in heterogeneous classes. Secondly, even if appropriate materials could be obtained, teachers had received no training to provide instruction in heterogeneous situations. And, finally, teachers whose only experience was with teaching in ability grouped classes were genuinely worried that students, particularly the higher and lower achieving ones, would be short changed in this arrangement. Mathematics teachers at The Partnership Middle School were supportive of the design of the project because they thought it had the potential to provide empirical information on the results of heterogeneous grouping. The reasoning went something like this. If students who were taught in heterogeneous groups did not suffer with regard to the kind of evidence that is meaningful to school boards and parents (i.e., standardized test results), then teachers could adopt the heterogeneous approach with some confidence that the change in grouping practices would not be harmful to the students in their charge. Moreover, the design would provide an opportunity to examine, against the same criterion, the learning of students in a Thematic approach (that incorporates the major reform initiatives being promoted by the National Council for Teachers of Mathematics) compared to the learning of students in an approach to instruction that reflects a more cautious transition toward reform in which the traditional textbook retains much of its traditional importance (i.e., the Blend treatment).

In this project, the importance of demonstrating that an innovation does not harm students cannot be over emphasized. Teachers who resist being forced to change their instructional practices are often persuaded to change when they witness demonstrable successes experienced by colleagues who do adopt reform recommendations. If students who experience an "experimental" approach profit on conventional measures as well or better than they could be expected to under a more traditional treatment, then the next task is to devise ways to assess the additional, unique benefits to be derived from the new approach, while continuing to monitor achievement on traditional tasks.

On these terms, the news from this first operational year of the project is quite good. It is important to consider the results in light of the fact that the Thematic approach was under development during this year. Implementation involved a considerable degree of trial and error, with the thematic teachers piloting much of the thematic curriculum for the first time. The outcomes reflect a school in which the mathematics teachers in all instructional conditions were dedicated and working hard to provide effective instruction to their students. Students in all of three treatment groups made significant gains. This is an impressive accomplishment, since pre- and posttest achievement as reflected by NCE scores is referenced to norm group performance at different points in the school year. Students who learned just as much as students in the norming sample, no more or no less, would attain the same NCE scores at pre- and posttesting. The fact that students in all treatment groups made significant gains reflects upward movement against the normative reference group. Although learning, overall.
did not differ across achievement groups, girls in all groups made larger gains than boys.

Analyses in which the Language Total pretest scores were used as a covariate did suggest that overall students in the Homogeneous group achieved at higher rates than those in the Blend condition, but their achievement did not differ from those in the Thematic treatment. We also learned from these analyses that the three treatments did not differ in effectiveness for the lower achieving students (those with scores below their treatment group median on the Concepts and Applications pretest), but higher achieving students appeared to profit most from the Homogeneous treatment condition. In these comparisons among higher achieving students, the Theme group was marginally higher than the Blend group. It should be noted, however, that the higher performance of higher achieving students in the traditional Homogeneous treatment may be attributable to factors other than classroom instruction. A strong complaint of some teachers at the PMS was that assertive parents with good knowledge of the workings of the system were those who pushed for assignment of their children to homogeneous classes. They were also through to be the parents who pressed their children to achieve well in school, and who provided close supervision of homework.

Comparisons with seventh grade cohorts in two schools that serve reasonably comparable populations also showed that the Partnership Middle School Thematic treatment was doing well. The slope of change for the PMS Thematic group was comparable (but with higher scores at both pre and posttesting) to that of one comparison school, and it was far superior to that of the second school which began with comparable pretest scores and actually showed a significant decline from pre- to posttesting.

The final examination of achievement on the standardized achievement measure compared the PMS Thematic group with previous cohorts of students who attended the three schools that were involved in the previous analysis. The pre- to posttest slope for PMS Theme students was parallel to that of the 1991-92 cohort. The pretest scores of both the 1992-93 and 1993-94 cohorts were lower than those of the 1991-92 cohort, confirming the belief held by teachers that incoming students were less well prepared than in former years. The positive pre- to posttest slope reflecting gains by the PMS Theme group was significantly steeper than that of the 1992-93 cohort, again providing cause to conclude that the Thematic treatment did no harm to student achievement, as assessed by traditional measures. In fact, Theme students made noteworthy gains in comparison to their predecessors from the previous year.

In addition to the evidence of learning provided by the standardized test, the criterion referenced test covering the skills and concepts taught in the Quincy Market unit showed favorable results. Although Theme teachers expressed disappointment that test results fell well short of a mastery level, the results did provide strong evidence of positive results for the Thematic approach. The teachers involved in this first year trial opted to continue with the thematic approach, and to work on enriching the thematic curriculum, while other teachers enlisted to pick up the students as they moved into eighth grade and to pilot new thematic curriculum with these students, thus providing continuity in the approach to mathematics instruction experienced by these students.

At the level of implementation reached during the year under review, the Theme approach appeared to exert little influence on mathematics attitudes or students' motivation-linked self-perceptions. An interesting exception was seen in the results for the Mathematics Future attitude subscale, which revealed a treatment by achievement level interaction. Whereas neither the Blend nor the Thematic treatment had differential effects on the interest of students of differing achievement levels in working in the future at occupations involving mathematics, the traditional treatment experienced by students in the Homogeneous classes did affect students differently, depending on their entry mathematics achievement level. Higher achieving students were more favorably influenced than were lower achieving students. The picture presented by overall responses (data pooled across achievement levels) to items comprising the Mathematics
Future scale, the majority of students were not highly enthusiastic about working at jobs requiring mathematics. Yet they strongly endorsed the idea that preparation in mathematics is important in "getting ahead."

Students at PMS did not appear to hold strongly negative stereotypes of mathematicians. While they viewed mathematicians as people who are very absorbed in their work, they did not seem to equate this tendency with mathematicians being loners with little interest in other people or with an inability to live a normal family life.

In general, students at the target school seemed to have reasonably strong mathematics self-concepts. A very small minority felt they did not usually understand what was being discussed in math class, and a large majority felt they were good with numbers and good at math. Those who thought math is easy and those who disagreed with this statement were about equally divided, as were those who agreed or disagreed that they were good at word problems. In one sense it is not clear what the ideal outcome for this set of items would be. For example, we would not want students to feel that mathematics is beyond their capability. On the other hand, we would hope that the instruction they receive is sufficiently challenging for them to recognize that mathematics can be difficult. And we hope for them to develop the motivation to persist in the face of difficulty and to enjoy the challenge.

In general, students expressed quite favorable Attitudes Toward Mathematics. It was especially interesting that a very large majority of students said they were willing to work hard to do well in mathematics, and that they felt good when they were able to solve a mathematics problem for themselves. Slightly more than 60 percent said mathematics is interesting, although about 10 percent fewer than that counted mathematics among their most interesting subjects.

Perhaps the most disappointing aspect of the attitudinal results was reflected in students' persistent views that mathematics mostly involves rote memorization, that there is always a rule to follow, and that mathematics consists largely of a set of unrelated topics. It is possible that students under the Thematic condition for the very first time did not perceive that their hands-on activities and opportunities for understanding and applying skills and concepts embedded in the thematic curriculum also qualified as mathematics. This possibility should be follow-up in interviews or focus groups with students.

Although the instructional treatments did not appear to influence changes in students' motivation-related self-perceptions, as predictors, Challenge-Seeking/Persistence and self-perception of Scholastic Competence did account for a significant portion of the variance in mathematics achievement. It seems likely that these characteristics are developed over a long period of time through the kinds of mathematics instruction students receive throughout their elementary school experience. It may be overly optimistic to believe that these predictors of achievement may be changed quickly or easily.

A primary lesson from our work in partnership with schools is that real change in instructional approaches comes very slowly, even among teachers who are philosophically committed to reform. The curriculum must be rich and intellectually engaging for teachers as well as students, and we believe there must be opportunities for teachers to adapt and elaborate existing thematic curricula in ways that fit with their own interests and priorities. The Quincy Market curriculum that was used during the initial relevant to the data presented here was found wanting in several ways. So, while the initial year of its implementation was uneven, the teacher who was most involved with it, and whose academic background was in economics, has embellished it and implemented it with increased confidence during the current year. As she has done so, we have explored the use of more performance-based assessment and have underway an investigation designed to clarify the influence of thematic instruction on students' communication of their problem-solving processes. We are also attempting to examine the separate and joint influences of thematic instruction and weekly testing with performance-based ap-
approaches. And we are beginning a study, informed by sociocultural theory, designed to provide a close examination of the construction of motivation in the context of small group activity.

The first year of Phase I involved preliminary developmental activities with the business partner. Our business partner developed the initial draft of a thematic unit entitled “Mathematics in Manufacturing,” which incorporates opportunities for students to have first-hand experiences in a high tech manufacturing plant. Over the summer mathematicians from the university worked with teachers to develop problems suitable for the “Mathematics in Manufacturing” theme. The developmental process has progressed through several iterations involving review and revision by teachers, the business partner, curriculum design people from our project, and mathematicians. Implementation of the unit has just begun in eighth grade, and teachers and project personnel are working on the articulation of the seventh grade Quincy Market unit to “Mathematics in Manufacturing.” We have also been able to assist on-site educators in the implementation of their existing priorities (such as “Family Math”), and the partnership has been the catalyst for developing a televised call-in program to assist students with homework. The program has enjoyed instant popularity throughout the community.

The most difficult challenges ahead appear to be political and structural. The political issues revolve around the inconsistency between an official district policy that mandates heterogeneous grouping, and the reality of assertive parents being successful in getting their children assigned to GATE classes, which skim the highest achieving students (but not limited to traditional definitions of gifted) off the pool, leaving a restricted range of students for assignment to the heterogeneous classes. The district appears to have trouble bringing reality into conformity with stated philosophy out of fear of losing the higher achieving students to private schools.

The structural problems involve the difficulty of finding time for teachers to work collaborative in the development and sharing of instructional resources, and in mutual mentoring.

The instructional and administrative personnel in this school seem more ready to work toward those ends that those of other schools in which we have worked, but it will be some time before we expect to be in a position to study the full potential of thematic instruction to influence learning and motivation among Latino students. And, from our experience to date, we find it hard to imagine that effective reform can come as the result of administrative mandate, or without major investments in professional development.
References


Fennema, E., & Sherman, J. (1976). Fennema-Sherman mathematics attitude scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. In JSAS Catalog of Selected Documents in Psychology [Ms. No. 1225].


Contextualizing Middle School Mathematics


Figures

Figure 1
Treatment Comparisons: Pre and Posttest Scores on Mathematics Concepts and Applications

Figure 2
Sex by Trials Interaction for Mathematics Concepts and Applications

Figure 3
Mathematics Concepts and Applications for Three Grade Seven Cohorts - Pooled Across Three Schools

Figure 4
Concepts and Applications Achievement by Treatment Groups: Quincy Market
Figure 5
Percent of Students Who Agree or Disagree With Mathematics Self-Concept Statements

Figure 6
Percent of Students Who Disagree or Agree With Mathematician Stereotype Statements

Figure 7
Mathematics Achievement Level by Treatment Interaction for Mathematics Future Subscale

Figure 8
Percent of Students Who Agree or Disagree with Mathematics Future Statements.
Figure 9
Percent of Students Who Agree or Disagree with Attitudes Toward Mathematics Statements

Figure 10
Percent of Students Who Disagree or Agree with Nature of Mathematics Statements