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

---

Monograph 7

2005

Differences in Student Performance for Three Treatment Groups

Thomas A. Romberg

Lorene Folgert

University of Wisconsin-Madison

Mary C. Shafer

Northern Illinois University

WCER

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., & Folgert, L. (Editors). (2005). *Differences in Student Performance for Three Treatment Groups* (Longitudinal/Cross-Sectional Study of the Impact of Teaching Mathematics using *Mathematics in Context* on Student Achievement: Monograph 7). Madison, WI: University of Wisconsin-Madison.

Romberg, T. A., Shafer, M. C., Folgert, L., & LeMire, S. (2005). Overall differences in achievement for the three treatment groups. In T. A. Romberg, M. C. Shafer, & L. Folgert (Editors) *Differences in Student Performance for Three Treatment Groups* (Longitudinal/Cross-Sectional Study of the Impact of Teaching Mathematics using *Mathematics in Context* on Student Achievement: Monograph 7) 8-38. Madison, WI: University of Wisconsin-Madison.

Shafer, M. C., Romberg, T. A., & Folgert, L. (2005). Classroom achievement of comparable classes. In T. A. Romberg, M. C. Shafer, & L. Folgert (Editors) *Differences in Student Performance for Three Treatment Groups* (Longitudinal/Cross-Sectional Study of the Impact of Teaching Mathematics using *Mathematics in Context* on Student Achievement: Monograph 7) 39-111. Madison, WI: University of Wisconsin-Madison.

Romberg, T. A., Shafer, M. C., Folgert, L., & LeMire, S. (2005). Other results. In T. A. Romberg, M. C. Shafer, & L. Folgert (Editors) *Differences in Student Performance for Three Treatment Groups* (Longitudinal/Cross-Sectional Study of the Impact of Teaching Mathematics using *Mathematics in Context* on Student Achievement: Monograph 7) 112-143. Madison, WI: University of Wisconsin-Madison.

Table of Contents

The L/CSS Monograph Series	**4**
Introduction to Monograph 7	**6**
Chapter 1: Overall Differences in Achievement for the Three Treatment Groups	**8**
Four Key Variables	**9**
Analysis of Covariance	**14**
Additional Analyses	**15**
Conclusions and Implications	**37**
Chapter 2: Classroom Achievement of Comparable Classes	**39**
Comparable Classes	**39**
Comparable Groups in Year 1	**39**
Comparable Groups in Year 2	**71**
Comparable Groups in Year 3	**95**
Conclusion	**110**
Chapter 3: Other Results	**112**
Differences in Standardized Test Percentiles	**112**
Differences in Performance on the Anchor Items	**115**
Differences in Change in Attitudes	**130**
Differences in Change in Mathematical Reasoning	**139**
Conclusion	**143**
References	**144**

The L/CSS Monograph Series

This is the seventh 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, 1997-98) 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 1997-98 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 methodologies used in and the results of scaling the instruction students experienced, their opportunity to learn comprehensive mathematics with understanding, and the capacity of their schools to support teaching and learning mathematics.

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 seventh monograph provides the results of our attempts to understand the variables that had an impact on student performance.

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

Monograph 1 — Purpose, Plan, Goals and Conduct of the Study

- Chapter 1. Standards-Based Reform and *Mathematics in Context*
- Chapter 2. Design of the Longitudinal/Cross-Sectional Study
- Chapter 3. Instrumentation, Sampling, and Operational Plan
- Chapter 4. Conduct of the Study

Monograph 2 — Backgrounds on Students and Teachers

- Chapter 1. Background Information on Students
- Chapter 2. Information on Teacher Background Variables

Monograph 3 — Instruction, Opportunity to Learn with Understanding, and School Capacity

- Chapter 1. The Quality of Instruction
- Chapter 2. Opportunity to Learn with Understanding
- Chapter 3. School Capacity

Monograph 4 — Measures of Student Performance

- Chapter 1. Classroom Achievement
- Chapter 2, The Development of a Single Scale for Mapping Progress in Mathematical Competence

Monograph 5 — The Impact of *Mathematics in Context* on Student Achievement

- Chapter 1. Grade-Level-by-Year Studies
- Chapter 2. Cross-Sectional Studies.
- Chapter 3. Longitudinal Studies

Monograph 6 — Differences in Performance between *Mathematics in Context* and Conventional Students

- Chapter 1: Differences in Experimental Treatments and Units
- Chapter 2. Contrast Between MiC, MiC (Conventional), and Conventional Student Performance in the Cross-Grade and Cross-Year Studies
- Chapter 3. Contrast Between MiC and Conventional Student Performance in the Longitudinal Studies

Monograph 7 — Differences in Student Performance for three Treatment Groups

- Chapter 1: Overall Differences in Achievement for the Three Treatment Groups
- Chapter 2: Classroom Achievement of Comparable Classes
- Chapter 3: Other Results

Monograph 8 Implications and Conclusions

Chapter 1. Implementation Stories

Chapter 2. Insights about Implementing a Standards-Based Curriculum in Schools

Chapter 3. What we have Learned.

Introduction to Monograph 7

This monograph contains three chapters. Their purpose is to summarize the variables that had an impact on changes in student performance for the three treatment groups identified in Monograph 6, Chapter 1. Data are reported in relationship to Research Question 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 during field-testing 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.

Information on five composite variables was gathered in the study. This simplified model presented in Chapter 2 of Monograph 1 describes the relationship between variations in classroom achievement (CA), aggregated by strand, or total performance, that can be attributed to variations in prior achievement (PA), method of instruction (I), opportunity to learn with understanding (OTL_u), and school context (SC). This relationship can be expressed as—

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

These composite indices were intended to capture the variability across classes (and schools). The details of how these composite indices were created are discussed in Monographs 3 and 4.

As explained in Monograph 6, Chapter 1, methodological issues that arise in conducting quasi-experimental studies involve attempts to document differences in both the treatments and the subjects. This can only be done by attempting 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 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 Research Question 3, data were collected for the three treatment groups MiC, MiC (Conventional), and Conventional students in two of the four districts involved in the study: Districts 1 and 2. In Chapter 1, the relationships of the key variables to the overall differences in performance on CA are examined. Then, for the three treatment groups contrasts in differences in performance are made on CA and its sub-scales. The contrasts come from an analysis of covariance on overall CA scores using PA as the covariate. In Chapter 2 of this monograph, the performance of groups of students who were comparable in PA is investigated with respect to instruction (I), OTL_u, and SC. In Chapter 3, contrasts of changes in standardized test percentiles, in the anchor items from the CA tests, and in attitudes towards mathematics, are reported for the three treatment groups. Also, changes in reasoning from the Collis-Romberg Profiles are reported for the cohorts of students studied longitudinally.

CHAPTER 1: OVERALL DIFFERENCES IN ACHIEVEMENT FOR THE THREE TREATMENT GROUPS

Thomas A. Romberg, Mary C. Shafer, Lorene Folgert, and Steven LeMire

As described in Monograph 6, Chapter 1, we concluded that there are three distinct treatment groups. The MiC group reflects the intended pedagogy and content of the program, and the Conventional group reflects conventional pedagogy and content. The third group taught MiC with more conventional pedagogy or nominally taught MiC and used a combination of conventional texts and MiC materials. This group is labeled MiC (Conventional). In Districts 1 and 2 there were 79 teachers instructing groups of students in a particular grade over the three years of the study:

- 28 of the groups were identified as MiC,
- 28 groups were identified as MiC (Conventional), and
- 23 groups were identified as Conventional.

In Chapters 2, 3, and 4 of Monograph 6 the differences in student achievement for these groups were examined for the eight grade-level-by-year studies, the cross-grade studies, the cross-year studies, and the longitudinal cohorts. In these studies there was considerable variation in the sets of scores for all three treatment groups. The means on classroom achievement (CA) often favored the MiC groups or were comparable to the Conventional groups, and were often greater than those for the MiC (Conventional) group. Where there are differences in means for different groups, they can best be attributed to differences in the groups of students taught by specific teachers, and not simply to the use of MiC, MiC (Conventional) or conventional instructional treatments.

In this chapter the overall differences in mean achievement are reported for these three treatment groups. Note that the groups come from all years and all grades. The unit of analysis is the teacher/student group. The measures of achievement are the classroom achievement (CA) scale and its sub-scales. Using CA for groups at different grades in different years is possible since we were able to calibrate a single proficiency scale that incorporated items from all the tests. Two assessment systems were developed for this study: the External Assessment System (Romberg & Webb, 1997–1998) and the Problem Solving Assessment System (Dekker et al., 1997–1998). Student responses on these assessments were used create the CA scale that provides a frame of reference for monitoring growth. (See Monograph 4, Chapter 2 for description of the development of the proficiency scale and progress maps for charting growth over time.) In this monograph, measures of classroom achievement are stated in terms of this proficiency scale. (Complete data sets for students can be found in Romberg, Webb, Folgert, & Shafer 2003a, b, c, d, e, and f.) Finally, because a priori the groups differed on several measures, the statistical model used for this analysis was analysis of covariance (ANCOVA).

Four Key Variables

Given the fact that this was a quasi-experimental study, neither teachers nor students were randomly assigned to treatment groups, and the detailed data reported in Monograph 6 made it clear that the treatment groups differed on four key variables: student prior achievement (PA), method of instruction (I), opportunity to learn with understanding (OTL_u), and school capacity for reform (SC). The research team used standardized test scores as measures of prior achievement. As part of the regular standardized testing programs in each district, the *TerraNova* (CTB/McGraw-Hill, 1997) and the *Stanford Mathematics Achievement Test* (Harcourt Brace Educational Measurement, 1997) were administered in Districts 1 and 2, respectively. National percentile rankings were used for purposes of comparison across groups (See Romberg, Webb, & Folgert, 2004 for complete data sets.) The instruction composite variable (I) included five major categories: unit planning, lesson planning, mathematical interaction during instruction, classroom assessment practice, and student pursuits during instruction. Six levels of instruction were identified to capture the variation among teachers in different grade levels and treatments. (Data for each teacher can be found in technical reports: Shafer, Marten, Folgert, & Kwako, 2003a, b; Shafer, Marten, Webb, Folgert, & Davis, 2003.) In this study, opportunity to learn is interpreted more broadly than as a gauge of content coverage and is viewed as a student's opportunity to learn mathematics *with understanding* (OTL_u). The OTL_u composite variable included three categories: curricular content, modification of curricular materials, and teaching for understanding. Teaching for understanding contained four subcategories: the development of conceptual understanding, the nature of student conjectures about mathematical ideas, the nature of connections within mathematics, and the nature of connections between mathematics and students' life experiences. Four levels of OTL_u were identified for capturing the variation among teachers. (Data for each teacher can be found in technical reports: Shafer, Folgert, Wagner, & Kwako, 2004; Shafer, Folgert, & Kwako, 2004a, b.) School capacity (SC) is the collective power of the school staff to improve student achievement (Newmann, King, & Youngs, 2000). In this study, school capacity was characterized through four subcategories of cultural conditions in the school (shared vision for mathematics teaching and learning between principal and teacher, administrative support, school as a workplace, support for innovation) and three subcategories of structural conditions (collaboration among teachers, work structure, influence of standardized testing). Five levels of school capacity were identified for study teachers. (Data for each teacher can be found in technical reports: Shafer, Folgert, Webb, & Kwako, 2003a, b; Shafer, Folgert, Webb, Kwako, Lee, & Wagner, 2003.)

The initial step in this analysis was to examine the overall impact of these four variables on CA scores via analysis of variance using a SAS program (SAS Institute, 2000; see Table 1-1).

Table 1-1.
Analysis of Variance of the Impact of Prior Achievement, Instruction, Opportunity to Learn with Understanding, and School Capacity on Classroom Achievement

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	21766.92	5441.73	9.76	<.0001
Error	70	39043.72	557.77		
Corrected Total	74	60810.64			

R-Square	Coeff Var	Root MSE	CA MEAN Mean
0.36	9.71	23.62	243.30

Source	DF	Type III SS	Mean Square	F Value	Pr > F
PA MEAN	1	16052.14	16052.14	28.78	<.0001
INSTRUCTION	1	276.98	276.98	0.50	0.48
OTLu	1	141.58	141.58	0.25	0.62
SCHOOL CAP	1	231.14	231.14	0.41	0.52

The overall impact of these variables on CA was significant. However, only 36% of the variance was attributable to these variables, and when the separate contribution of the four variables was examined, only the impact of prior achievement (PA) was significant (see Table 1-2).

Table 1-2.
The Contribution of Prior Achievement, Instruction, Opportunity to Learn with Understanding, and School Capacity to the Overall Variance

Variable	Estimate	Standard error	t-value	Pr > t
PA	0.90	0.17	5.36	<.0001
INSTRUCTION	0.39	0.55	0.70	0.48
OTLu	0.51	1.01	0.50	0.62
SCHOOL CAP	-0.55	0.86	-0.64	0.52

The lack of significant contribution to the overall variance for both Instruction (I) and Opportunity to Learn *with understanding* (OTL_u) given the contribution of PA was surprising given that differences in these variables were used to identify the three treatment groups. To further examine this fact, correlations between the variables were found (see Table 1-3).

Table 1-3.
Pearson Correlation Coefficients

	CA_MEAN	PA_MEAN	INSTRUCTION	OTLu	CAP
CA MEAN					
PA MEAN	0.58				
INSTRUCTION	0.27	0.16			
OTL*u*	0.24	0.13	0.87		
SCHOOL CAP	0.08	0.13	0.36	0.29	

Four things are apparent from these correlations. First, there is a strong relationship between the CA and PA means; second, both I and OTL_u are significantly correlated with CA; third, I and OTL_u are strongly correlated; and SC is weakly correlated to both CA and PA. In fact the strong correlation between I and OTL_u (0.87) indicates they are measures of the same domain. This validates the use of variation in these two measures to define the three treatment groups.

Given that PA contributes the most to CA scores, we examined the overall CA means and the PA means for the three treatment groups (see Table 1-4 and Figure 1-1). While the CA mean for the MiC groups was higher than those of the other two groups, their PA means were also higher than those of both the MiC (Conventional) and Conventional groups. Thus, it was decided to use the PA scores as a covariate and adjust the groups' CA means to account for these differences in PA.

Table 1-4.
Classroom Achievement and Prior Achievement Means and Standard Deviations for the Three Treatment Groups

		CA		PA	
Treatment	N	Mean	Std Dev	Mean	Std Dev
MiC	28	254.53	27.71	52.21	16.50
MiC (Conventional)	28	241.98	27.85	47.76	15.09
Conventional	23	236.45	31.36	48.92	19.39

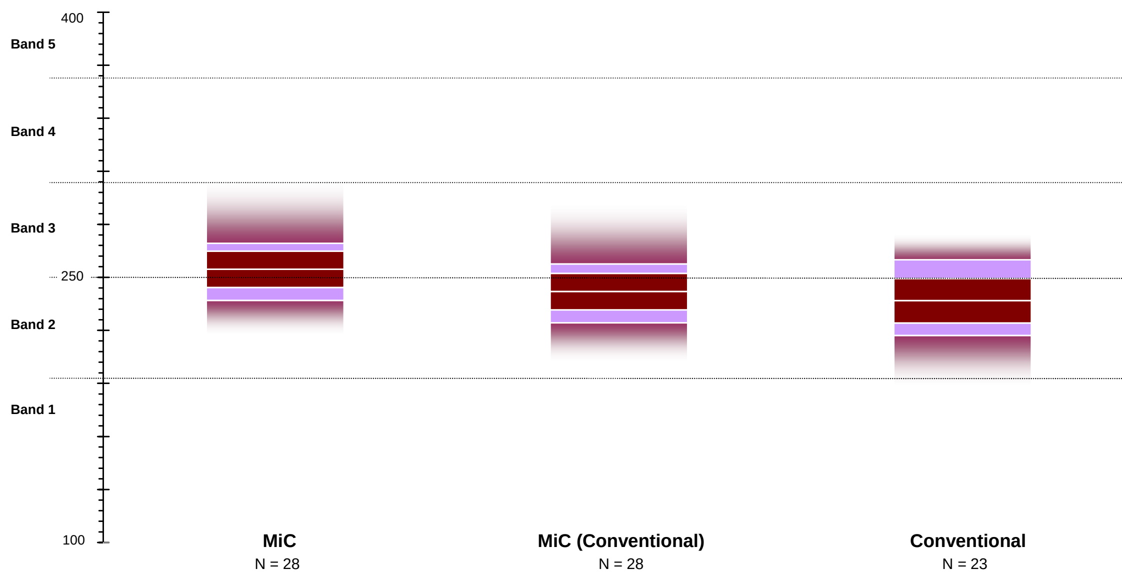


Figure 1-1. Distribution of classroom achievement in MiC, MiC (Conventional), and Conventional classrooms.

The least squares (LS) means for the three treatment groups based on the adjustment due to differences in PA are shown in Table 1-5. The order of the means is the same, but the differences between the means are reduced. This fact further justifies the use of PA as a covariate in the following analysis.

Table 1-5.
Least Squares Means for the Three Treatment Groups

Treatment	CA LSMEAN
MiC	252.01
MiC (Conventional)	243.89
Conventional	237.20

Analysis of Covariance (ANCOVA)

Analysis of covariance (ANCOVA) was then carried out using a SAS program (SAS Institute, 2000). The results are shown in Table 1-6.

Table 1-6.
ANCOVA for the Three Treatment Groups with CA LS Means as the Dependent Variable and PA as the Covariate.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	26026.34	8675.45	15.58	<.0001
Error	75	41772.19	556.96		
Corrected Total	78	67798.53			

R-Square	Coeff Var	Root MSE	CA MEAN Mean
0.38	9.64	23.60	244.82

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMENT	2	2786.66	1393.33	2.50	0.09
PA MEAN	1	21545.44	21545.44	38.68	<.01

When the contrasts between the three treatment groups is examined (see Table 1-7), one finds a significant difference between the means of the MiC and Conventional groups with a substantial effect size and a moderate effect size correlation.[1] There are no important differences between the MiC and MiC (Conventional) groups and between the MiC (Conventional) and Conventional groups.

[1] Given the power of the statistical tests in this study p-values <.10 are considered to imply important differences, and p-values <.05 are statistically significant differences. Effect sizes (Cohen's d) > 0.50 are considered substantial, and effect sizes > 0.30 are considered moderate. Also, an effect size correlation > 0.30 is considered strong (Cohen, 1988).

Table 1-7.
Contrasts in the CA means of the three treatment groups

Parameter	Estimate	Standard Error	t Value	Pr>|t|	Cohen's d	Effect Size r
MiC vs MiC (Conv)	8.12	6.35	1.28	0.20	0.34	0.17
MiC vs Conventional	14.81	6.66	2.22	0.03	0.63	0.30
MiC (Conv) vs Conventional	6.69	6.64	1.01	0.32	0.29	0.14

In summary, the overall difference in CA means significantly favors the MiC group over the Conventional group. When initial differences in achievement are taken into account, the overall performance of the MiC and MiC (Conventional) are different with a moderate effect size, and the performance of the MiC (Conventional) and the Conventional groups are similar.

Additional Analyses

Because the CA index was created from two types of assessments (External Assessment System [EAS] and Problem Solving Assessment [PSA]) and items for four content areas (number, algebra, geometry, and statistics) are included in the assessments, we examined the overall differences in achievement between the three treatment groups on the sub-scores for each assessment and on the four content areas. We used ANCOVA with PA as the covariate for these analyses.

Differences in Achievement on the EAS

The raw means and the adjusted means of student performance on the EAS are shown in Table 1-8 and illustrated in Figure 1-2. Again, the rank order of the means was not changed, but the distance between the means was reduced.

Table 1-8.
EAS Means and Adjusted Means for the Three Treatment Groups

TREATMENT	N	EAS Mean	EAS LSMEAN
MiC	28	253.88	251.89
MiC (Conventional)	28	241.56	242.89
Conventional	23	238.29	239.10

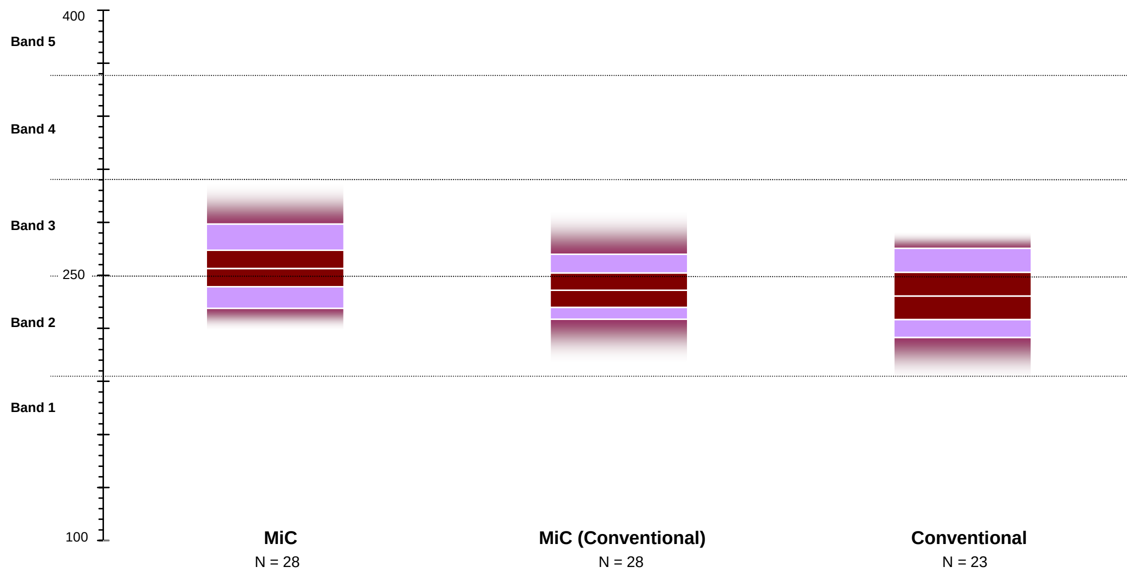


Figure 1-2. Distribution of External Assessment Means in MiC, MiC (Conventional), and Conventional classrooms.

The ANCOVA for this analysis is shown in Table 1-9.

Table 1-9.
ANCOVA for the Three Treatment Groups with EAS LS means as the Dependent Variable and PA as the Covariate

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	19822.44	6607.48	10.56	<.0001
Error	75	46922.28	625.63		
Corrected Total	78	66744.72			

R-Square	Coeff Var	Root MSE	EAS Mean
0.30	10.21	25.01	244.98

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMENT	2	2234.64	1117.32	1.79	0.17
PA MEAN	1	16246.20	16246.20	25.97	<.0001

The contrasts between the means of the three treatment groups are shown in Table 1-10.

Table 1-10.
Contrasts in the EAS Means for the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC VS MiC (Conventional)	9.00	6.72	1.34	0.18	0.36	0.18
MiC vs Conventional	12.79	7.06	1.81	0.07	0.51	0.25
MiC (Conventional) vs Conventional	3.79	7.04	0.54	0.59	0.15	0.08

There are important differences between the MiC group and the Conventional group with a substantial effect size. There is a moderate difference between the MiC and MiC(Conventional) groups. And, there are no difference between the MiC (Conventional) and Conventional groups on the EAS index.

Differences in Achievement on the PSA

The raw means and the adjusted means of student performance on the PSA are shown in Table 1-11 and illustrated in Figure 1-3. Again, the rank order of the means was not changed, but the distance between the means was.

Table 1-11.
PSA Means and Adjusted Means for the Three Treatment Groups

TREATMENT	N	PSA Mean	PSA LSMEAN
MiC	28	254. 19	251.88
MiC (Conventional)	28	243.17	244.71
Conventional	23	237.24	238.17

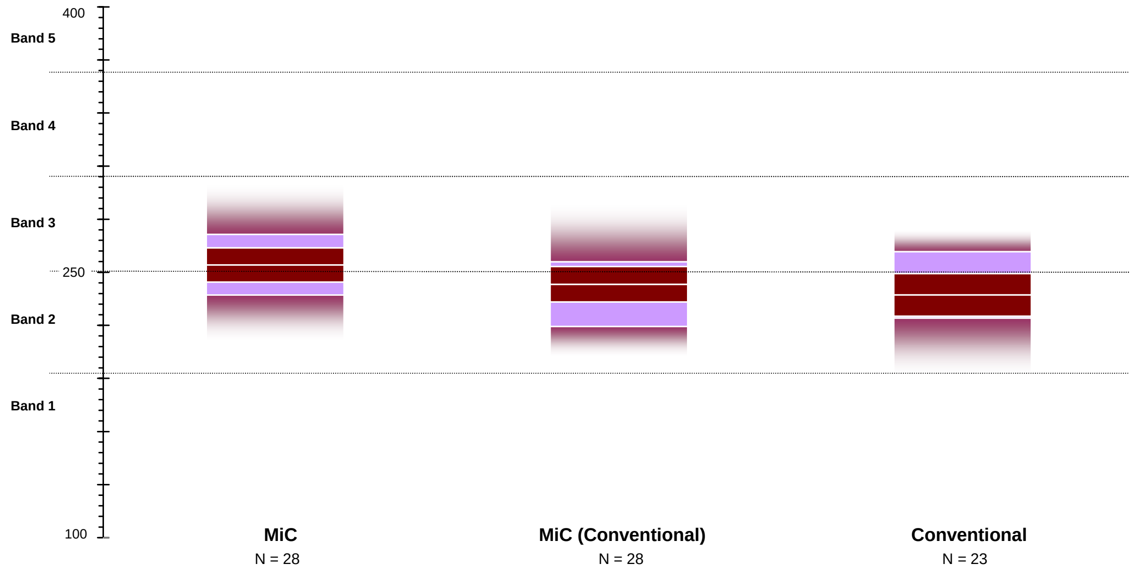


Figure 1-3. Distribution of Problem Solving Assessment Means in MiC, MiC (Conventional), and Conventional classrooms.

The ANCOVA for this analysis is shown in Table 1-12.

Table 1-12.
ANCOVA for the Three Treatment Groups with PSA LS Means as the Dependent Variable and PA as the Covariate

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	25597.83	8532.61	18.07	<.0001
Error	75	35408.01	472.11		
Corrected Total	78	61005.84			

R-Square	Coeff Var	Root MSE	PSA Mean
0.42	8.86	21.73	245.35

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMENT	2	2374.90	1187.45	2.52	0.09
PA MEAN	1	21762.27	21762.27	46.10	<.0001

The contrasts between the means of the three groups are shown in Table 1-13.

Table 1-13.
Contrasts in the PSA Means of the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	7.17	5.83	1.23	0.22	0.33	0.16
MiC vs Conventional	13.71	6.13	2.24	0.02	0.63	0.30
MiC (Conventional) vs Conventional	6.54	6.11	1.07	0.29	0.30	0.15

There are significant differences between the MiC and Conventional groups with substantial effect size. There are moderate effect size differences between the MiC and MiC (Conventional) groups and between the MiC (Conventional) and Conventional groups on the PSA index.

The following pair of items from the Grade 6 PSA illustrates the basis of the differences in problem solving between the three groups. The first question assesses students' abilities to interpret a problem situation, identify appropriate arithmetic calculations (multiplication and subtraction), use whole and decimal numbers, and provide a correct answer with clear and supporting work. The first item assesses students' ability to interpret a problem situation; identify appropriate arithmetic calculations (multiplication and subtraction); and use whole and decimal numbers.

The ranger station publishes a magazine about birds called *Bird Watchers' Bulletin*. The price of one issue is $2.95. To increase the number of issues of *Bird Watchers' Bulletin* that are sold, the rangers start an advertising campaign. The following commercial is played daily on the local radio station.

> You heard it right! For only $21.95 a year (that's 12 fabulous issues), you can have *Bird Watchers' Bulletin* delivered to your home every month, rain or shine! And being a subscriber means you'll receive each issue before it becomes available at the newsstand!

14. How much money do you save buying a one-year subscription instead of buying a single copy each month for one year worth of *Bird Watchers' Bulletin* at a newsstand? Show your work.

Figure 1-4. Item 14 from Grade 6 Problem Solving Assessment.

The second item assesses students' abilities to interpret a problem situation, identify appropriate arithmetic calculations (multiplication, division, and subtraction), use whole and decimal numbers, compare decimals, find unit price as a part of solution, and provide correct answer with clear supporting work.

The ranger station publishes a magazine about birds called *Bird Watchers' Bulletin*. The price of one issue is $2.95. To increase the number of issues of *Bird Watchers' Bulletin* that are sold, the rangers start an advertising campaign. The following commercial is played daily on the local radio station.

You heard it right! For only $21.95 a year (that's 12 fabulous issues), you can have *Bird Watchers' Bulletin* delivered to your home every month, rain or shine! And being a subscriber means you'll receive each issue before it becomes available at the newsstand!

Gill does not buy *Bird Watchers' Bulletin* every month. He wonders if he will save money by buying a one-year subscription.

15. After how many newsstand issues does it become cheaper to buy a subscription? Explain your answer.

Figure 1-5. Item 15 from Grade 6 Problem Solving Assessment.

The scoring guide for these items allows for 0, 1, or 2 points depending on both answer and work. The difference in means for the three groups on these items is shown in Table 1-14.

Table 1-14.
Means of Grade 6 Problem Solving Assessment Sample Items 14 and 15 by Treatment

Item	Content Strand(s)	***MiC*** (N)	***MiC*** Mean	***MiC (Conventional)*** (N)	***MiC (Conventional)*** Mean	***Conventional*** (N)	***Conventional*** Mean
#14	Number	267	1.02	368	0.63	231	0.68
#15	Number/Algebra	267	0.51	368	0.25	231	0.25

Differences in Achievement on the Number Scale

The raw means and the adjusted means for the number scale are shown in Table 1-15 and illustrated in Figure 1-6. Again, the rank order of the means was not changed, but the distance between the means was.

Table 1-15.
Number Scale Means and Adjusted Means for the Three Treatment Groups

TREATMENT	N	Mean	NS LSMEAN
MiC	28	254.60	252.49
MiC (Conventional)	28	241.81	243.22
Conventional	23	241.76	242.62

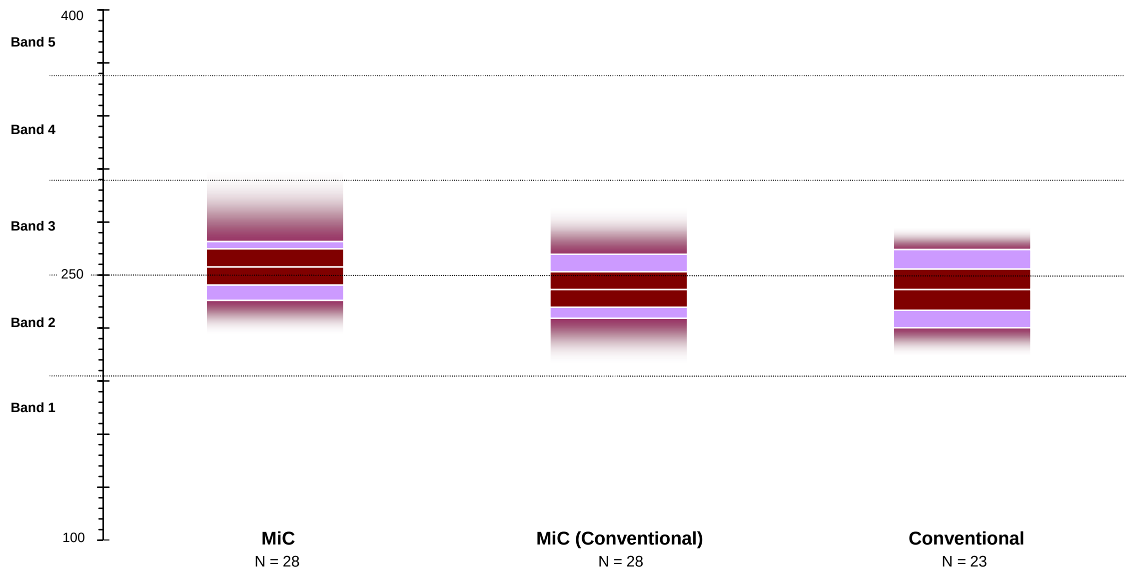


Figure 1-6. Distribution of number scale of classroom achievement in MiC, MiC (Conventional), and Conventional classrooms.

The ANCOVA for this analysis is shown in Table 1-16.

Table 1-16

ANCOVA for the Three Treatment Groups with NS LS Means as the Dependent Variable and PA as the Covariate

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	21210.66	7070.22	13.18	<.0001
Error	75	40231.21	536.42		
Corrected Total	78	61441.88			

R-Square	Coeff Var	Root MSE	NS Mean
0.34	9.40	23.16	246.33

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMENT	2	1663.12	816.56	1.52	0.22
PA MEAN	1	18242.65	18242.65	34.01	<.0001

The contrasts between the means of the three groups are shown in Table 1-17.

Table 1-17.

Contrasts in the NS Means of the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC VS MiC (Conventional)	9.27	6.22	1.49	0.14	0.40	0.20
MiC vs Conventional	9.87	6.54	1.51	0.13	0.43	0.21
MiC (Conventional) vs Conventional	0.59	6.52	0.09	0.93	0.03	0.01

The MiC group's performance is higher (but not significantly higher) with a moderate effect size than the other two groups.

The following item from the Grade 6 PSA illustrates the basis of the differences in number between the three groups. The item assesses students' abilities to identify an appropriate series of arithmetic calculations, add and subtract fractions, rewrite fractions with common denominators to compute, and provide a correct answer with clear supporting work.

> Beth and Gill work as volunteers at the ranger station in Wingra Park. They keep records of the number of visitors who come to the park. The park is open seven days a week. In the first week of April, 134 people visited the park.
>
> There are three entrances to the park, the north, south, and west entrances. On Tuesday, $\frac{1}{9}$ of the people came in through the *north* entrance, and $\frac{2}{3}$ of the visitors arrived through the *south* entrance.
>
> **5.** What fraction of the visitors arrived through the *west* entrance? Show your work.

Figure 1-7. Item 5 from Grade 6 Problem Solving Assessment.

The scoring guide for this item allows for 0, 1, or 2 points depending on both answer and work. The difference in means for the three groups on this item is given in Table 1-18.

Table 1-18.
Means of Grade 6 Problem Solving Assessment Sample Item 5 by Treatment

Item	**Content Strand(s)**	***MiC***		***MiC (Conventional)***		***Conventional***	
		(N)	**Mean**	***(N)***	**Mean**	***(N)***	**Mean**
#5	Number	267	0.48	368	0.22	231	0.36

Differences in Achievement on the Algebra Scale

The raw means and the adjusted means for the algebra scale are shown in Table 1-19 and illustrated in Figure 1-8. Again, the rank order of the means was not changed, but the distance between the means was.

Table 1-19.
Algebra Scale Means and Adjusted Means for the Three Treatment Groups

TREATMENT	N	Mean	ALG LSMEAN
MiC	28	255.88	254.05
MiC (Conventional)	28	244.31	245.53
Conventional	23	239.59	240.33

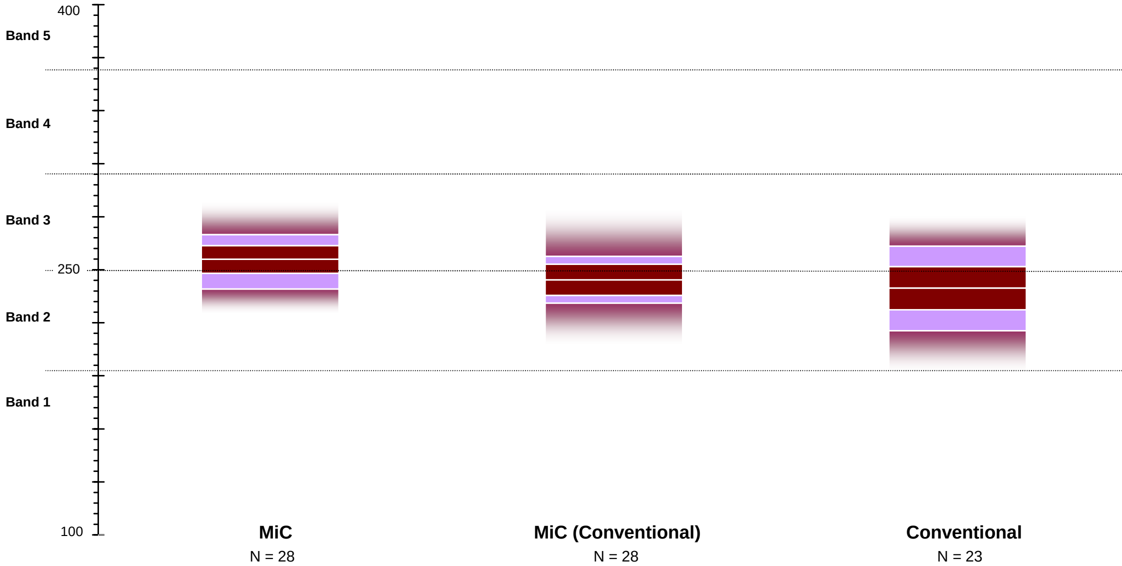


Figure 1-8. Distribution of algebra scale of classroom achievement in MiC, MiC (Conventional), and Conventional classrooms.

The ANCOVA for this analysis is shown in Table 1-20.

Table 1-20.
ANCOVA for the Three Treatment Groups with ALG LS Means as the Dependent Variable and PA as the Covariate

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	17334.62	5778.21	12.97	<.0001
Error	75	33410.29	445.47		
Corrected Total	78	50744.92			

R-Square	Coeff Var	Root MSE	ALG Mean
0.34	8.54	21.11	247.04

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMENT	2	2456.65	1228.33	2.76	0.07
PA MEAN	1	13660.08	13660.08	30.66	<.0001

The contrasts between the means of the three groups are shown in Table 1-21.

Table 1-21.
Contrasts in the ALG Means of the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	8.52	5.67	1.50	0.14	0.40	0.20
MiC vs Conventional	13.72	5.96	2.30	0.02	0.65	0.31
MiC (Conventional) vs Conventional	5.20	5.94	0.88	0.38	0.25	0.12

There is a significant difference between the MiC and Conventional groups with a substantial effect size and effect size correlation. There is a moderate effect size difference between the MiC and MiC (Conventional) groups, but there are no differences between the MiC (Conventional) and Conventional groups on the algebra scale.

The following pair of items from the Grade 6 PSA illustrate the basis of the differences in algebra between the three groups. The first item assesses students' abilities to use a formula, calculate (multiply and subtract) with whole and decimal numbers, use order of operations, provide a correct answer, and show appropriate work.

> Birds lose weight when they fly. For example, a swan loses about 0.1 kilograms of weight for each hour flying.
>
> When you know the weight of a swan at the beginning of a flight—the starting weight in kilograms—you can compute the landing weight with the following formula:
>
> **landing weight = starting weight – N** x **0.1**
>
> In this formula the landing weight is the weight in kilograms after a flight of **N** hours.
>
> **11.** If a swan has a starting weight of 10.5 kilograms, how much will it weigh after flying 7 hours? Show your work.

Figure 1-9. Item 11 from Grade 6 Problem Solving Assessment.

The second item assesses students' abilities to solve for the unknown in a given formula, calculate (subtract and divide) with whole and decimal numbers, and provide a correct answer with clear supporting work.

Birds lose weight when they fly. For example, a swan loses about 0.1 kilograms of weight for each hour flying.

When you know the weight of a swan at the beginning of a flight—the starting weight in kilograms—you can compute the landing weight with the following formula:

landing weight = **starting weight – N** x **0.1**

In this formula the landing weight is the weight in kilograms after a flight of **N** hours.

A swan weighed 13 kilograms. After a flight its weight dropped to 11.9 kilograms.

12. How many hours has this swan been flying? Show your work.

Figure 1-10. Item 11 from Grade 6 Problem Solving Assessment.

The scoring guide for these items allows for 0, 1, or 2 points depending on both answer and work. The difference in means for the three groups on these items is given in Table 1-22.

Table 1-22.
Means of Grade 6 Problem Solving Assessment Sample Items 11 and 12 by Treatment

Item	Content Strand(s)	*MiC*		*MiC (Conventional)*		*Conventional*	
		(N)	**Mean**	*(N)*	**Mean**	*(N)*	**Mean**
#11	Algebra/Number	267	0.73	368	0.37	231	0.64
#12	Algebra/Number	267	0.56	368	0.21	231	0.38

Differences in Achievement on the Geometry Scale

The raw means and the adjusted means for the geometry scale are shown in Table 1-23 and illustrated in Figure 1-11. Again, the rank order of the means was not changed, but the distance between the means was.

Table 1-23.
Geometry Scale Means and Adjusted Means for the Three Treatment Groups

TREATMENT	N	Mean	GEOM LSMEAN
MiC	28	250.73	248.88
MiC (Conventional)	28	244.33	245.56
Conventional	23	235.61	236.36

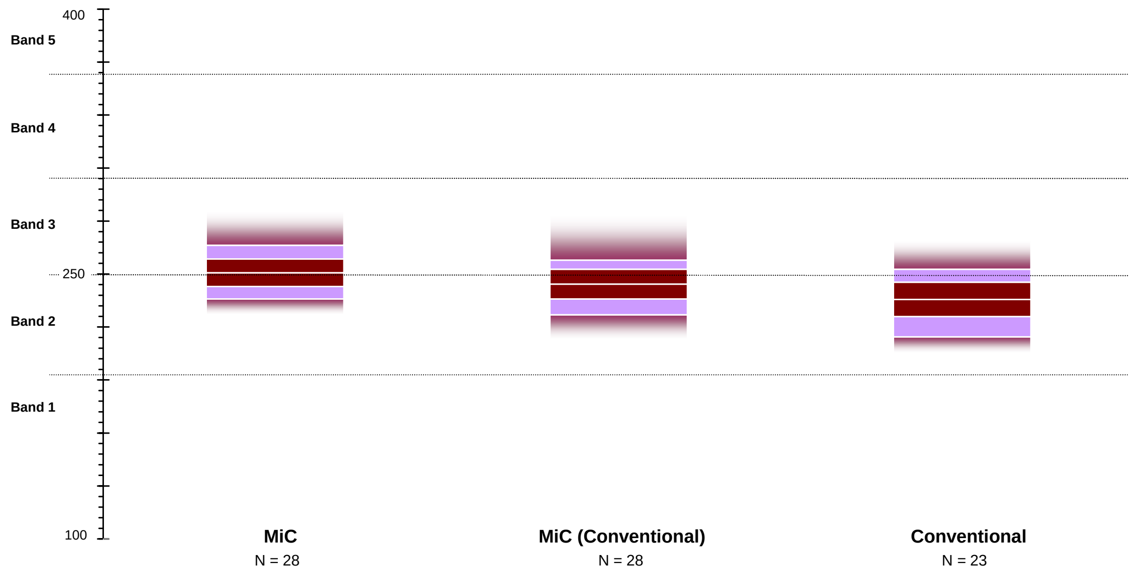


Figure 1-11. Distribution of geometry scale of classroom achievement in MiC, MiC (Conventional), and Conventional classrooms.

The ANCOVA for this analysis is shown in Table 1-24.

Table 1-24.
ANCOVA for the Three Treatment Groups with GEOM LS Means as the Dependent Variable and PA as the Covariate

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	16847.31	5615.77	17.36	<.0001
Error	75	24257.81	323.44		
Corrected Total	78	41105.12			

R-Square	Coeff Var	Root MSE	GEOM Mean
0.41	7.37	17.98	244.06

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMENT	2	2069.16	1034.58	3.20	0.05
PA MEAN	1	13958.15	13958.15	43.16	<.0001

The contrasts between the means of the three groups are shown in Table 1-25.

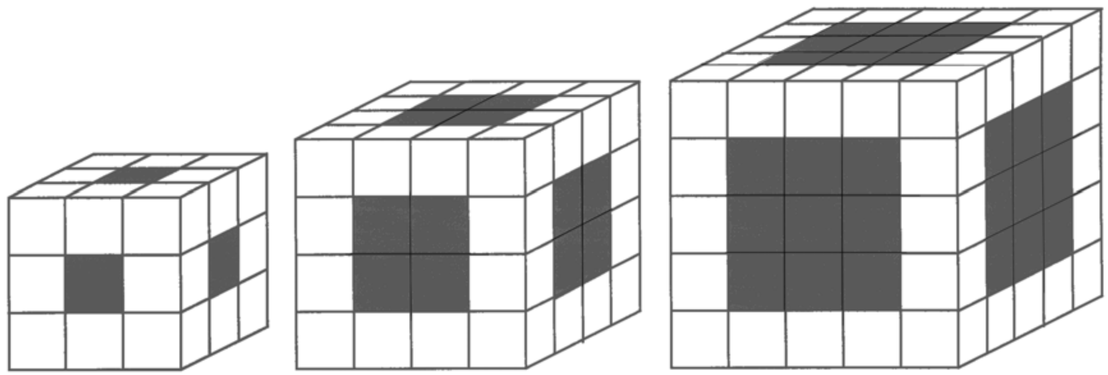
Table 1-25
Contrasts in the GEOM Means of the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	3.32	4.83	0.69	0.49	0.18	0.09
MiC vs Conventional	12.52	5.08	2.47	0.02	0.70	0.33
MiC (Conventional) vs Conventional	9.20	5.06	1.82	0.07	0.51	0.25

There are significant differences between the MiC and Conventional groups with a substantial effect size. There are no differences between the MiC and MiC (Conventional) groups, but there are important differences between and the MiC (Conventional) and Conventional groups on the geometry scale.

The following item from the Grade 8 PSA illustrates the basis of the differences in geometry between the three groups. The item assesses students' abilities to interpret a pattern demonstrated in diagrams and provide a correct answer with explanation.

The first three cubes of a larger sequence are shown below.



The first cube is a 3 x 3 x 3 cube. The second cube is a 4 x 4 x 4 cube.

Small white cubes form the edges of the large cubes. The other cubes on the outside are black.
Ignore the cubes on the inside.

18. How many white cubes are needed for the 4 x 4 x 4 cube?
    Show how you found your answer.

Figure 1-12. Item 18 from Grade 8 Problem Solving Assessment.

The scoring guide for this item allows for 0, 1, or 2 points depending on both answer and work. The difference in means for the three groups on this item is given in Table 1-26.

Table 1-26
Means of Grade 8 Problem Solving Assessment Sample Item 18 by Treatment

Item	**Content Strand(s)**	***MiC***		***MiC (Conventional)***		***Conventional***	
		(N)	**Mean**	***(N)***	**Mean**	***(N)***	**Mean**
#18	Algebra/Geometry	214	0.50	108	0.31	138	0.20

Differences in Achievement on the Statistics Scale

The raw means and the adjusted means for the statistics scale are shown in Table 1-27 and illustrated in Figure 1-13. Again, the rank order of the means was not changed, but the distance between the means was.

Table 1-27.
Statistics Scale Means and Adjusted Means for the Three Treatment Groups

TREATMENT	N	Mean	STAT LSMEAN
MiC	28	253.87	252.05
MiC (Conventional)	28	241.67	242.89
Conventional	23	237.15	237.90

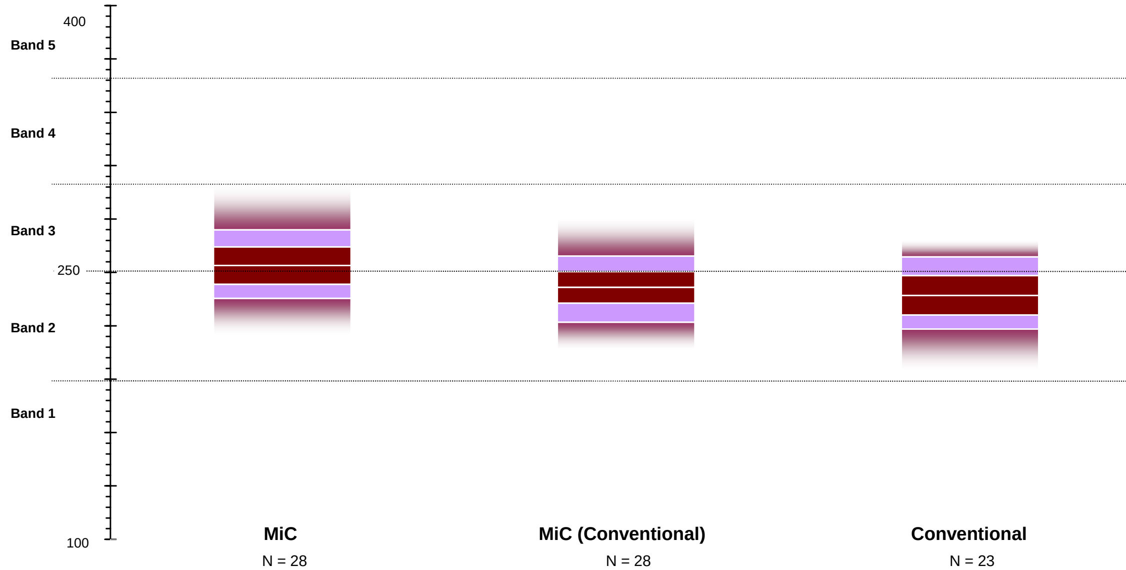


Figure 1-13. Distribution of geometry scale of classroom achievement in MiC, MiC (Conventional), and Conventional classrooms.

The ANCOVA for this analysis is shown in Table 1-28.

Table 1-28.
ANCOVA for the Three Treatment Groups with STAT LS Means as the Dependent Variable and PA as the Covariate

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	17679.63	5893.21	11.21	<.0001
Error	75	39432.07	525.76		
Corrected Total	78	57111.70			

R-Square	Coeff Var	Root MSE	STAT Mean
0.31	9.37	22.93	244.68

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMENT	2	2647.48	1323.74	2.52	0.09
PA MEAN	1	13749.22	13749.22	26.15	<.0001

The contrasts between the means of the three groups are shown in Table 1-29.

Table 1-29.
Contrasts in the STAT Means of the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	9.16	6.16	1.49	0.14	0.40	0.20
MiC vs Conventional	14.16	6.47	2.19	0.03	0.62	0.30
MiC (Conventional) vs Conventional	4.99	6.45	0.77	0.44	0.22	0.11

There are significant differences between the MiC and Conventional groups with a substantial effect size and strong effect size correlation. There is a moderate effect size difference between the MiC and MiC (Conventional) groups, and there are no differences between the MiC (Conventional) and Conventional groups on the statistics scale.

The following item from the Grade 6 PSA illustrates the basis of the differences in statistics between the three groups. The item assesses students' abilities to identify the appropriate arithmetic calculations to determine a mean, use whole numbers in the calculation, and provide a correct answer with clear supporting work.

> Beth and Gill work as volunteers at the ranger station in Wingra Park. They keep records of the number of visitors who come to the park. The park is open seven days a week. In the first week of April, 134 people visited the park.
>
> What is the average number of people per day who visited the park this week in April?

Figure 1-14. Item 1 from Grade 6 Problem Solving Assessment.

The scoring guide for this item allows for 0 or 1 point depending on the answer. The difference in means for the three groups on this item is given in Table 1-30.

Table 1-30.
Means of Grade 6 Problem Solving Assessment Sample Item 1 by Treatment

Item	Content Strand(s)	*MiC*		*MiC (Conventional)*		*Conventional*	
		(N)	**Mean**	***(N)***	**Mean**	***(N)***	**Mean**
#1	Statistics/Number	214	1.19	108	1.01	138	0.96

Conclusions and Implications

From this overall analysis of the performance for the three treatment groups, the initial conclusion is that the variation in classroom achievement mean scores is best predicted by the prior achievement mean scores for the groups. Not surprisingly, this implies that regardless of treatment differences, groups of students with higher prior achievement scores are likely to achieve more than students with lower prior achievement.

The primary conclusion about the differences between the three treatment groups is that the overall achievement for the MiC teacher/student groups was higher than the achievement of either the MiC (Conventional) or the Conventional groups when differences for prior achievement that favored the MiC group were taken into account. Although this overall analysis masks the within-group variation due to differences in school districts, grade levels, or teachers in the different treatment groups, this finding implies that if one is going to implement a reform program like MiC, it is important to implement it as intended. Regardless of other differences, if MiC is implemented as intended, students will do better than partial implementation of the curriculum and do better than using a conventional program.

Furthermore, by examining for differences in achievement on the CA subscales for the three treatment groups, the overall differences favor the MiC teacher/student groups on External Assessment System and the Problem Solving Assessment indices. Similarly, on the content subscales the MiC group's achievement was higher than that of the MiC (Conventional) group's achievement on the number scale. This result is contrary to the voiced intent of many MiC (Conventional) teachers who deliberately augmented the MiC materials with conventional materials on number because they believed that MiC was weak on number. On both the algebra and statistics scales, the MiC group's achievement was significantly higher than that of the other groups. This is consistent with the differences in content coverage between MiC and conventional materials. Finally, on the geometry scale, the MiC group's achievement was significantly higher that that of the Conventional group. Again, this is consistent with the differences in content coverage between MiC and conventional materials. However, the MiC level of achievement was similar to that of the MiC (Conventional) group. This implies that many of the MiC (Conventional) teachers taught some of the MiC geometry units.

A final conclusion about the differences between the treatment groups is that it is difficult to distinguish the differences in achievement between the MiC (Conventional) and Conventional groups. With the exception of the differences in achievement on the geometry scale, their performances were very similar. The implication of this finding is that augmenting a reform curriculum with conventional materials or augmenting a conventional curriculum with reform materials is not likely to result in different student performance.

CHAPTER 2: CLASSROOM ACHIEVEMENT AND COMPARABLE CLASSES

Mary C. Shafer, Thomas A. Romberg, and Lorene Folgert

Given the overall differences in achievement between the three treatment groups reported in Chapter 1 and because there was no possibility of random assignment of either students or teachers in the two districts, it is imperative to examine in detail the differences in the classroom achievement for comparable classes in the treatment groups. In this chapter we go beyond the overall treatment effects to explore the effect of other variables. Furthermore, given the strong relationship between group means on prior achievement (PA) and group means on classroom achievement (CA) reported in Chapter 1 changes in student achievement are examined for classes that were matched with respect to prior achievement. Because the differences in achievement in these classes involve more than prior achievement contributing factors such as instruction, opportunity to learn with understanding, and school capacity need to be explored. Moreover, other factors, such as administration of study assessments and teachers' perceptions of the capabilities of their students, also deserve consideration. The analysis focuses on groups of students who began with comparable prior achievement and investigates the impact of what transpired during the school year that affected student performance. Data were collected through classroom observation reports, teacher logs and journal entries, interviews, and questionnaires for the three treatment groups in two of the four districts involved in the study, Districts 1 and 2.

Comparable Classes

Initially, we identified student groups in Districts 1 and 2 at a particular grade level that began with comparable prior achievement (PA). We selected clusters of groups, organized by teacher. First, the instruction students experienced is described for each teacher in a particular cluster, and contrasts are drawn among teachers. Second, classroom achievement (CA) scores derived from study assessments in the spring of the year are discussed with respect to PA and the instruction experienced. Then, we examine the impact of opportunity to learn with understanding (OTL_u) and school capacity on changes in CA.

Comparable Groups in Year 1

In Year 1 (1997-1998), MiC and MiC (Conventional) teachers taught MiC for the first time during the entire school year, as the commercial version had just become available. Prior to that, these teachers had limited or no experience teaching MiC units. Conventional teachers taught the curricula already available in their schools.[2] Students were in Grades 5, 6, and 7. Fifth-grade students

[2] 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 (1998), *Mathematics: Applications and Connections* (District 1) and Charles, et al (1998), *Middle School Mathematics* (District 2). In Grade 8, both districts used Price,

were in elementary schools that fed into study middle schools. Note that the number of students for whom we have data for both prior achievement and classroom achievement varies for many teachers at all grade levels. Two factors are relevant in understanding the differences in the number of students reported. First, standardized test scores for some students were unavailable. Second, some students did not complete both study assessments and classroom achievement scores were not be calculated for them.

Grade 5

Table 2-1 and Figure 2-1a show the mean PA scores from district standardized tests administered in the prior spring and the level of instruction assigned to each teacher as instruction unfolded during the school year. Two clusters are of interest. One cluster includes students with mean PA scores between the 70th and 95th percentiles who experienced different levels of instruction during the school year, classes taught by LaSalle,[3] MiC Teacher 31; Kipling, MiC (Conventional) Teacher 19; and Fulton, Conventional Teacher 42.[4] The second cluster includes students with mean PA scores around the 75th percentile and lower levels of instruction, classes taught by Conventional Teachers 17 and 43.

---

et al (1992), *Prealgebra*. 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.

[3] Names of teachers and schools are pseudonyms.

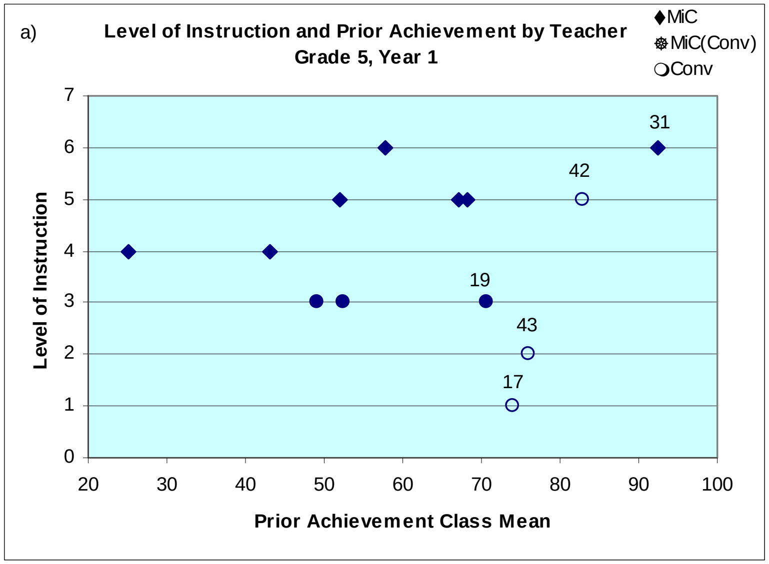
[4] In Monograph 6, Chapter 1, Fulton, Conventional Teacher 42, was classified as an outlier in the cross-tabulation of the levels of instruction and opportunity to learn. Therefore, he was not included in the analysis of student performance by treatment group. However, we decided to include his classes in the analysis of comparable groups discussed in this chapter.

Table 2-1

Level of Instruction, Prior Achievement, and Classroom Achievement for Grade 5 Students in 1997-98, by District and by Teacher

District/ Teacher ID	Level of Instruction	PA* (N)	PA* Mean	CA (N)	CA Mean
MiC					
District 1					
14	4	46	43.04	53	238.96
31	6	30	92.37	31	316.80
49	4	9	25.11	9	242.47
District 2					
2	5	25	68.24	32	231.94
12	5	30	52.00	41	222.12
22	5	19	67.00	23	241.13
26	6	63	57.71	72	241.82
MiC (Conventional)					
District 1					
4	3	17	49.18	15	252.58
19	3	25	70.76	26	291.47
27	3	15	52.53	19	252.62
District 2					
(None)					
Conventional					
District 1					
17	1	20	74.10	22	272.22
42	5	29	82.90	30	286.82
District 2					
43	2	40	75.98	49	259.50

* The mean preceding achievement score for 1997-1998 is the mean standardized test score from the previous spring.



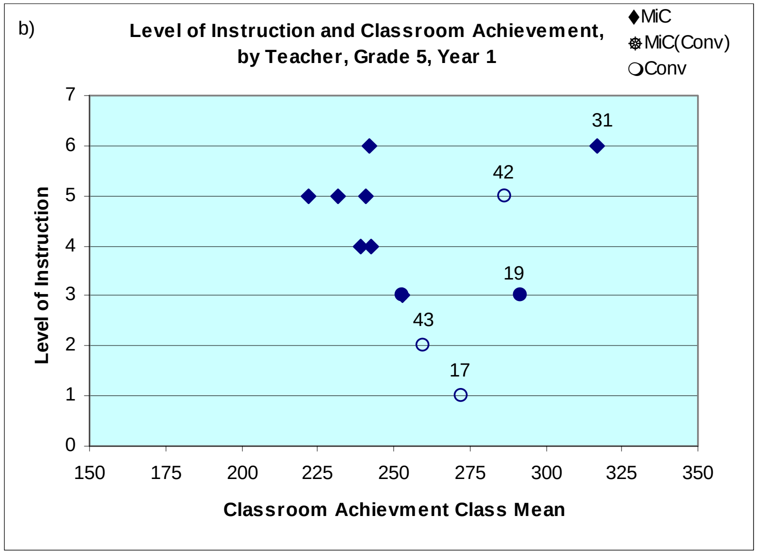


Figure 2-1. Level of instruction for Grade 5 teachers in Year 1 compared with a) prior achievement; b) classroom achievement at the end of the year

Teachers 19, 31, and 42. LaSalle, MiC Teacher 31, and Fulton, Conventional Teacher 42, taught students with high mean PA scores (92nd and 83rd percentiles, respectively), and the instruction experienced by their students was of high quality (Levels 6 and 5, respectively). LaSalle presented lessons in ways that supported conceptual understanding and mathematical reasoning. She set the stage for students to work on their own. Students were expected to carry out the mathematical work and develop their own conjectures, solution strategies, and generalizations. They presented their work during class discussions in which they compared various strategies. The presentations afforded opportunities for students to assess their own work and for the teacher to assess their progress toward instructional goals. Substantive feedback was offered by students and teacher.

Fulton presented lessons that emphasized conceptual understanding, and he actively participated in lessons with his students. Although students tended to give steps in a procedure as evidence of their thinking, he encouraged students to talk about why procedures worked and to consider reasonableness of answers. He modified lessons based on student statements or inquires and promoted connections among mathematical ideas or between mathematics and students' lives. Student–student conversation occurred on a limited basis and usually consisted of sharing answers.

Kipling taught students with high average mean PA scores (70th percentile), and her students experienced instruction more reflective of good conventional pedagogy. She generally presented problem-solving strategies or procedures to her class. At times, however, her questions focused on mathematical processes such as reflection on the lesson or articulation of thinking. However, she expected meaningful explanations. She pressed students to use precise mathematical terminology and to think about alternate solution strategies. She sought explanations and accurate procedures as evidence of student learning. Feedback was primarily teacher-directed and responsive to inaccuracies or misunderstandings. Although students were seated in small groups, students did not work cooperatively and student–student conversation was not encouraged.

Consistent with the mean PA score, LaSalle's students had the highest mean CA score among all fifth-grade groups (see Table 2-1 and Figure 2-1b; note the different scales for PA and CA in these plots). This contrasts with the change Fulton's students experienced from the second highest mean PA score to the third highest mean CA score. Kipling's students outperformed Fulton's students in CA, even though they began with lower PA scores. This is interesting because Kipling tended to present procedures and problem-solving strategies to students, although she encouraged and expected that students explain their thinking. In contrast, Fulton continually requested that students understand the meaning underlying procedures. The instruction students experienced may not be enough to account for the lower CA scores in this case.

Table 2-2 and Figure 2-2 show mean CA scores and the level of OTL_u assigned to each teacher. Students in classes taught by Kipling, MiC (Conventional) Teacher 19, and LaSalle, MiC Teacher 31, experienced high OTL_u. Both presented a comprehensive, integrated curriculum, teaching six units involving four content strands. They used the sequence of units recommended in teacher support materials and occasionally supplemented the curriculum with activities disconnected from the curriculum, such as a game based on the stock market (LaSalle) and competitions and practice for upcoming district standardized tests (Kipling). During LaSalle's lessons, content was explored in enough detail for students to think about relationships among mathematical ideas or to link procedural and conceptual knowledge, and students were encouraged to make generalizations about mathematical ideas. Connections between mathematics and students' daily lives were clearly apparent in lessons. The main focus of Kipling's lessons was on procedural understanding, although at times she emphasized conceptual understanding. Connections among mathematical ideas and between mathematics and students' lives were not discussed in detail. Fulton used materials from a variety of resources in lieu of a textbook. Mathematics was taught in depth, but the content was restricted to number and some geometry topics, and lessons emphasized conceptual and procedural understanding. Instructional materials featured connections among mathematics topics and connections to the students' lives. Thus, differences in CA scores between classes taught by Kipling and Fulton seem to be attributable to the differences in content taught. For LaSalle, the high quality of instruction and OTL_u students experienced likely influenced their high CA performance.

Table 2-2
Level of Opportunity to Learn with Understanding, Prior Achievement, and Classroom Achievement for Grade 5 Students in 1997-98, by District and by Teacher

District/ Teacher ID	Level of OTLu	PA		CA	
		(N)	Mean	*(N)*	Mean
		MiC			
District 1					
14	4	46	43.04	53	238.96
31	4	30	92.37	31	316.80
49	3	9	25.11	9	242.47
District 2					
2	4	25	68.24	32	231.94
12	4	30	52.00	41	222.12
22	4	19	67.00	23	241.13
26	4	63	57.71	72	241.82
		MiC (Conventional)			
District 1					
4	4	17	49.18	15	252.58
19	4	25	70.76	26	291.47
27	3	15	52.53	19	252.62
District 2					
(None)					
		Conventional			
District 1					
17	2	20	74.10	22	272.22
42	3	29	82.90	30	286.82
District 2					
43	2	40	75.98	49	259.50

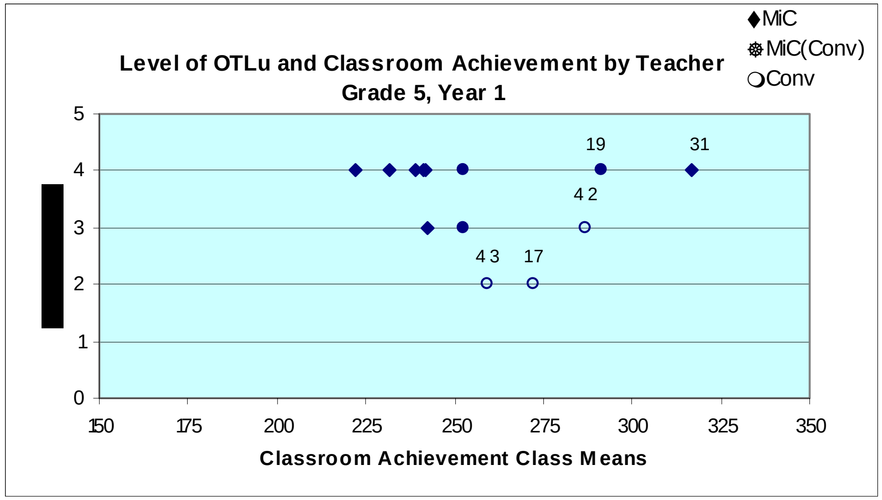


Figure 2-2. Level of opportunity to learn with understanding for Grade 5 teachers in Year 1 compared with classroom achievement at the end of the year

With respect to school capacity, LaSalle, MiC Teacher 31, and Kipling, MiC (Conventional) Teacher 19, taught at the same school, but perceived limited and average levels of school capacity, respectively (see Table 2-3 and Figure 2-3). These teachers taught at Beethoven Elementary, a school for Grades 3–5, that had an enrollment of 900 students with 300 students in Grade 5. The student population had about equal number of Minority and White students. Approximately 42% of the students were African American. In Grade 5, students were grouped homogeneously for mathematics instruction. LaSalle taught the advanced students, and Kipling taught students of high average abilities. Both teachers felt that incompatible visions of teaching and learning mathematics were held by the principal and teachers in their school, and that the principal's vision was not clearly communicated to teachers. They differed in their perceptions of administrative support regarding choice of instructional materials, change in instructional practice, and changes in policy. Collaboration among teachers rarely transpired. Both teachers reported that they had a high influence in planning and teaching mathematics, but they differed in perceptions of their influence over educational policy such as discipline and the content of professional development programs and administrative decisions such as hiring new faculty. However, both felt that faculty and staff were committed to academic excellence and that teachers supported one another in their efforts to improve instruction. Professional development specifically related to mathematics was offered at the school and substitute time was allotted for external professional development. These teachers believed that the district and state standardized testing programs had little influence in their planning because they emphasized problem solving and written explanations of solutions (LaSalle) or felt MiC was well aligned with the testing programs (Kipling).

Table 2-3

Level of School Capacity, Prior Achievement, and Classroom Achievement for Grade 5 Students in 1997-98, by District and by Teacher

District/ Teacher ID	Level of School Capacity	PA *(N)*	PA Mean	CA *(N)*	CA Mean
		MiC			
District 1					
14	3	46	43.04	53	238.96
31	2	30	92.37	31	316.80
49	3	9	25.11	9	242.47
District 2					
2	5	25	68.24	32	231.94
12	2	30	52.00	41	222.12
22	5	19	67.00	23	241.13
26	5	63	57.71	72	241.82
		MiC (Conventional)			
District 1					
4	4	17	49.18	15	252.58
19	3	25	70.76	26	291.47
27	3	15	52.53	19	252.62
District 2					
(None)					
		Conventional			
District 1					
17	2	20	74.10	22	272.22
42	4	29	82.90	30	286.82
District 2					
43	4	40	75.98	49	259.50

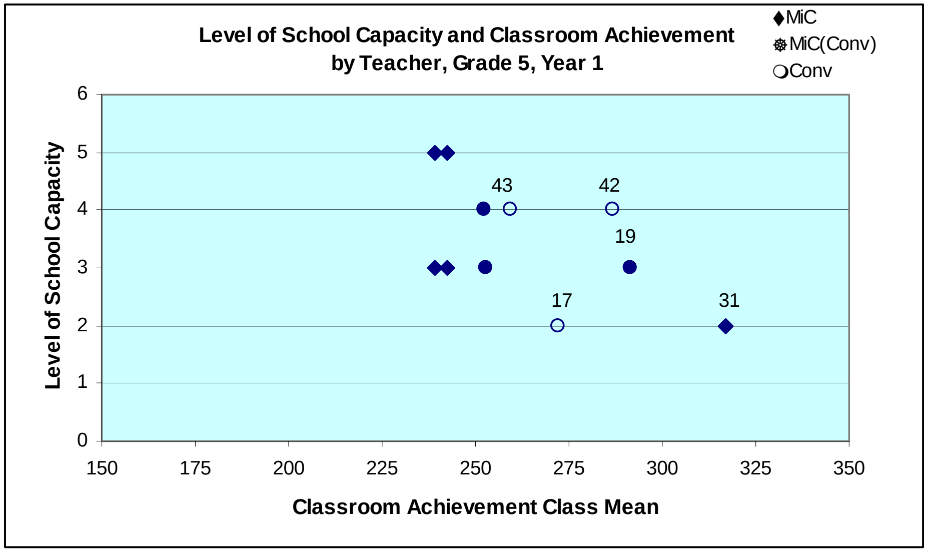


Figure 2-3. Level of school capacity for Grade 5 teachers in Year 1 compared with classroom achievement at the end of the year

In contrast, Fulton, Conventional Teacher 42, perceived a moderately high level of school capacity. He taught at River Forest Elementary, a school for Grades 4 and 5, that had an enrollment of 320 students. Approximately 35% of the student population was minority with 20% African American students and 10% Hispanic students. Ten to 20% of the students were eligible for government-funded lunch programs. Fulton felt that the principal and teachers had visions of teaching and learning mathematics that were partially aligned because teachers adopted the initiatives put forth by their principal who was new to the school that year. He felt that he received very strong administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. Collaboration among teachers was encouraged. Fulton felt that he had a high influence in planning and teaching mathematics, educational policy such as curriculum and the content of professional development programs, and less influence over discipline. He felt that faculty and staff were committed to academic excellence. Teachers supported one another in their efforts to improve instruction. Teachers were able to suggest agendas for in-service workshops, and substitute time was allotted for external professional development. Fulton believed that the district and state standardized testing programs had little influence in his planning because he emphasized problem solving and written explanations of solutions.

One might expect to see increased performance when teachers perceived average and moderately high school capacity and a decline in student performance when teachers perceived limited school capacity. This was not evident for this cluster of teachers and students. Despite differences in school capacity between LaSalle and Kipling, these felt highly supported by other faculty and staff. Because student performance in LaSalle's classes remained at a high level and performance in Kipling's classes outperformed Fulton's classes, school capacity seemed not to be a factor in the CA performance. Rather, the differences in CA scores seem to be attributable to the quality of instruction and OTL_u experienced.

Conventional Teachers 17 and 43. Kershaw and Gant, Conventional Teachers 17 and 43, taught students who had mean PA scores around the 75th percentile (see Table 2-1 and Figure 2-1a). As instruction unfolded during the school year, Kershaw presented procedures or problem-solving strategies, and students practiced them in rote ways. At times, however, lessons were underdeveloped, and students lacked the support they needed to understand the mathematics on their own. Inquiry during lessons was limited to lower order thinking, and lessons did not promote conceptual understanding. Questions required numerical answers, and students were not expected to elaborate on their reasoning. Student–student conversation was not encouraged. Procedural understanding was accepted as evidence of student learning, and instructional decisions were based on student homework and classwork. Feedback was limited to accuracy of answers. In contrast, Gant frequently devoted a major portion of the class period to review of a previous lesson, homework, or a warm-up activity. Inquiry during lessons was limited to lower order thinking, and conceptual understanding was not promoted. Questions required limited responses or steps in procedures. Student–student conversation was not encouraged. Feedback was indirectly responsive to student needs in that it involved "more of the same," such as additional instruction and practice sets. Feedback was often directed toward the format of the answer (such as simplified form or labeling the answer) rather than clarifying explanations or working toward student understanding.

Although close in mean PA scores and the quality of instruction was lower for Kershaw's classes, mean CA scores were further apart, with Kershaw's students having the higher mean CA score (see Table 2-1 and Figure 2-1b). And, even though the quality of instruction was not as robust for these teachers, their students still outperformed students in classes taught by MiC and MiC (Conventional) teachers who began with average or low PA scores. This seems to suggest that other factors had an important effect on student performance in these teachers' classes. These classes experienced a limited level of OTL_u (see Table 2-2 and Figure 2-2). Curricular content spanned a vast content plane with little or no depth. Content was presented as disparate pieces of knowledge heavily laden with vocabulary and prescribed algorithms. Few questions fostered conceptual understanding, making conjectures was not encouraged, and connections between mathematics and students' lives were not discussed. Because students in these classes had similar experiences relative to OTL_u, it is unlikely that OTL_u influenced CA results.

Similarly, teachers' perceptions of school capacity did not seem to influence CA results (see Table 2-3 and Figure 2-3). Kershaw perceived a low level of school capacity. She taught at Dewey Elementary School, a school for Grades 1–5 with an enrollment of 690 students. About 50% of the fifth-grade students were minority with 36% African American students and 13% Hispanic students. More than 50% of the students were eligible for government-funded lunch programs. Kershaw felt that incompatible visions of teaching and learning mathematics were held by the principal and teachers in her school. However, she felt that she received strong administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. Kershaw felt that she had a high level of influence in planning and teaching mathematics, educational policy such as curriculum and the content of professional development programs, and less influence over discipline. However, she felt little support for innovation in terms of professional development opportunities at the school or external professional development. In lieu of common planning time, teachers informally met regularly for collaboration. District and state standardized tests influenced the content and sequence of content she taught during the school year.

Gant perceived a moderately high level of school capacity. She taught at Von Steuben Community School, a school for Grades K–8. Ten to 20% of the students were eligible for government-funded lunch programs. Parents were encouraged to serve as classroom volunteers, tutors, and mentors. Gant felt that incompatible visions of teaching and learning mathematics were held by the principal and teachers in her school. However, she felt that she received strong administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. She felt that she had a high level of influence in planning and teaching mathematics, educational policy such as curriculum and the content of professional development programs, and less influence over discipline. In-service workshops related to general teacher methods were available at the school, and substitute time was allotted for external professional development. In lieu of common planning time, teachers informally met regularly for collaboration. She believed that her curriculum was well aligned with the district and state standardized testing programs, and these tests had little influence over curriculum and instruction.

Consideration of school capacity for these teachers did not yield information that might shed light on factors that influenced student performance. Thus, in these teachers' classes, the results seem to suggest that factors that were not researched in this study had an important effect on student performance.

Grade 6

Table 2-4 and Figure 2-4a show the mean PA score from district standardized tests and the level of instruction assigned to each sixth-grade teacher. One cluster is described here: students with mean PA scores around the 70th percentile who experienced different levels of instruction, classes taught by Dillard, MiC Teacher 45, and Renlund, Conventional Teacher 21. As instruction unfolded during the school year, Dillard attempted to teach mathematics for understanding, and he allowed students to model particular mathematical processes for the class. By doing so, he created situations in which students could communicate mathematical ideas and make decisions about appropriate approaches to problems. Even with this interaction, his feedback was directed toward procedures. Conventional Teacher 21 presented particular procedures or strategies, and students practiced them in rote ways. At times, lessons mainly involved review of previously learned content. Inquiry during lesson was limited to lower order thinking, and lessons did not promote conceptual understanding or connections among mathematical ideas. Questions required only answers or steps in procedures. Feedback was teacher-directed and consisted of providing "more of the same thing," rather than an emphasis on understanding the underlying concepts. Student–student conversation was not encouraged.

Table 2-4

Level of Instruction, Prior Achievement, and Classroom Achievement for Grade 6 Students in 1997-98, by District and by Teacher

District/ Teacher ID	Level of Instruction	PA*		CA	
		(N)	Mean	*(N)*	Mean
		MiC			
District 1					
33	5	50	44.98	54	215.44
District 2					
45	4	36	71.94	36	267.20
51	5	37	51.27	47	231.92
		MiC (Conventional)			
District 1					
30	2	39	51.95	8	232.86
38	1	38	45.74	56	209.15
District 2					
37	3	38	55.92	45	228.04
52	3	31	33.94	33	205.50
		Conventional			
District 1					
18	1	10	43.00	17	221.54
32	2	38	38.16	59	253.88
District 2					
10	1	22	38.05	27	193.17
21	2	18	68.89	21	228.97

* The mean preceding achievement score for 1997-1998 is the mean standardized test score from the previous spring.

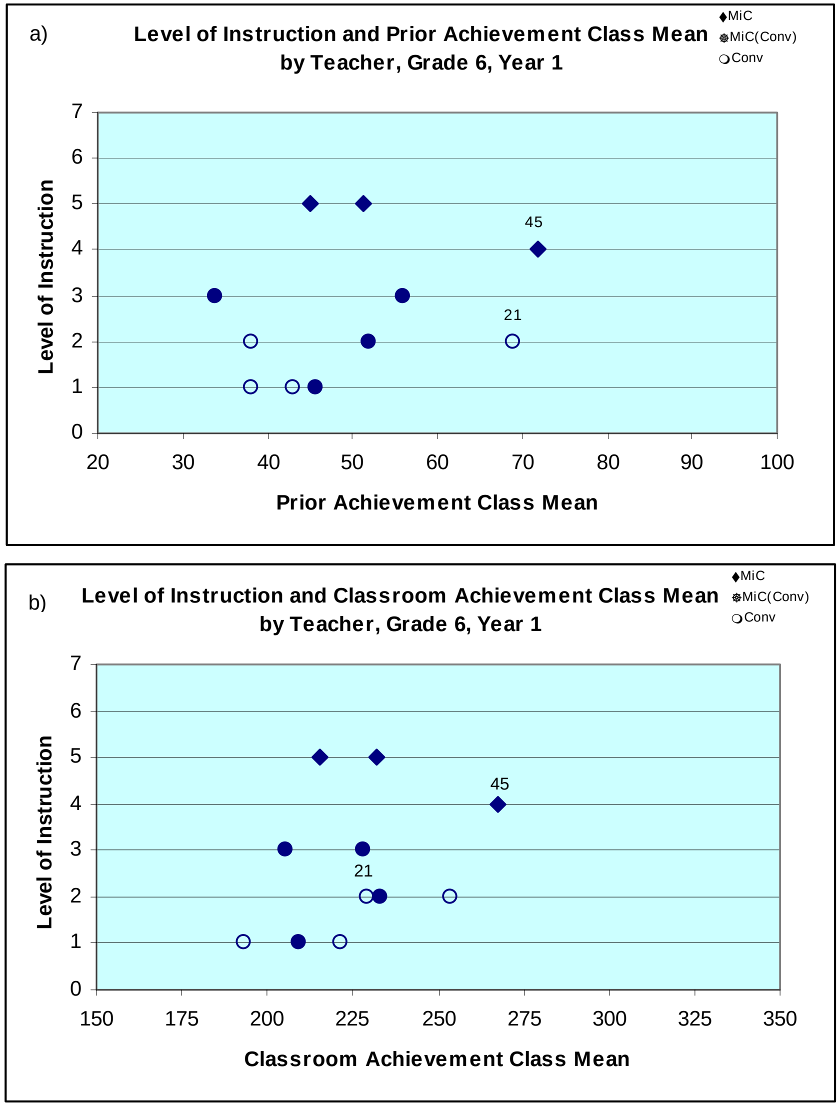


Figure 2-4. Level of instruction for Grade 6 teachers in Year 1 compared with a) prior achievement; b) classroom achievement at the end of the year

Consistent with the mean PA score, Dillard's students had the highest mean CA score among all sixth-grade groups (see Table 2-4 and Figure 2-4b). This contrasts with the considerable change experienced by Renlund's students from the second highest PA score to the fifth highest mean CA score. The quality of instruction students experienced with these teachers likely affected student performance, but there may also be influences with respect to OTL_u and school capacity.

With respect to OTL_u, students in Dillard's classes experienced a moderate level of OTL_u (see Table 2-5 and Figure 2-5). Dillard taught a comprehensive, integrated curriculum, teaching six units in three content strands. Occasionally, he supplemented MiC with activities disconnected from the curriculum, such as a district-mandated reading program for 30 minutes in every class and practice for standardized tests. His lessons at times emphasized conceptual understanding, but the main focus was on building students' procedural understanding. Connections among mathematical ideas and connections between mathematics and students' lives were not discussed in detail. Students in Renlund's classes experienced limited OTL_u. Curricular content spanned a vast content plane with little or no depth, and the textbook was followed with few modifications. The content was presented as disparate pieces of knowledge heavily laden with vocabulary and prescribed algorithms. Few questions fostered conceptual understanding, making conjectures was not encouraged, and connections between mathematics and students' lives were not discussed in lessons.

Table 2-5

Level of Opportunity to Learn with Understanding, Prior Achievement, and Classroom Achievement for Grade 6 Students in 1997-98, by District and by Teacher

District/ Teacher ID	Level of OTLu	PA *(N)*	PA Mean	CA *(N)*	CA Mean
		MiC			
District 1					
33	4	50	44.98	54	215.44
District 2					
45	3	36	71.94	36	267.20
51	3	37	51.27	47	231.92
		MiC (Conventional)			
District 1					
30	2	39	51.95	8	232.86
38	2	38	45.74	56	209.15
District 2					
37	2	38	55.92	45	228.04
52	2	31	33.94	33	205.50
		Conventional			
District 1					
18	1	10	43.00	17	221.54
32	2	38	38.16	59	253.88
District 2					
10	2	22	38.05	27	193.17
21	2	18	68.89	21	228.97

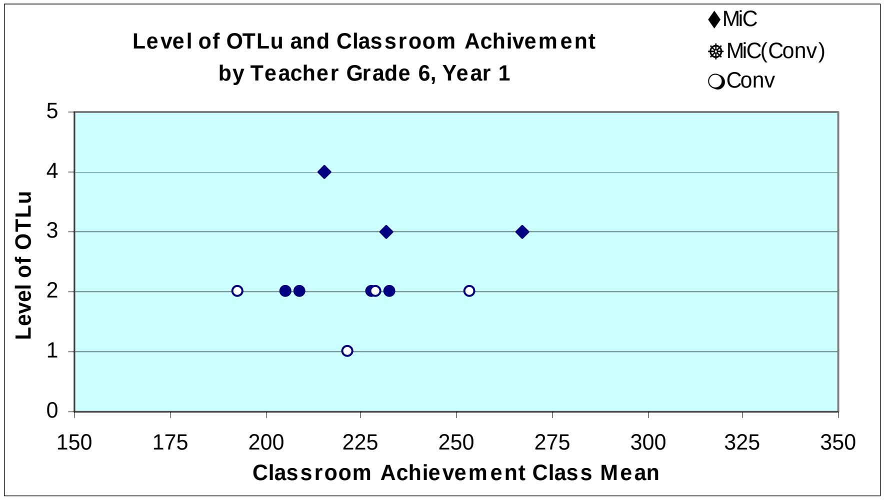


Figure 2-5. Level of opportunity to learn with understanding for Grade 6 teachers in Year 1 compared with classroom achievement at the end of the year

As with instruction, the OTL_u students experienced in these classes likely affected their performance. In Dillard's classes, the combination of comprehensive content taught in some depth and instruction that attempted to promote conceptual understanding seemed to have a positive effect on student performance. In contrast, performance declined for Renlund's students who experienced breadth in content rather than depth and instruction that did not promote conceptual understanding.

The school capacity perceived by these teachers may have added to this impact. Dillard and Renlund perceived average and low school capacity, respectively (see Table 2-6 and Figure 2-6). Dillard taught at Guggenheim Middle School, which had an enrollment of 1285 students. Classes met every other day, alternating two times a week and three times a week, for 110-minute periods. More than 60% of the student population was minority with 32% African American students and 32% Hispanic students. More than 50% of the students were eligible for government-funded lunch programs. Dillard felt that the principal and teachers had visions of teaching and learning mathematics that were aligned on some ideas, but were incompatible on others. He felt that he received an average level of administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. He met regularly with other teachers for collaboration in lieu of professional development at the school. Common planning time was available for mathematics teachers on his teaching team during which they discussed content, instructional and assessment methods, and program evaluation. Dillard felt that he had a high level of influence in planning and teaching mathematics, and a low level of influence over educational policy such as curriculum and discipline. He believed that MiC was well aligned with the district and state standardized testing programs, and these tests had little influence over curriculum and instruction.

Table 2-6

Level of School Capacity, Prior Achievement, and Classroom Achievement for Grade 6 Students in 1997-98, by District and by Teacher

District/ Teacher ID	Level of School Capacity	PA *(N)*	PA Mean	CA *(N)*	CA Mean
			MiC		
District 1					
33	3	50	44.98	54	215.44
District 2					
45	3	36	71.94	36	267.20
51	3	37	51.27	47	231.92
			MiC (Conventional)		
District 1					
30	2	39	51.95	8	232.86
38	2	38	45.74	56	209.15
District 2					
37	3	38	55.92	45	228.04
52	2	31	33.94	33	205.50
			Conventional		
District 1					
18	5	10	43.00	17	221.54
32	2	38	38.16	59	253.88
District 2					
10	2	22	38.05	27	193.17
21	2	18	68.89	21	228.97

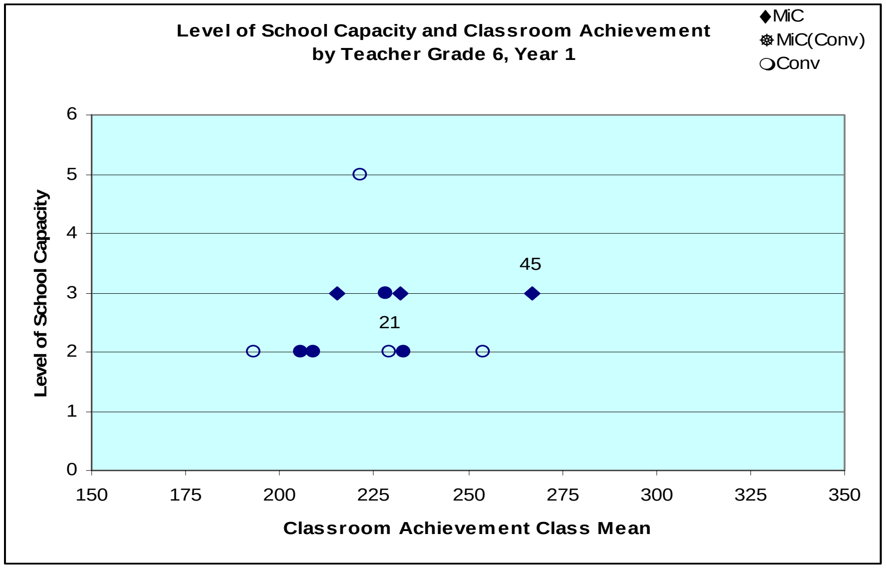


Figure 2-6. Level of school capacity for Grade 6 teachers in Year 1 compared with classroom achievement at the end of the year

Relund taught at Newberry Middle School, which had an enrollment of 1185 students. Classes met five times a week for 50-minute periods. More than 50% of the students were eligible for government-funded lunch programs. Renlund perceived low school capacity. She felt that the principal did not have a clear vision of mathematics teaching and learning. However, she felt that she received an average level of administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. She met regularly with other teachers for collaboration, but no common planning time was provided for them. The teacher leader in her department was a mentor who observed other teachers, provided feedback, modeled instruction, and discussed curriculum and/or instruction with other teachers. No in-service workshops were offered for teachers at the school, but substitute time was allotted for external professional development. District and state standardized tests significantly influenced instruction, as large amounts of time were devoted to test preparation.

Students in this cluster were new to their schools this year, as they made transitions from elementary to middle schools. That is, both groups experienced changes in the cultural and structural conditions of schools, which may have had an impact on their performance that year. However, differences in the types of collaboration experienced by these teachers were noted. Dillard had multiple opportunities during the school year to discuss mathematics curriculum, instruction, and assessment. Similar opportunities were unavailable to Renlund. The type of collaboration Dillard experienced likely supported the instruction and OTL_u his students experienced and, therefore, mitigated the impact of students' transition from elementary to middle school. Given the descriptions of instruction, OTL_u, and school capacity for these teachers, it is likely that all three variables were influential in the changes in student performance over the school year.

Grade 7

Table 2-7 and Figure 2-7a show the mean PA score from district standardized tests and the level of instruction assigned to each seventh-grade teacher. Two clusters are described. The first cluster included students with mean PA scores around the 50th percentile who experienced different levels of instruction, classes taught by Keeton, MiC Teacher 20, and Donnely, MiC (Conventional) Teacher 44. The second cluster included students with mean PA scores around the 40th percentile who experienced different levels of instruction: Heath and McFadden, MiC Teachers 8 and 11, respectively; Draski and Teague, MiC (Conventional) Teachers 35 and 47, respectively; and McLaughlin, Conventional Teacher 53.

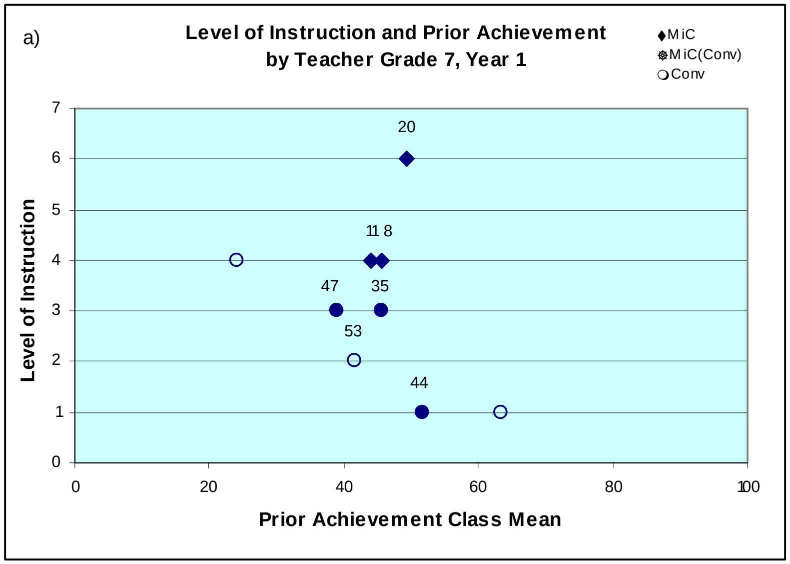
Teachers 20 and 44. Students in this cluster had mean PA scores around the 50th percentile, and they experienced different levels of instruction during the school year. Keeton, MiC Teacher 20, worked toward developing conceptual understanding. She shared the mathematical work with her students, posed questions that encouraged students to explain their thinking, and helped students make connections among mathematical ideas. Students were encouraged to make generalizations. She pressed students to effectively communicate their thinking processes, and she modified lessons in response to students' methods. Keeton used students' explanations as evidence of their learning, and interactions were used to promote making sense of the mathematics. A blend of clear and mathematically sound feedback addressed skills, procedures, and concepts. In contrast, Donnely, MiC (Conventional) Teacher 44, presented lessons, inquiry was limited to lower order thinking, and lessons did not promote conceptual understanding. Lessons routinely featured review of previous lessons, discussion of homework, or a warm-up activity for most of the class period. Frequently, formal lessons were not presented. Students were given an assignment, but the content was not discussed prior to students completing it. Donnely sought correct answers as evidence of student learning. When questions were posed by the teacher, students usually stated only answers or relayed steps in procedures. Feedback was very teacher-directed, as he believed that the teacher was the ultimate authority. Accuracy was the goal, and understanding the meaning behind the procedures was not addressed. Student–student conversation was limited and usually consisted of sharing answers.

Table 2-7

Level of Instruction, Prior Achievement, and Classroom Achievement for Grade 7 Students in 1997-98, by District and by Teacher

District/ Teacher ID	Level of Instruction	PA*		CA	
		(N)	Mean	*(N)*	Mean
MiC					
District 1					
8	4	35	45.54	38	250.92
District 2					
11	4	37	43.97	41	217.02
20	6	41	49.22	45	256.32
MiC (Conventional)					
District 1					
44	1	47	51.60	50	244.44
District 2					
35	3	35	45.54	50	249.14
47	3	42	39.05	48	235.87
Conventional					
District 1					
13	1	33	63.42	39	268.58
53	2	36	41.53	45	227.08
District 2					
5	4	24	24.17	22	197.75
24	2	2	25.00	1	236.39

* The mean preceding achievement score for 1997-1998 is the mean standardized test score from the previous spring.



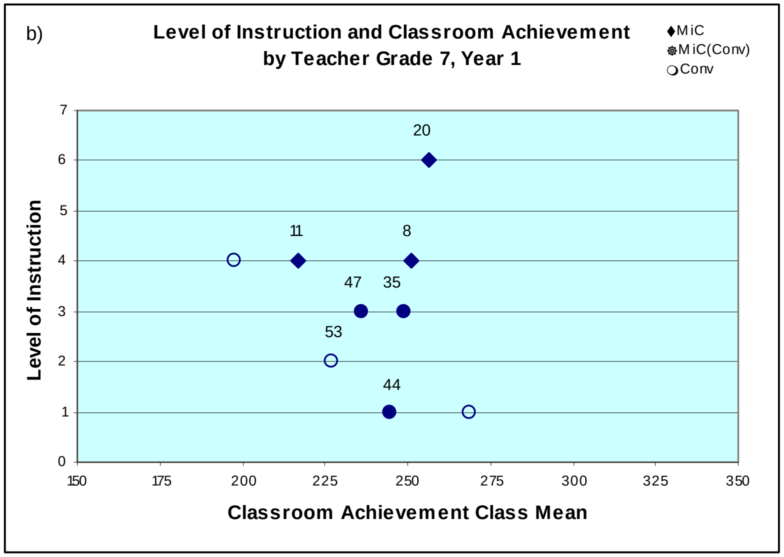


Figure 2-7. Level of instruction for Grade 7 teachers in Year 1 compared with a) prior achievement; b) classroom achievement at the end of the year

Although PA performance was similar, Keeton's students outperformed Donnely's students in CA (see Table 2-7 and Figure 2-7b). In fact, students in one MiC group and one MiC (Conventional) group outperformed Donnely's students, even though the other classes began with lower mean PA scores. Therefore, it is likely that the quality of instruction experienced by students in this cluster had an impact on their CA performance.

Exploration of factors related to OTL_u is also of interest (see Table 2-8 and Figure 2-8). Keetons's students experienced a high level of OTL_u. She presented a comprehensive, integrated curriculum, teaching six units in three content strands. Mathematics was explored in enough detail for students to think about relationships among mathematical ideas or to link procedural and conceptual knowledge. Students were encouraged to make generalizations. Occasionally, she supplemented MiC with activities disconnected from the curriculum, such as a district-mandated reading program for 30 minutes in every class and practice for upcoming district standardized tests. Students in Donnely's classes experienced limited OTL_u. He presented three MiC units in different content strands and supplemented heavily with drill-and-practice materials. Lessons provided little attention to conceptual understanding, and students were not encouraged to make conjectures. Connections among mathematical ideas and connections between mathematics and students' lives were not discussed. In this cluster, it is likely that the high quality of instruction and OTL_u experienced by students in Keeton's classes influenced their gains in performance, whereas the combination of low quality of instruction and OTL_u experienced by Donnely's students influenced their decline in performance.

Table 2-8

Level of Opportunity to Learn with Understanding, Prior Achievement, and Classroom Achievement for Grade 7 Students in 1997-98, by District and by Teacher

District/ Teacher ID	Level of OTLu	PA		CA	
		(N)	Mean	*(N)*	Mean
		MiC			
District 1					
8	4	35	45.54	38	250.92
District 2					
11	2	37	43.97	41	217.02
20	4	41	49.22	45	256.32
		MiC (Conventional)			
District 1					
44	2	47	51.60	50	244.44
District 2					
35	2	35	45.54	50	249.14
47	4	42	39.05	48	235.87
		Conventional			
District 1					
13	2	33	63.42	39	268.58
53	2	36	41.53	45	227.08
District 2					
5	3	24	24.17	22	197.75
24	2	2	25.00	1	236.39

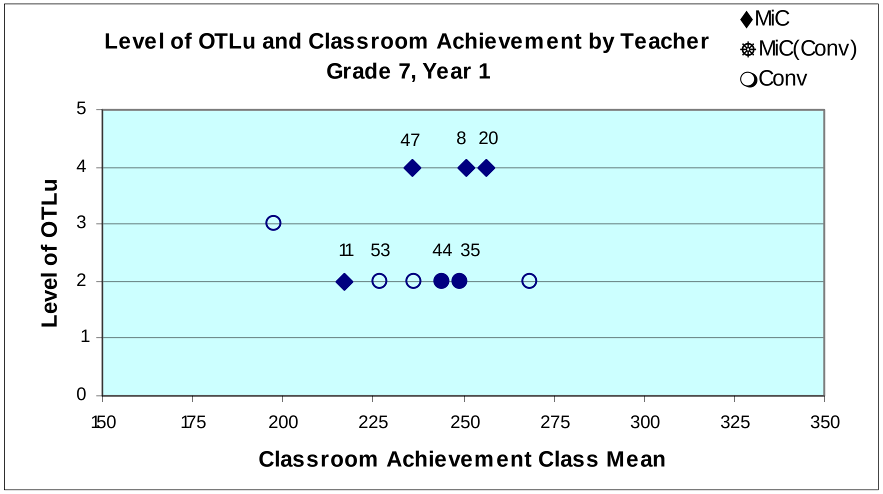


Figure 2-8. Level of opportunity to learn with understanding for Grade 7 teachers in Year 1 compared with classroom achievement at the end of the year

The school capacity these teachers perceived (high and average for Keeton and Donnely, respectively) may also have contributed to the differences in CA scores (see Table 2-9 and Figure 2-9). Keeton taught in the same school as Dillard, MiC Teacher 45. Keeton perceived a high level of school capacity. She felt that principal and teacher visions for mathematics teaching and learning were clearly defined and generally aligned. She felt that she received very strong administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. Keeton felt that she had a high influence in planning and teaching mathematics, an average level of influence over curriculum, and less influence over the content of professional development programs and discipline. She felt that faculty and staff were committed to academic excellence and that teachers supported one another in their efforts to improve instruction. Although professional development opportunities were limited at her school, monetary support was available for other opportunities. Keeton met regularly with other teachers for collaboration. Common planning time was available for mathematics teachers on her teaching team during which they discussed content, particular MiC units, instructional and assessment methods, and program evaluation. As the team leader, Keeton made instructional decisions relative to curricular scope and sequence, integration of innovative curricula, and assessment preparation. District and state standardized tests influenced the content and sequence of content she taught during the school year.

Table 2-9

Level of School Capacity, Prior Achievement, and Classroom Achievement for Grade 7 Students in 1997-98, by District and by Teacher

District/ Teacher ID	Level of School Capacity	PA (N)	PA Mean	CA (N)	CA Mean
MiC					
District 1					
8	4	35	45.54	38	250.92
District 2					
11	4	37	43.97	41	217.02
20	5	41	49.22	45	256.32
MiC (Conventional)					
District 1					
44	3	47	51.60	50	244.44
District 2					
35	4	35	45.54	50	249.14
47	1	42	39.05	48	235.87
Conventional					
District 1					
13	4	33	63.42	39	268.58
53	2	36	41.53	45	227.08
District 2					
5	2	24	24.17	22	197.75
24	1	2	25.00	1	236.39

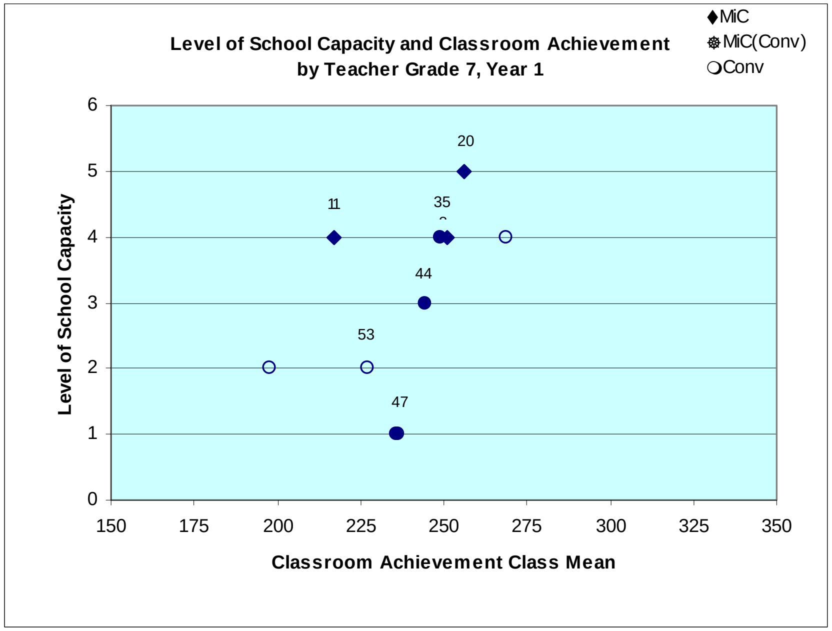


Figure 2-9 Level of school capacity for Grade 7 teachers in Year 1 compared with classroom achievement at the end of the year

Donnely taught at Von Humboldt Middle School, which had an enrollment of 740 students. Classes met four times a week for 50-minute periods. Approximately 60% of his students were White and the largest minority was African American students. Ten to 20% of the students in the school were eligible for government-funded lunch programs. Donnely perceived an average level of school capacity. He felt that the principal and teachers had incompatible visions for mathematics teaching and learning. However, he felt that he received very strong administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. Donnely felt that he had a high influence in planning and teaching mathematics and an average level of influence over educational policies related to curriculum, the content of professional development programs, and discipline. He felt that faculty and staff were committed to academic excellence and that teachers generally supported one another in their efforts to improve instruction. Donnely met regularly with other teachers for collaboration, and common planning time was available. Professional development took the form of informal meetings with other teachers at his school, although substitute time was allotted for external professional development activities. As a new teacher in the district, district and state standardized tests had little influence on Donnely because he was unfamiliar with them.

Students in this cluster were in the same schools for both Grades 6 and 7 so they did not experience changes in structural and cultural conditions in the school. However, differences in the types of collaboration experienced by these teachers were noted. Keeton regularly discussed mathematics curriculum, particular MiC units, instruction, and assessment with other mathematics teachers on her teaching team. This type of collaboration was unavailable to Donnely. Thus, it is likely that CA performance for students in this cluster was influenced by the quality of instruction and OTL_u students experienced, and their teachers' perceptions of school capacity.

Teachers 8, 11, 35, 47, and 53. Students in this cluster had mean PA scores around the 40th percentile, and they experienced different levels of instruction: Heath and McFadden, MiC Teachers 8 and 11, respectively; Draski and Teague, MiC (Conventional) Teachers 35 and 47, respectively; and McLaughlin, Conventional Teacher 53; see Table 2-7 and Figure 2-7a).

As instruction unfolded during the school year, Heath, MiC Teacher 8, worked toward developing conceptual understanding. She included additional exercises based on students' need for practice, mini lessons on procedures, and different contexts to clarify and extend student's understanding of particular concepts and procedures. However, she rarely elicited different problem-solving strategies from students, and student explanations were focused on procedures rather than on elaboration of reasoning. McFadden, MiC Teacher 11, attempted to work toward developing conceptual understanding. She modified lessons based on student inquiries, and she pointed out pertinent elements of tasks that enabled students to continue the mathematical work on their own. She shifted students' work from small groups to whole-class instruction in attempt to increase understanding of the mathematics. Student–student conversation was limited and usually consisted of sharing answers, and student explanations were focused on procedures rather than on elaboration of reasoning. McFadden sought evidence of procedural competence through students' written work and sharing answers in class. Feedback was teacher-directed and was directed toward skills and procedures.

Lessons presented by Draski, MiC (Conventional) Teacher 35, frequently contained review of previous lessons, discussion of homework or a warm-up activity. On some occasions, she posed questions that encouraged students to consider connections among mathematical ideas. Inquiry during lessons included limited attention to conceptual understanding. Student explanations were focused on procedures rather than on elaboration of reasoning or solution strategies; and multiple strategies were not elicited. Draski sought evidence of procedural understanding through students' written work. Feedback was teacher-directed and was directed toward skills and procedures. Some students exchanged ideas or provided assistance to their classmates, but contributions to solving problems were not made equally by all students. Teague, MiC (Conventional) Teacher 47, attempted to teach for conceptual understanding through direct teaching, questioning, and having students explain their answers. She tended to add mini-lessons focused on particular procedures or reviews of other lessons based on students' lack of understanding. She maintained that students should be told this information and implied that they should memorize it, rather than investigate the concepts in real world contexts. At times, however, her questions or explanations promoted conceptual understanding. Even when she posed yes/no questions, students had to justify their responses. Thus, at times their explanations often focused on solution strategies and mathematical reasoning. Teague sought procedural understanding as evidence of learning. Feedback was often teacher-directed and was focused on explanations as well as procedures.

Mc Laughlin, Conventional Teacher 53, presented procedures or strategies, and students practiced them in rote ways. Inquiry during lessons was limited to lower order thinking, and lessons did not promote conceptual understanding. When McLaughlin posed questions, she sometimes answered the questions herself, and, although she frequently asked students if they solved problems in different ways, she eventually led the class to use one specific procedure for solving a type of problem. Rules and procedures were emphasized. As a result, students' explanations were focused on procedures, not mathematical reasoning. Most instruction was presented in a didactic format, followed by students completing assignments individually. This type of instruction and the attention given to correct computational procedures limited the type and quality of the feedback offered to students. McLaughlin responded to students' incorrect answers with mini-lessons in which more examples of correct procedures were given and subsequently reinforced through practice. Because few opportunities were available for student–student conversation, feedback was solely teacher-directed.

With respect to CA scores, performance Heath's and Draski's classes were similar in both PA and CA. However, CA performance these classes surpassed the performance Donnely's (MiC (Conventional) Teacher 44) classes. This is interesting because Donnely's students had higher PA scores but experienced a lower quality of instruction during the school year (see Table 2-7 and Figure 2-7b). Also, although the mean PA performance of MiC Teacher McFadden's classes was similar to MiC (Conventional) Teacher Draski's classes, the mean CA performance for McFadden's classes was thirty points less than the mean CA for Draski's classes. Also, Teague's students outperformed McFadden's classes by nearly 20 points even though students in McFadden's classes had higher mean PA scores. In comparison, for students in McLaughlin's classes, little change in performance was evident. This is of interest because these students experienced instruction that was reflective of conventional pedagogy.

Exploration of factors related to OTL_u is also important (see Table 2-8 and Figure 2-8). Students in Heath's classes experienced a high level of OTL_u. Heath, MiC Teacher 8, presented a comprehensive, integrated curriculum with attention to all content areas, teaching two geometry, two algebra, one number, and one statistics unit. She supplemented MiC with tasks or multiple models that emphasized connections among ideas and connections to students' lives. She occasionally added competitions and projects such as constructing mobiles of three-dimensional solids before studying one of the MiC geometry units. Heath worked toward developing conceptual and procedural understanding. Observed student conjectures consisted mainly of making connections between a new problem and problems already seen.

Students in Teague's classes experienced a moderate level of OTL_u. Teague, MiC (Conventional) Teacher 47, taught four units: two number units, one geometry unit, and one algebra unit. She supplemented the curriculum with skill-based practice worksheets and activities such as a district-mandated reading program for 30 minutes in every class and practice for upcoming district standardized tests. Teague attempted to teach for conceptual understanding, although procedural understanding was deemed very important. She frequently worked with students to look for connections among mathematical ideas. Observed student conjectures consisted mainly of making connections between a new problem and problems already seen. Connections between mathematics and students' lives were not discussed in detail.

Students in other classes in this cluster experienced limited OTL_u. McFadden, MiC Teacher 11, and Draski, MiC (Conventional) Teacher 35, taught two MiC units during the first semester, one number unit, and one geometry unit. After withdrawing from the study in the second semester, they taught a conventional curriculum for the balance of the school year. McFadden attempted to teach for understanding, and for Draski, conceptual understanding was a small part of lesson design. McLaughlin, Conventional Teacher 53, generally followed the adopted textbook with few modifications, teaching number theory, operations with fractions, decimals, and integers, and some geometry. She presented vast content as disparate pieces of knowledge heavily laden with vocabulary and prescribed algorithms. She supplemented the textbook with logic puzzles, competitions, projects, and games. Conceptual understanding was a small part of the lesson design; lessons focused on building students' procedural understanding without meaning. Students were generally expected to use the method she presented. Limited changes were made in instruction even when student responses were unreasonable. Observed student conjectures consisted mainly of making connections between a new problem and problems already seen. Connections among mathematical ideas and between mathematics and students' lives were rarely discussed.

In this cluster, it is likely that the combination of comprehensive content taught in some depth and instruction that attempted to promote conceptual understanding seemed to have a positive effect on student performance for students in MiC Teacher Heath's and MiC (Conventional) Teacher Teague's classes. The declining performance in MiC Teacher McFadden's classes and similar performance in MiC (Conventional) Teacher Draski's and Conventional Teacher McLaughlin's classes likely were influenced by breadth in content rather than depth and instruction that attempted to teach for understanding or did not promote conceptual understanding. Thus, when MiC was used throughout the school year in MiC (Conventional) classes, students experienced increased performance. On the other hand, when MiC was abandoned in favor of a conventional curriculum, performance stayed about the same or declined considerably.

With respect to school capacity, Heath, MiC Teacher 8, perceived a moderately high level of school capacity in her school (see Table 2-9 and Figure 2-9). Heath taught at the same school as Weatherspoon, MiC Teacher 33. Heath felt that principal and teacher visions for mathematics teaching and learning were partially aligned because the teachers had influenced the principal's vision of mathematics instruction. She felt that she received very strong administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. Heath felt that she had a high influence in planning and teaching mathematics, an average level of influence over curriculum, and less influence over the content of professional development programs and discipline. She felt that the faculty and staff were committed to academic excellence and that teachers supported one another in their efforts to improve instruction. Professional development opportunities at her school were related to general teaching methods, and monetary support was available for other opportunities. Heath met informally with other teachers on a regular basis to discuss content, instructional and assessment methods, and program evaluation. Heath felt that the district standardized testing program did not influence curriculum and instruction because she emphasized problem solving and written explanations during instruction.

Teague, MiC (Conventional) Teacher 47, had a very low perception of school capacity in her school. Teague taught on the same teaching team and Keeton and Dillard, MiC Teachers 20 and 45. Teague felt that the principal and teachers had incompatible expectations for mathematics teaching and learning. However, she felt that she received very strong administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. Heath felt that she had a high influence in planning and teaching mathematics, but little influence over the content of professional development programs and discipline. She felt that faculty and staff were committed to academic excellence and that teachers supported one another in their efforts to improve instruction. Although professional development opportunities were limited at her school, monetary support was available for other opportunities. Teague met regularly with other teachers for collaboration. Common planning time was available for mathematics teachers on her teaching team during which they discussed content, particular MiC units, instructional and assessment methods, and program evaluation. District and state standardized tests influenced the content Teague taught, as she supplemented MiC with computation and skill practice.

McFadden, MiC Teacher 11, and Draski, MiC (Conventional) Teacher 35, taught at Hirsch Metro Middle School, which had an enrollment of 2190 students. Classes met every other day, alternating two times a week and three times a week, for 127-minute periods. The student population was 98% minority with 96% Hispanic students. More than 50% of the students were eligible for government-funded lunch programs. McFadden and Draski perceived moderately high school capacity. They felt that the principal and teachers had visions of teaching and learning mathematics that were aligned on some ideas, but were incompatible on others. They felt that they received very strong administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. The principal encouraged teachers to observe one another's teaching, provide meaningful feedback to others, and join networks of teachers in the district. McFadden and Draski felt that they had a high influence in planning and teaching mathematics, an average level of influence over educational policies related to curriculum, the content of professional development programs, and discipline. They felt that faculty and staff were committed to academic excellence and that teachers supported one another in their efforts to improve instruction. Many opportunities for professional development were available in their school, and monetary resources were available for other opportunities. These teachers had formal meetings on a regular basis, and generally also met in informal gatherings. Common planning time was not available for mathematics teachers. District and state standardized tests significantly influenced the instruction of these teachers, as large amounts of time were devoted to test preparation. In fact, this became such a serious issue that they withdrew from the study in the second semester. The principal was not convinced that mathematics education reform should be supported wholeheartedly. His interpretations of state standards led to a different approach for reaching expectations of improved student achievement. He stated that 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. He expressed reluctance for his teachers to use MiC and strongly encouraged them to withdraw from participation in the study. Consequently, McFadden and Draski withdrew from the study in the second semester, and a conventional curriculum was used for the remainder of the school year. The teachers agreed to have students complete study assessments in May, and CA scores were computed for them.

McLaughlin, Conventional Teacher 53, perceived a low level of school capacity. McLaughlin taught at Wacker Middle School, which had an enrollment of 650 students. Classes met five times a week for 42-minute periods. Approximately 40% of the student population was minority, with 30% African American students. McLaughlin felt that the principal and teachers had incompatible expectations for mathematics teaching and learning. However, she felt that she received strong administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. On the other hand, she felt that she had a little over curriculum and the content of professional development programs. She felt that high standards were stated for academic work, but faculty and staff were only somewhat committed to academic excellence and that teachers did not support one another in improving instruction. Although professional development opportunities at her school consisted of informal sharing of ideas among teachers, monetary support was available for external opportunities. During informal meetings, teachers discussed content, instruction, and assessment. McLaughlin felt that the district standardized testing program did not influence curriculum and instruction because she emphasized problem solving and written explanations during instruction.

The strong influence of school capacity for McFadden and Draski had a direct and major impact on the curriculum they taught, which is an important aspect of the OTL_u students experienced. For other teachers in this cluster, the school capacity they perceived did not have such an impact. For these teachers, performance was influenced by instruction and OTL_u.

Summary of Year 1 Results

In Year 1, MiC and MiC (Conventional) teachers taught MiC for the first time, as the commercial version had just become available. Their own understanding of new mathematics content and ways it was presented, and their pedagogical content knowledge, all of which are central to effective instruction, were being developed as teachers taught MiC for the first time. This alone may have been enough to affect student performance. However, in the clusters we examined, when teachers attempted to teach or taught mathematics for understanding (Levels 4-6) or when instruction was more reflective of good conventional pedagogy (Level 3), students in MiC classes showed gains in performance or remained at roughly the same level. The same was true with respect to OTL_u. When MiC classes experienced OTL_u at Levels 3 and 4, the combination of comprehensive content taught in some depth and instruction that attempted to promote conceptual understanding had a positive effect on student performance. Generally, the quality of instruction and OTL_u had more influence than school capacity. However, the type of collaboration experienced by teachers made a difference. At Grade 6, when teacher collaboration involved discussion of mathematics curriculum, instruction, and assessment, classroom achievement increased. For classes using conventional curricula, student performance varied considerably. The performance of some Grade 5 students who experienced a high quality of instruction and moderate OTL_u declined. However, the performance of other conventional groups at different grade levels remained at about the same level after experiencing low quality of instruction and OTL_u. At Grade 6, because students made a transition from elementary to middle schools, they experienced changes in structural and cultural conditions of their schools that were not an issue for students at other grade levels. In the cluster of students with comparable prior achievement at Grade 6, classroom achievement of MiC students remained at a high level, whereas performance students studying a conventional curriculum declined substantially. Although these groups experienced major differences in instruction and OTLu, teacher collaboration, a dimension of school capacity, may also have had an important effect. The MiC teacher in this cluster had multiple opportunities to collaborate with mathematics teachers on his teaching team regarding mathematics curriculum, instruction, and assessment. These opportunities were not available to the conventional teacher in this cluster. At Grade 7 in the examined clusters, student performance was positively affected when MiC (Conventional) teachers taught MiC throughout the school year, even when instruction was reflective of good conventional pedagogy. On the other hand, student performance stayed the same or declined considerably when MiC (Conventional) teachers substantially supplemented MiC with skill practice or when they stopped teaching MiC and used a conventional textbook.

Comparable Groups in Year 2

In Year 2, some teachers in the clusters we examine taught MiC for the first time during the entire school year, and they faced the same challenges that MiC and MiC (Conventional) teachers struggled with during the first study year. In addition, some students were new to the study, while other students participated in the study both years. Classroom achievement from Year 1 was used as a measure of prior achievement in Year 2. In most cases, student groups were larger in determining the classroom achievement mean score for Year 2 than groups used in determining the prior achievement mean score. One factor that influenced sample size was the completion of both study assessments in Year 1. (Recall that results from both assessments were used in defining the proficiency scale that describes classroom achievement.)

Grade 6

Table 2-10 and Figure 2-10a show the mean PA score (CA from the prior spring) and the level of instruction assigned to each sixth-grade teacher. Two clusters are described. The first cluster had mean PA scores around 270 and experienced different levels of instruction during the school year, classes taught by Gollen, MiC Teacher 67, and Friedman, Conventional Teacher 82. The second cluster of students had mean PA scores between 240 and 270 and experienced high levels of instruction during the school year, classes taught by Weatherspoon, Gollen, and Redling, MiC Teachers 33, 67, and 69, respectively.

Teachers 67 and 82. Students in this cluster had mean PA scores around 270, but they experienced different levels of instruction. Gollen, MiC Teacher 67, taught MiC for the first time this year. Although instruction was of high quality (Level 5), Gollen had difficulty discerning the alignment of lessons with the unit goals and leading classroom discussions. Student-student interactions rarely went beyond sharing answers or procedures and were not orchestrated to promote sense making. Feedback was clear and mathematically sound and was directed toward the mathematics. Friedman, Conventional Teacher 82, presented particular procedures, and students practiced them in a rote fashion. Inquiry during lessons was limited to lower order thinking, and little attention was given to conceptual understanding. In response to Friedman’s questions, students stated answers or provided steps in procedures, and they were not expected to elaborate on their reasoning. Procedural understanding was sought as evidence of student learning, and greater attention was given to student homework and classwork for instructional decision-making. Feedback was indirectly responsive to student needs in that it involved “more of the same,” such as additional instruction and practice sets. Feedback was limited to checking for accurate answers and seldom addressed student misconceptions. Student–student conversation was not encouraged.

Table 2-10
Level of Instruction, Prior Achievement, and Classroom Achievement for Grade 6 Students in 1998-99, by District and by Teacher

District/ Teacher ID	Level of Instruction	PA		CA	
		(N)	Mean	*(N)*	Mean
MiC					
District 1					
33	5	27	259.56	43	267.54
67	5	15	270.01	24	266.13
78	4	3	302.65	22	311.99
District 2					
69	5	8	242.67	41	281.62
MiC (Conventional)					
District 1					
38	1	13	285.87	44	292.00
73	1	12	306.27	29	209.77
District 2					
52	3	1	298.51	41	204.00
83	2	0		40	201.19
85	3	0		49	282.96
Conventional					
District 1					
79	1	5	230.29	18	212.70
District 2					
21	2	0		20	265.27
82	2	18	271.78	30	294.20

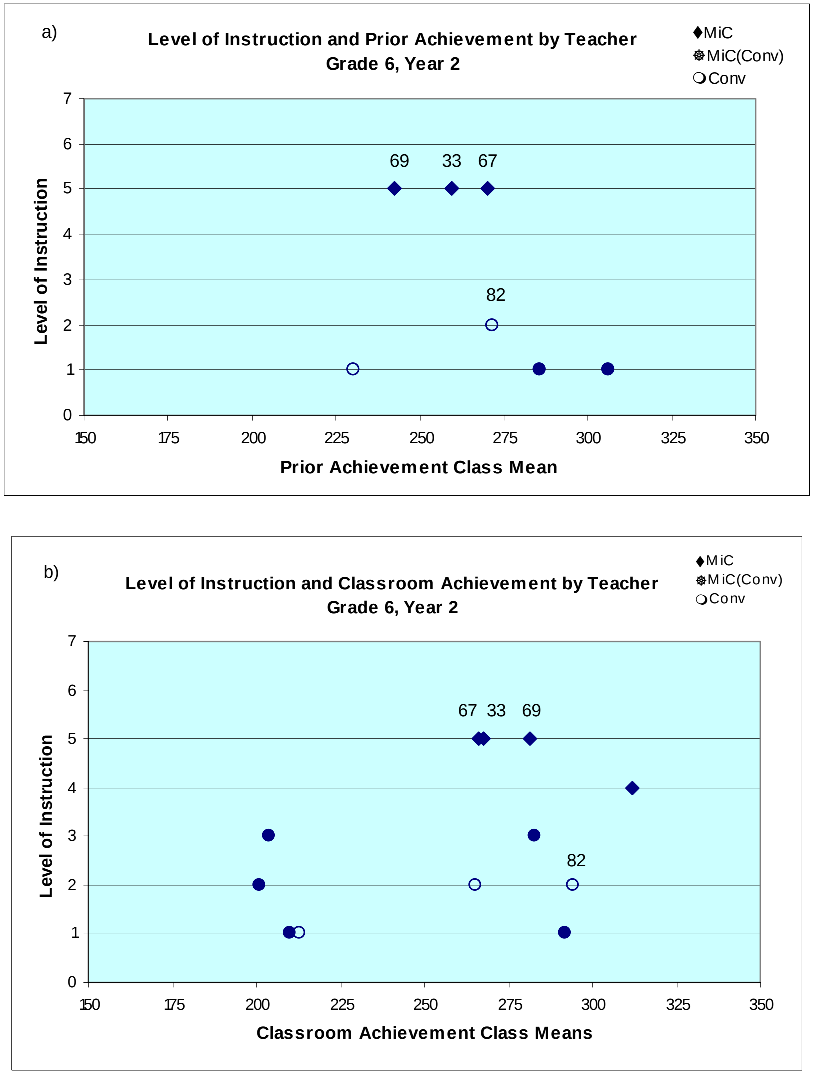


Figure 2-10. Level of instruction for Grade 6 teachers in Year 2 compared with a) prior achievement; b) classroom achievement at the end of the year

With respect to CA scores (see Table 2-10 and Figure 2-10b), students in Gollen's class decreased slightly in performance in comparison to the 22-point gain in mean scores for students in Friedman's classes. This result is surprising because of the differences in instruction noted for these teachers. Also noteworthy is that students in Friedman's classes were in classes with Gant, Conventional Teacher 43, during the previous year in which instruction frequently was characterized as emphasizing review of procedures. The higher quality of instruction in the second year for these students may have had an impact on their performance.

Students in Gollen's classes experienced a moderate level of OTL_u (see Table 2-11 and Figure 2-11). Gollen taught a comprehensive, integrated curriculum, teaching five units in three content strands. During instruction, she worked toward conceptual understanding. Observed student conjectures consisted mainly of making connections between a new problem and problems previously seen. Connections among mathematical ideas and connections between mathematics and students' lives were not discussed in detail. Friedman's students experienced limited OTL_u. Curricular content spanned a vast content plane with little or no depth, and the textbook was followed with few modifications. The content was presented as disparate pieces of knowledge heavily laden with vocabulary and prescribed algorithms. Few questions fostered conceptual understanding, and making conjectures was not encouraged. Connections between mathematics and students' lives were not discussed in lessons.

Table 2-11

Level of Opportunity to Learn with Understanding, Prior Achievement, and Classroom Achievement for Grade 6 Students in 1998-99, by District and by Teacher

District/ Teacher ID	Level of OTLu	PA		CA	
		(N)	Mean	*(N)*	Mean
		MiC			
District 1					
33	4	27	259.56	43	267.54
67	3	15	270.01	24	266.13
78	3	3	302.65	22	311.99
District 2					
69	3	8	242.67	41	281.62
		MiC (Conventional)			
District 1					
38	1	13	285.87	44	292.00
73	2	12	306.27	29	209.77
District 2					
52	3	1	298.51	41	204.00
83	2	0		40	201.19
85	3	0		49	282.96
		Conventional			
District 1					
79	2	5	230.29	18	212.70
District 2					
21	2	0		20	265.27
82	2	18	271.78	30	294.20

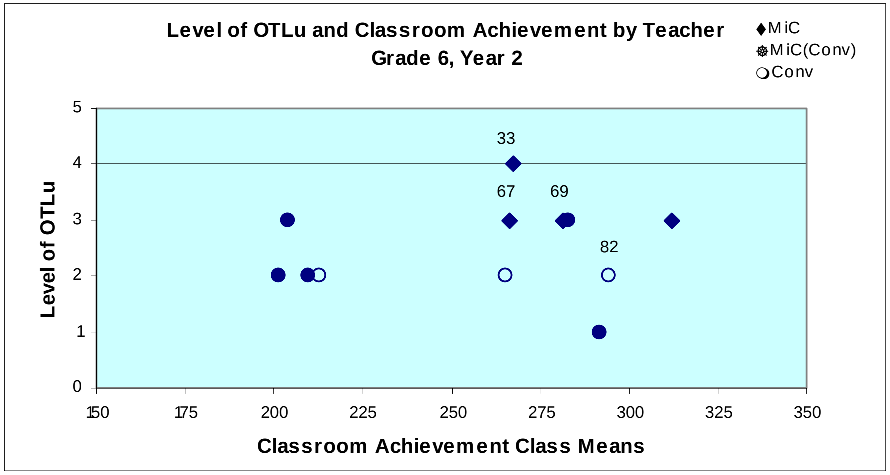


Figure 2-11. Level of opportunity to learn with understanding for Grade 6 teachers in Year 2 compared with classroom achievement at the end of the year

What is interesting here is that even though students in Gollen's class experienced higher OTL_u, their performance decreased slightly in comparison to the substantial gain for students in Friedman's classes. OTL_u seemed to have little influence on student achievement. Investigation of school capacity is warranted. Table 2-12 and Figure 2-12 show the levels of school capacity for sixth-grade teachers.

Gollen taught at Addams Middle School with an enrollment of 500 students in Grades 6–10. Classes met five times a week for 43-minute periods. Because it was a magnet school for the performing arts, sixth-grade students were drawn from the entire district to attend this middle school. In sixth-grade approximately 55% of the 115 students were minority with 15% African American students, 10% Hispanic students, and 25% Multiracial students. Twenty to 30% of the students were eligible for government-funded lunch programs. Gollen perceived moderately high school capacity. She felt that the principal and teachers had visions of teaching and learning mathematics that were aligned on some ideas expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. She felt that she had an average level of influence in planning and teaching mathematics and curriculum, and a high level of influence over the content of professional development programs and discipline. She felt that faculty and staff were committed to academic excellence and that teachers supported one another in their efforts to improve instruction. Although formal meetings were held for mathematics teachers, teachers met informally on a limited basis to discuss mathematics curriculum and instruction. Substitute time was allotted for external professional development. Gollen was the only sixth-grade MiC teacher in her school, and she did not have the opportunity to collaborate with other teachers who were implementing MiC.

She believed that MiC was well aligned with the district and state standardized testing programs, and these tests had little influence over curriculum and instruction.

Table 2-12
Level of School Capacity, Prior Achievement, and Classroom Achievement for Grade 6 Students in 1998-99, by District and by Teacher

District/ Teacher ID	Level of School Capacity	PA (N)	PA Mean	CA (N)	CA Mean
MiC					
District 1					
33	2	27	259.56	43	267.54
67	4	15	270.01	24	266.13
78		3	302.65	22	311.99
District 2					
69	5	8	242.67	41	281.62
MiC (Conventional)					
District 1					
38		13	285.87	44	292.00
73	3	12	306.27	29	209.77
District 2					
52	5	1	298.51	41	204.00
83	5	0		40	201.19
85	5	0		49	282.96
Conventional					
District 1					
79	3	5	230.29	18	212.70
District 2					
21	2	0		20	265.27
82	3	18	271.78	30	294.20

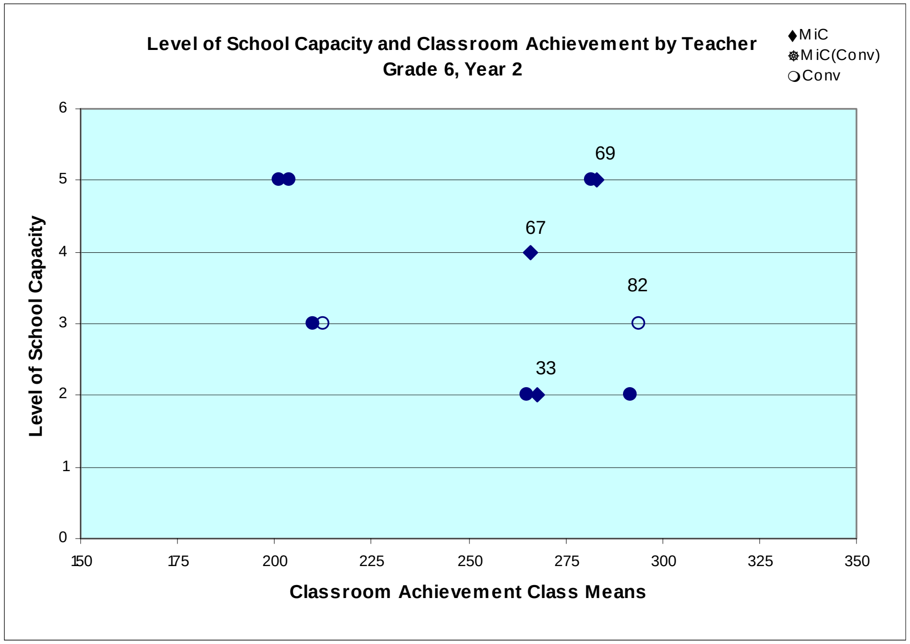


Figure 2-12. Level of school capacity for Grade 6 teachers in Year 2 compared with classroom achievement at the end of the year

Friedman, Conventional Teacher 82, taught at the same school for Grades K-8 as Gant, Conventional Teacher 43. Friedman perceived an average level of school capacity. She felt that the principal and teachers had incompatible visions for mathematics teaching and learning. However, she felt that she received very strong administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. Her principal encouraged teachers to observe one another and provide meaningful feedback on instructional practice. Friedman felt that she had a high level of influence in planning and teaching mathematics, and limited influence over educational policies related to curriculum, content of professional development programs, and discipline. She felt that faculty and staff were committed to academic excellence and that teachers supported one another in their efforts to improve instruction. Formal meetings for teachers were held on a regular basis throughout the school year, and teachers met informally on occasion in lieu of common planning time. Professional development at the school took the form of informal collaborations among teachers, although the principal allotted substitute time to support teachers in obtaining external professional development. District and state standardized tests had little influence because she was unfamiliar with them.

Two factors related to school capacity may have influenced CA results. First, Gollen was the only sixth-grade MiC teacher in her school, and she reported infrequent meetings with other mathematics teachers and few opportunities to collaborate with other MiC teachers in the district. Second, her students transitioned from various elementary schools across the district to this magnet middle school in which the structural and cultural conditions were different from what they experienced in elementary grades. In contrast, students in Friedman's classes did not encounter changes in school context, as they were in the same school for both Grades 5 and 6 in a K-Grade 8 community school. As noted earlier, Friedman's students experienced instruction focused on procedures and limited OTL_u throughout the school year. These students had also been in Gant's classes the previous year during which they experienced similar OTL_u and instruction. Thus, it is surprising that with these levels of instruction and OTL_u over a two-year period Friedman's students demonstrated substantial gains in performance. The fact that they continued their education in the same school likely influenced their performance in significant ways.

MiC Teachers 67, 33, and 69. The cluster of students with Weatherspoon, Gollen, and Redling, MiC Teachers 33, 67, and 69, experienced instruction at Level 5, but they varied in their PA performance (mean scores between 240 and 270; see Table 2-10 and Figure 2-10a). Gollen and Redling taught MiC for the first time this year. As noted earlier, Gollen taught for conceptual understanding, although she had difficulty orchestrating classroom discussion. Weatherspoon and Redling worked with students to develop conceptual understanding, and they shared the mathematical work with their students. They emphasized connections among mathematical ideas. Of interest here is the substantial gain in CA mean score for students in Redling's classes—nearly a 29-point gain, outperforming the other classes in this cluster who had higher mean PA scores by roughly 14 and 15 points, respectively (see Table 2-10 and Figure 2-10b). Because students in this cluster experienced a high quality of instruction, further investigation into OTL_u and school capacity is warranted.

This cluster of students experienced a high or moderate OTL_u (see Table 2-11 and Figure 2-11). Gollen taught a comprehensive, integrated curriculum, teaching five units in three content strands. She worked toward conceptual understanding. Observed student conjectures consisted mainly of making connections between a new problem and problems previously seen. Connections among mathematical ideas and connections between mathematics and students' lives were not discussed in detail. Weatherspoon presented a comprehensive, integrated curriculum, teaching four units in three content areas. She supplemented the units with activities that emphasized connections among mathematical ideas and connections to students' lives. Lessons also promoted procedural understanding. Redling taught mathematics in depth but restricted content primarily to three of the four sixth-grade algebra units. She supplemented MiC with preparation for the district standardized tests. At times, observed student conjectures consisted of investigating the truthfulness of particular statements, while at other times, conjectures stemmed from making connections between a new problem and problems previously seen. Connections among mathematical ideas and connections between mathematics and students' lives were not discussed in detail.

Brief descriptions of OTL_u for these students do not help to understand the factors that contributed to the substantial gain in performance for students in Redling's classes. An examination of school capacity is warranted. Table 2-12 and Figure 2-12 show information about school capacity for sixth-grade teachers.

Weatherspoon, Teacher 33, taught at the same school as Heath, MiC Teacher 8. The school's population was approximately 50% minority, and more than 50% of the students were eligible for government-funded lunch programs. Classes met five days a week for 90-minute periods. Weatherspoon perceived an average level of school capacity in the previous school year, but this year she perceived low school capacity. She was the only MiC teacher at her grade level in the school. She felt that principal and teacher visions for mathematics teaching and learning were clearly defined and generally aligned. She felt that she received strong administrative support in terms of clearly communicated expectations. However, limited support was available with respect to selecting instructional materials, changes in instructional practice, and changes in policy. The principal did not encourage teachers to observe one another's teaching, provide meaningful feedback to others, or join networks of teachers in the district. Weatherspoon felt that she had a high level of influence in planning and teaching mathematics, and an average level of influence over educational policies related to curriculum, content of professional development programs, and discipline. She felt that faculty and staff were committed to academic excellence and that teachers supported one another in their efforts to improve instruction. Professional development at the school included sessions on general teaching methods. The principal did not observe her teaching, and resources were not available for teachers to attend external professional development. Formal meetings for mathematics teachers were held infrequently, and teachers met informally on occasion in lieu of common planning time. Weatherspoon believed that the district and state standardized testing programs had little influence in her planning because she emphasized problem solving and written explanations of solutions during instruction. Also noteworthy is that class periods during administration of study assessments for Weatherspoon's students were interrupted by multiple bomb scares (in the aftermath of the school shootings in Colorado that spring). This situation likely affected their performance.

As noted in the previous section, Gollen, Teacher 67, perceived a moderately high level of school capacity. She taught in a magnet school for the performing arts; sixth-grade students were drawn from the entire district to attend this middle school. She was the only sixth-grade MiC teacher in her school, and she did not have the opportunity to collaborate with other teachers who were implementing MiC. Although formal meetings were held for mathematics teachers at her school, teachers met informally on a limited basis to discuss mathematics curriculum and instruction. Substitute time was allotted for external professional development.

Redling, Teacher 69, taught at the same school as Keeton and Dillard, MiC Teachers 20 and 45. Redling perceived high school capacity. She felt that the principal and teachers had visions of teaching and learning mathematics that were aligned on some ideas, but were incompatible on others. She felt that she received very strong administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. Redling felt that she had a high level of influence over educational policies related to curriculum, content of professional development programs, and discipline, and a high level of influence in planning and teaching mathematics. She felt that faculty and staff were committed to academic excellence and that teachers supported one another in their efforts to improve instruction. Redling reported that numerous professional development opportunities were available in the school, and the principal allotted substitute time for external professional development. During formal meetings, mathematics teachers on her teaching team discussed content, instructional and assessment methods, and program evaluation. Teachers also met informally to discuss a variety of topics. District and state standardized tests significantly influenced instruction, as large amounts of time were devoted to test preparation.

Three themes emerged from the descriptions of these teachers' perceptions of school capacity. First, students in classes with these teachers made transitions from elementary to middle schools in which the structural and cultural conditions were different from what they experienced in elementary grades. However, despite changes in these conditions, the performance of Redling's students increased significantly. The transition from elementary to middle schools may have influenced students in some schools and not others. Second, Weatherspoon and Gollen were the only sixth-grade MiC teachers in their schools, and both reported infrequent meetings with other mathematics teachers. This contrasts with Redling's participation in formal meetings with mathematics teachers on her teaching team. Although this was also the first year Redling taught MiC, other teachers on the team had taught the sixth-grade units. Their discussions not only included content but instructional and assessment methods. This type of collaboration likely supported the quality of instruction and OTL_u her students experienced and mitigated the impact of students' transition from elementary to middle schools. This finding is consistent with Year 1 results for Dillard, MiC Teacher 45, and a teacher on her teaching team. Third, another important consideration is the undesirable condition in which Weatherspoon's students completed study assessments. Class periods during administration of study assessments were interrupted with unavoidable and troubling external conditions, and performance was likely affected by these adverse conditions.

Grade 7

Table 2-13 and Figure 2-13a show the mean PA score (CA from the prior spring) and the level of instruction assigned to each seventh-grade teacher. One cluster is described here: the cluster with mean PA scores between 195 and 215 and who experienced similar instruction during the school year, classes taught by Muldoon and Bartlett, MiC (Conventional) Teachers 74 and 88. Both teachers presented particular strategies, and students practiced them in a rote fashion. Inquiry during class included limited attention to conceptual understanding. When questions were posed, students stated answers or provided steps in procedures, and they were not expected to elaborate on their reasoning. Student–student conversation was not encouraged. Muldoon sought procedural understanding as evidence of student learning. Feedback was teacher-directed and was limited to checking for accurate answers. Feedback seldom addressed student misconceptions. For Bartlett, evidence of student learning included student explanations in addition to procedural understanding and answers. Although feedback was directed toward skills and procedures, it rarely addressed the meaning of procedures and was often directed toward the format of the answer rather than clarifying explanations or developing student understanding.

Table 2-13

Level of Instruction, Prior Achievement, and Classroom Achievement for Grade 7 Students in 1998-99, by District and by Teacher

District/ Teacher ID	Level of Instruction	PA		CA	
		(N)	Mean	*(N)*	Mean
MiC					
District 1					
8	4	19	223.81	26	237.58
District 2					
45	5	10	267.28	31	263.57
62	5	0		58	237.65
70	4	22	201.04	72	218.64
MiC (Conventional)					
District 1					
74	2	12	196.62	50	257.69
88	2	19	211.42	45	259.91
District 2					
80	2	0		49	234.21
Conventional					
District 1					
13	1	3	257.28	47	265.94
76	2	12	244.81	22	234.12
77	2	9	209.22	16	249.79
District 2					
5	3	12	188.15	16	195.87

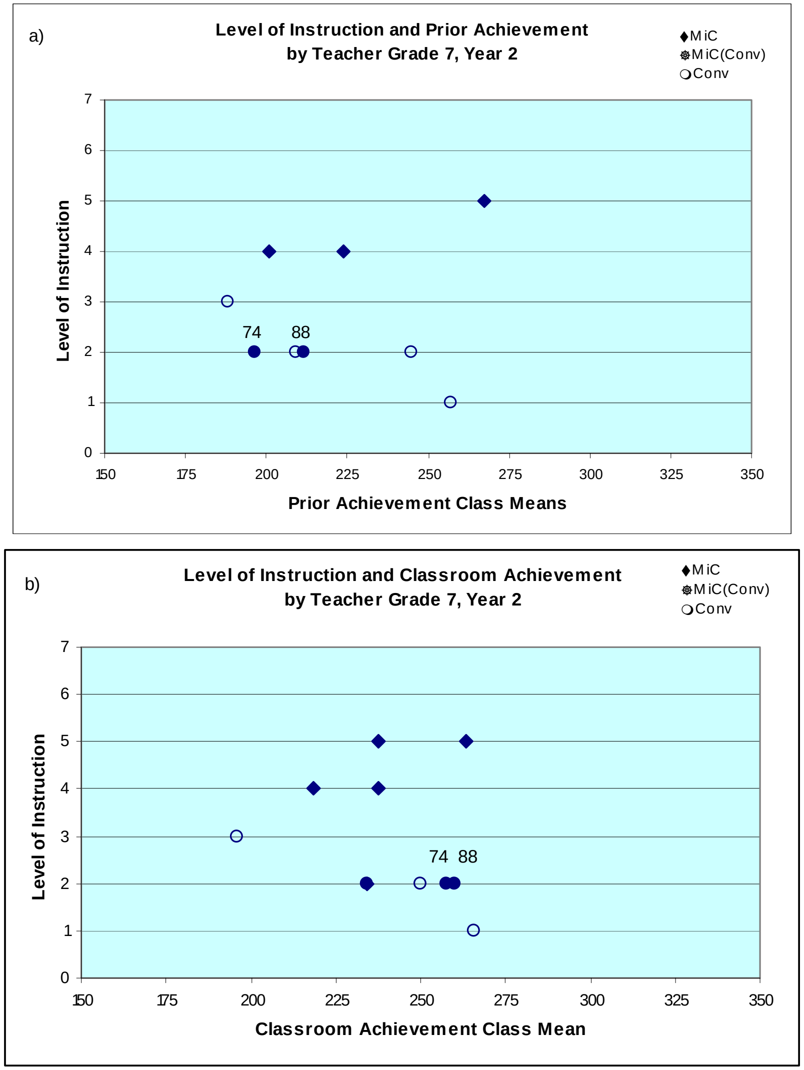


Figure 2-13. Level of instruction for Grade 7 teachers in Year 2 compared with a) prior achievement; b) classroom achievement at the end of the year

With respect to CA, student performance in classes of both teachers increased significantly—a 61-point increase in mean scores for Muldoon's students and a 48-point gain for Bartlett's students (see Table 2-13 and Figure 2-13b). During the school year, these students had experienced instruction that was more reflective of conventional pedagogy. However, it is important to note that in Grade 6, students in this cluster experienced a low quality of instruction. The change to a higher quality of instruction in Grade 7 was likely pivotal in the substantial improvement in CA.

With respect to OTL_u (see Table 2-14 and Figure 2-14), Muldoon, Teacher 74, taught three units in two content strands. She supplemented MiC with practice on computation. However, she also attended to such details as teaching sections of fifth- and sixth-grade MiC units that contained prerequisite skills for the seventh-grade units. Bartlett, Teacher 88, taught five units in four content strands. He also supplemented with MiC computation practice, but to a lesser extent than Muldoon. Lessons presented by both teachers provided limited attention to conceptual understanding. Making conjectures was generally not encouraged, but when observed, student conjectures consisted mainly of making connections between a new problem and problems previously seen. Connections among mathematical ideas and between mathematics and students' lives were not discussed in detail. In Muldoon's classes, instruction was more reflective of conventional pedagogy, but the seventh-grade units were supplemented with portions of units from previous grade levels. This combination of events may have influenced the substantial gains for her students. Bartlett's classes studied more MiC units, and although instruction was similar to Muldoon, modifications of the curriculum occurred to a much lesser extent. Although the quality of instruction was more reflective of conventional pedagogy, the number of units taught likely influenced the substantial gains in performance for Bartlett's classes.

Table 2-14

Level of Opportunity to Learn with Understanding, Prior Achievement, and Classroom Achievement for Grade 7 Students in 1998-99, by District and by Teacher

District/ Teacher ID	Level of OTLu	PA *(N)*	PA Mean	CA *(N)*	CA Mean
MiC					
District 1					
8	3	19	223.81	26	237.58
District 2					
45	3	10	267.28	31	263.57
62	4	0		58	237.65
70	3	22	201.04	72	218.64
MiC (Conventional)					
District 1					
74	2	12	196.62	50	257.69
88	3	19	211.42	45	259.91
District 2					
80	2	0		49	234.21
Conventional					
District 1					
13	2	3	257.28	47	265.94
76	2	12	244.81	22	234.12
77	3	9	209.22	16	249.79
District 2					
5	3	12	188.15	16	195.87

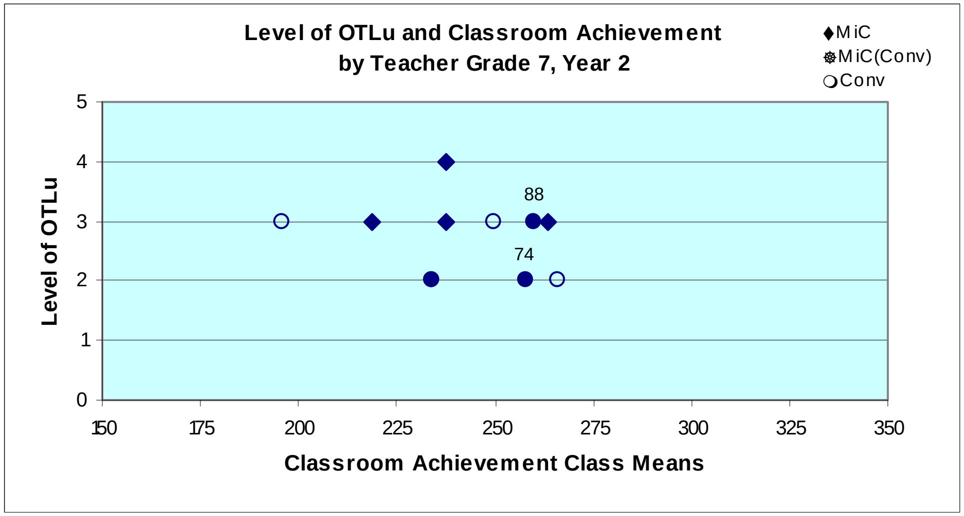


Figure 2-14. Level of opportunity to learn with understanding for Grade 7 teachers in Year 2 compared with classroom achievement at the end of the year

Table 2-15 and Figure 2-15 show data on school capacity for seventh-grade teachers. Muldoon and Bartlett, Teachers 74 and 88, taught at the same school as Donnely, MiC (Conventional) Teacher 44. Classes met five times a week for 63-minute periods. Muldoon and Bartlett perceived low and average levels of school capacity, respectively. Both teachers felt that the principal and teachers had visions of teaching and learning mathematics that were aligned on some ideas, but were incompatible on others. They felt that they received strong administrative support in terms of clearly communicated expectations, support for selecting instructional materials, and changes in instructional practice, and limited support for implementing changes in policy. Muldoon and Bartlett felt that they had limited and moderate influence, respectively, over educational decisions such as curriculum, discipline, and the content of professional development programs. Both teachers felt that they had limited influence over textbook selection and the skills to be taught at their grade level, but moderate to high influence over mathematics planning and teaching. Both felt that faculty and staff was committed to academic excellence, although they varied on the level of support they perceived teachers provided one another in their efforts to improve instruction. Some professional development activities were held at the school, and the principal provided substitute time for external professional development. Formal meetings with other mathematics teachers in the school were held infrequently, and teachers varied in the amount of times they met informally with other teachers to discuss mathematics curriculum, instruction, and assessment. Common planning time was available for teachers. District and state standardized tests had little influence because these teachers believed that their curriculum was well aligned with the testing programs.

The perceptions of school capacity reported by these teachers are similar in many ways and do not provide insight into the performance of their students. The substantial gains in performance for students in both teachers' classes were likely due to the combination of the instruction and OTL_u they experienced.

Table 2-15
Level of School Capacity, Prior Achievement, and Classroom Achievement for Grade 7 Students in 1998-99, by District and by Teacher

District/ Teacher ID	Level of School Capacity	PA (N)	PA Mean	CA (N)	CA Mean
MiC					
District 1					
8	2	19	223.81	26	237.58
District 2					
45	5	10	267.28	31	263.57
62	5	0		58	237.65
70	4	22	201.04	72	218.64
MiC (Conventional)					
District 1					
74	2	12	196.62	50	257.69
88	3	19	211.42	45	259.91
District 2					
80	2	0		49	234.21
Conventional					
District 1					
13	4	3	257.28	47	265.94
76	2	12	244.81	22	234.12
77	3	9	209.22	16	249.79
District 2					
5	3	12	188.15	16	195.87

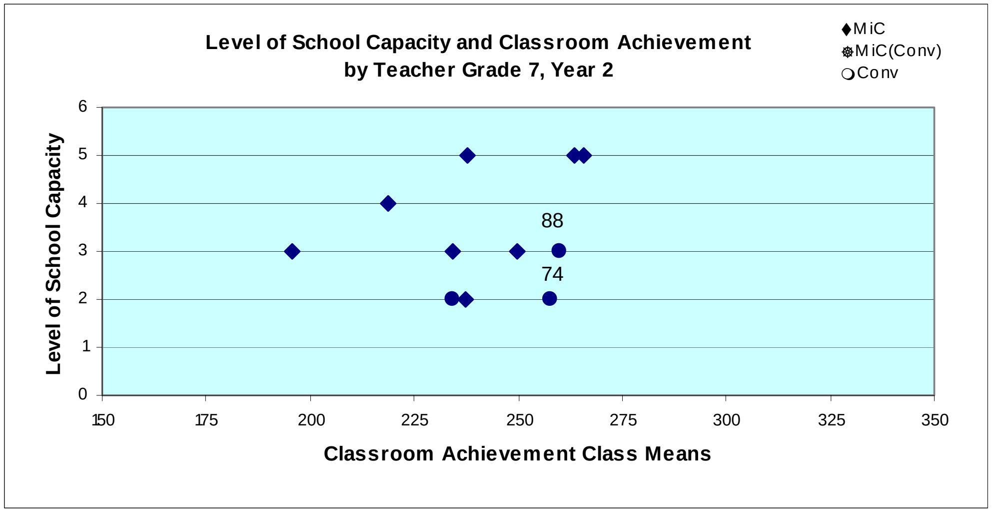


Figure 2-15. *Level of school capacity for Grade 7 teachers in Year 2 compared with classroom achievement at the end of the year*

Grade 8

Table 2-16 and Figure 2-16a show the mean PA score (CA from the prior spring) and the level of instruction assigned to each eighth-grade teacher. One cluster is described here: the cluster with mean PA scores between 250 and 275 and who experienced different levels of instruction during the school year, classes taught by Keeton, MiC Teacher 20, and Wolfe, Conventional Teacher 86. Keeton worked toward developing conceptual understanding. She shared the mathematical work with her students, posed questions that encouraged students to explain their thinking, and helped students make connections among mathematical ideas. She pressed students to effectively communicate their thinking processes, and she modified lessons in response to students' methods. She used students' explanations as evidence of their learning, and interactions were used to promote making sense of the mathematics. A blend of clear, mathematically sound feedback addressed skills, procedures, and concepts. Student–student conversation was frequent. In contrast, Wolfe presented particular procedures, and students practiced them in a rote fashion. Inquiry during lessons was limited to lower order thinking, and little attention was given to conceptual understanding. Students stated answers or provided steps in procedures when they answered Wolfe's questions, and they were not expected to elaborate on their reasoning. Procedural understanding was sought as evidence of student learning, and greater attention was given to student homework and classwork for instructional decision-making. Feedback was indirectly responsive to student needs in that it involved "more of the same" instruction or practice sets.

Table 2-16

Level of Instruction, Prior Achievement, and Classroom Achievement for Grade 8 Students in 1998-99, by District and by Teacher

District/ Teacher ID	Level of Instruction	PA *(N)*	PA Mean	CA *(N)*	CA Mean
		MiC			
District 1					
87	4	28	245.83	53	234.80
District 2					
20	5	22	255.02	40	272.15
47	4	18	231.43	52	239.67
		MiC (Conventional)			
District 1					
75	2	11	237.31	33	240.82
District 2					
(None)					
		Conventional			
District 1					
60	1	9	205.58	22	214.66
86	1	34	272.75	34	255.33
District 2					
5	3	12	188.15	24	189.70
24	2	1	203.63	5	192.22

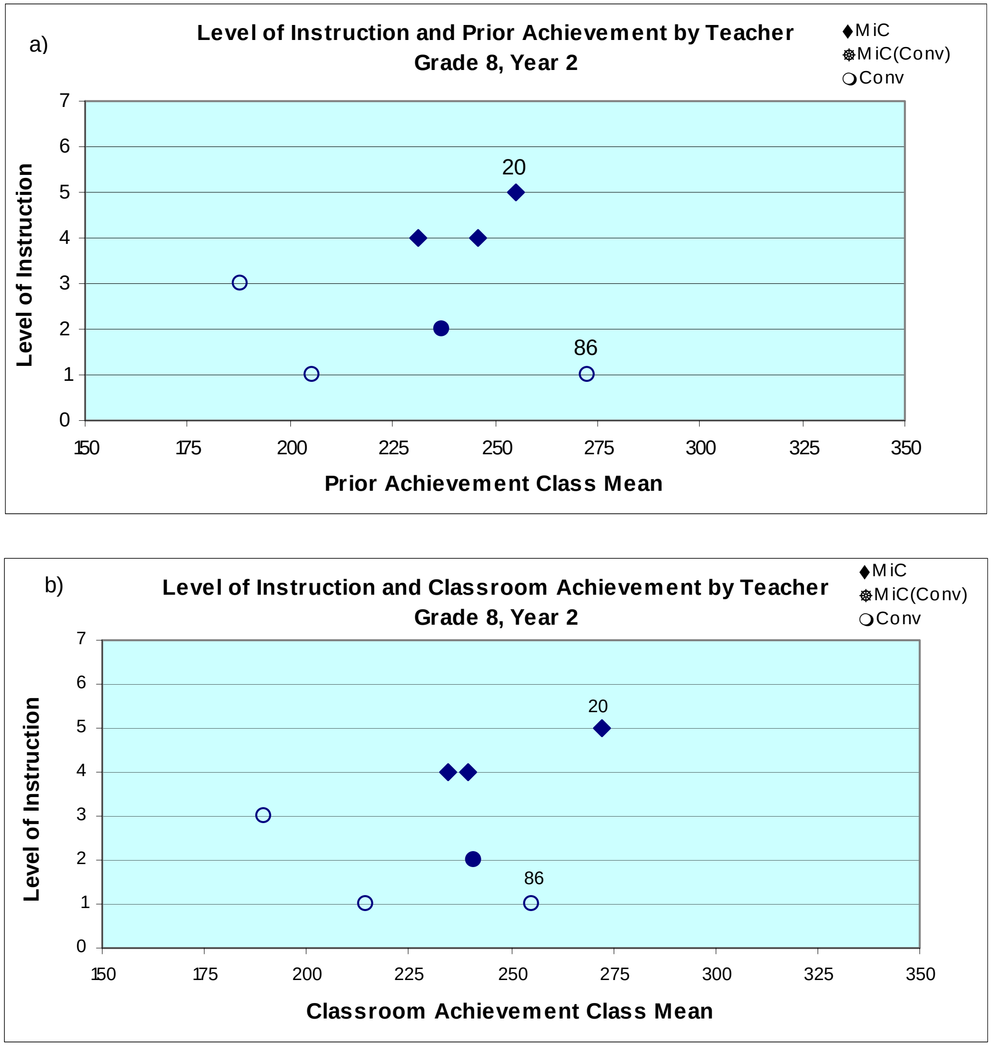


Figure 2-16. Level of instruction for Grade 8 teachers in Year 2 compared with a) prior achievement; b) classroom achievement at the end of the year

With respect to CA scores (see Table 2-16 and Figure 2-16b); Keeton's (Teacher 20's) students gained 17 points, whereas Wolfe's (Teacher 86's) students declined the same amount. Note that Keeton "looped" with her classes. That is, she moved with her students from one grade level to the next. She taught the same core of students in both Grades 7 and 8. These students experienced instruction of high quality for two years. Wolfe's students were in classes taught by another study teacher during the previous year, and they had experienced similar instruction that year. That is, students in Wolfe's classes experienced a low quality of instruction for two years. Although it is likely that the instruction students experienced in these classes made a difference in student performance, it is also important to examine differences in OTL_u for these teachers.

Students in Keeton's classes experienced high OTL_u (see Table 2-17 and Figure 2-17). Keeton presented a comprehensive, integrated curriculum, teaching five units in four content strands. Mathematics was explored in enough detail for students to think about relationships among mathematical ideas or to link procedural and conceptual knowledge. Students were encouraged to make generalizations. Occasionally, Keeton supplemented MiC with activities disconnected from the curriculum such as practice for district standardized tests. Students in Wolfe's classes experienced limited OTL_u. Curricular content spanned a vast content plane with little or no depth, following the textbook with few modifications. The content was presented as disparate pieces of knowledge heavily laden with vocabulary and prescribed algorithms. Few questions fostered conceptual understanding, and making conjectures was not encouraged. Connections between mathematics and students' lives were not discussed in lessons.

Table 2-17

Level of Opportunity to Learn with Understanding, Prior Achievement, and Classroom Achievement for Grade 8 Students in 1998-99, by District and by Teacher

District/ Teacher ID	Level of OTLu	PA (N)	PA Mean	CA (N)	CA Mean
		MiC			
District 1					
87	3	28	245.83	53	234.80
District 2					
20	4	22	255.02	40	272.15
47	3	18	231.43	52	239.67
		MiC (Conventional)			
District 1					
75	3	11	237.31	33	240.82
District 2					
(None)					
		Conventional			
District 1					
60	2	9	205.58	22	214.66
86	2	34	272.75	34	255.33
District 2					
5	3	12	188.15	24	189.70
24	2	1	203.63	5	192.22

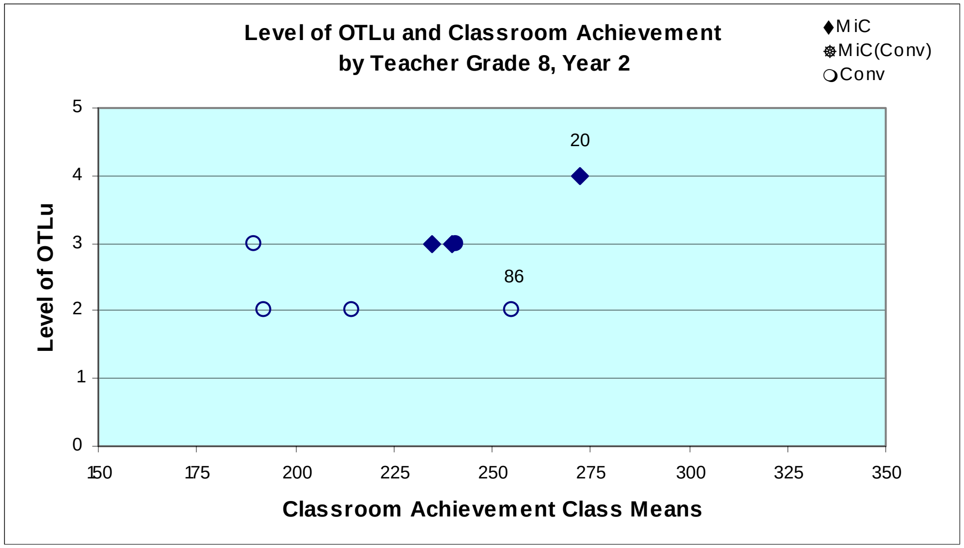


Figure 2-17. Level of opportunity to learn with understanding for Grade 8 teachers in Year 2 compared with classroom achievement at the end of the year

In both Grades 7 and 8, Keeton's students experienced a high quality of instruction and OTL_u, which likely contributed to their increased performance. In contrast, students in Wolfe's classes in both Grades 7 and 8 experienced a low quality of instruction and limited OTL_u. Because both teachers perceived an average level of school capacity (see Table 2-18 and Figure 2-18), the quality of instruction and OTL_u students experienced likely had more impact on student performance than school capacity.

Table 2-18

Level of School Capacity, Prior Achievement, and Classroom Achievement for Grade 8 Students in 1998-99, by District and by Teacher

District/ Teacher ID	Level of School Capacity	PA (N)	PA Mean	CA (N)	CA Mean
MiC					
District 1					
87	3	28	245.83	53	234.80
District 2					
20	3	22	255.02	40	272.15
47	4	18	231.43	52	239.67
MiC (Conventional)					
75	3	11	237.31	33	240.82
District 2					
(None)					
Conventional					
District 1					
60	2	9	205.58	22	214.66
86	3	34	272.75	34	255.33
District 2					
5	3	12	188.15	24	189.70
24	2	1	203.63	5	192.22

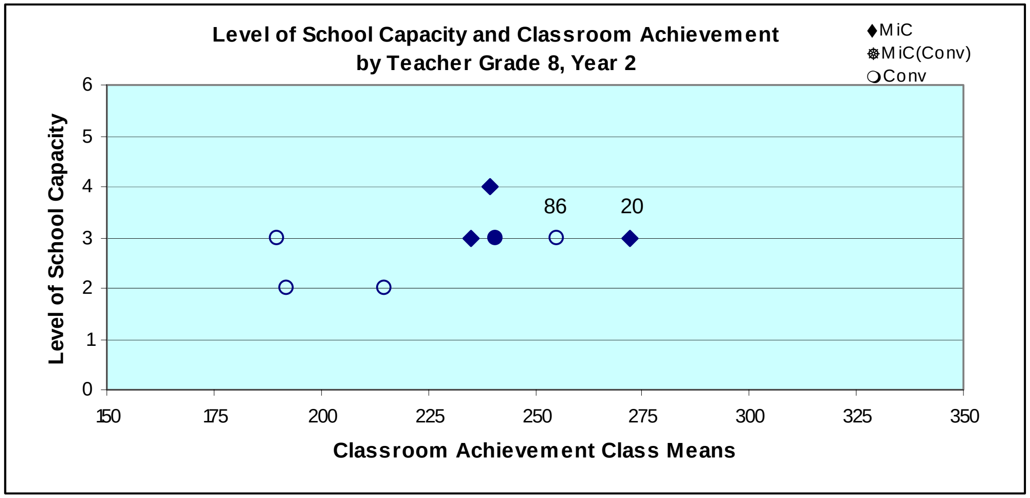


Figure 2-18. Level of school capacity for Grade 8 teachers in Year 2 compared with classroom achievement at the end of the year

Summary of Year 2 Results

In Year 2, some teachers taught MiC for the first time during the entire school year, and they faced the same challenges that MiC teachers experienced during the first study year. In addition, some students were new to the study, while other students were in the study both years. In Grade 6, all MiC students made transitions from elementary to middle schools in which the structural and cultural conditions were different from what they experienced in elementary grades. Despite these changes, the performance of students in only one teacher's classes increased significantly, even with high quality of instruction and moderate or high levels of OTL_u. Students in the examined conventional group did not experience such changes because they were in the same school for both Grades 5 and 6. Their performance increased substantially, even though the instruction and OTL_u they experienced were at low levels. A second factor likely affected student performance. When MiC teachers worked in isolation, student performance showed little improvement. In contrast, when one MiC teacher participated in formal discussions about curriculum, instruction, and assessment with other mathematics teachers who had previously taught the sixth-grade units, student performance increased substantially. This type of collaboration likely supported the quality of instruction and OTL_u students experienced. In Grade 7, when the quality of instruction MiC (Conventional) students experienced surpassed the instruction experienced in the prior year, students showed substantial gains in performance. In the case of one teacher, these gains may also have been influenced by supplementing seventh-grade MiC units with portions of units from different grade levels. Furthermore, when the quality of instruction and OTL_u were high in both Grades 7 and 8, MiC students experienced considerable gains in performance. Conversely, when the quality of instruction and OTL_u were low in both Grades 7 and 8, students who studied a conventional curriculum experienced considerable losses in student performance.

Comparable Groups in Year 3

In Year 3, MiC (Conventional) and Conventional teachers joined the study in order to follow students longitudinally. New students were not added to the study. At the end of the year, students had participated in the study either two or three years. Classroom achievement from Year 2 was used as a measure of prior achievement in Year 3. In many cases, student groups varied in sample size from prior achievement to classroom achievement. Two factors influenced sample size: completion of both study assessments in Year 2 and change in parental consent from Year 2 to Year 3.

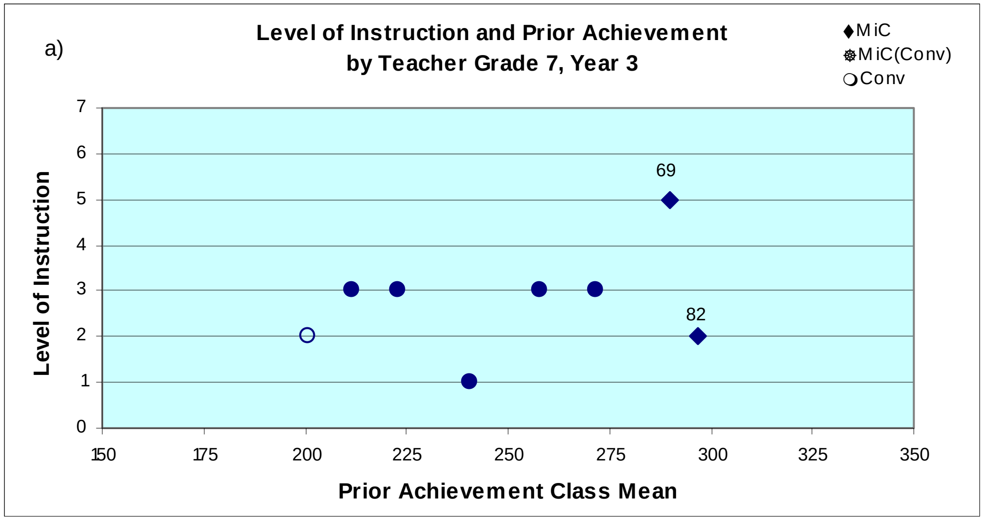
Grade 7

Table 2-19 and Figure 2-19a show the mean PA score (CA from the prior spring) and the level of instruction assigned to each seventh-grade teacher. Of interest here is the performance of students in classes of Redling, MiC Teacher 69, and Friedman, Conventional Teacher 82. Mean PA scores for these groups are comparable (289.94 and 296.92, respectively), and students experienced different quality of instruction. In addition, these students experienced substantial gains in CA performance in Year 2. Redling looped with her students from Grade 6 to Grade 7, and Friedman taught the same students in Grades 6 and 7. Students in both groups experienced comparable instruction in Years 2 and 3. With respect to CA (see Table 2-19 and Figure 2-19b), although mean score declined 2 points in Redling's classes, the mean score for students in Friedman's classes decreased 22 points. With the high quality of instruction (level 5) in Redling's classes over the two-year period, one would expect to see increases in student performance but performance stayed roughly the same. The OTL_u students experienced is of interest here. Friedman also taught the same students for two years, but the quality of instruction was low. In this case, more of the same kind of instruction over two years resulted in substantially lower student performance.

Table 2-19

Level of Instruction, Prior Achievement, and Classroom Achievement for Grade 7 Students in 1999-2000, by District and by Teacher

District/ Teacher ID	Level of Instruction	PA *(N)*	PA Mean	CA *(N)*	CA Mean
		MiC			
District 1					
(None)					
District 2					
69	5	30	289.94	29	287.49
		MiC (Conventional)			
District 1					
13	1	8	240.62	7	263.74
74	3	41	258.01	43	291.06
90	3	24	271.50	29	285.95
District 2					
52	3	14	211.78	16	230.81
94	3	15	223.10	16	238.40
		Conventional			
District 1					
77	2	3	200.81	12	232.39
District 2					
82	2	16	296.92	17	274.52



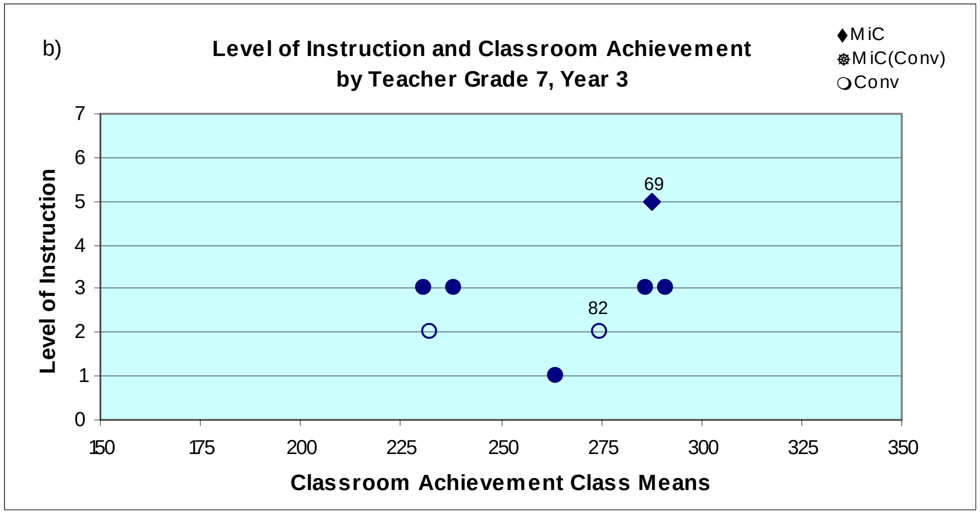


Figure 2-19. Level of instruction for Grade 7 teachers in Year 3 compared with a) prior achievement; b) classroom achievement at the end of the year

Students in Redling's (Teacher 69's) classes experienced a moderate level of OTL_u in Year 3 (see Table 2-20 and Figure 2-20). Redling taught units in depth, but only three units were taught this year. She supplemented MiC with preparation for the district standardized tests, a district initiative to have 30 minutes of silent reading in every subject class, and a project involving construction of kites with anti-tobacco messages from a grant she was awarded. She presented lessons that emphasized conceptual understanding, but she also emphasized procedural understanding. Connections among mathematical ideas and connections between mathematics and students' lives were not discussed in detail. The lower level of OTL_u, compared with high OTL_u in Year 2, likely influenced the decrease in performance. Friedman's (Teacher 82's) students experienced limited OTL_u in both years. Curricular content spanned a vast content plane with little or no depth, few questions fostered conceptual understanding, and connections were not discussed in lessons. These students experienced the low levels of instruction and OTL_u over the two-year period, which likely influenced declining CA performance.

Table 2-20
Level of Opportunity to Learn with Understanding, Prior Achievement, and Classroom Achievement for Grade 7 Students in 1999-2000, by District and by Teacher

District/ Teacher ID	Level of OTLu	PA *(N)*	PA Mean	CA *(N)*	CA Mean
MiC					
District 1					
(None)					
District 2					
69	3	30	289.94	29	287.49
MiC (Conventional)					
District 1					
13	2	8	240.62	7	263.74
74	3	41	258.01	43	291.06
90	4	24	271.50	29	285.95
District 2					
52	2	14	211.78	16	230.81
94	2	15	223.10	16	238.40
Conventional					
District 1					
77	2	3	200.81	12	232.39
District 2					
82	2	16	296.92	17	274.52

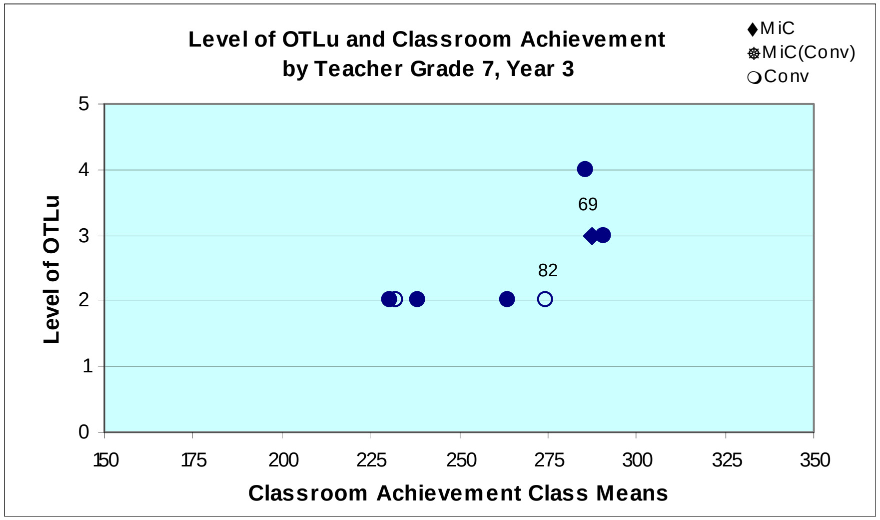


Figure 2-20. Level of opportunity to learn with understanding for Grade 7 teachers in Year 3 compared with classroom achievement at the end of the year

Because both teachers perceived an average level of school capacity (see Table 2-21 and Figure 2-21), and students were in the same schools both years, the changes in performance are likely attributable to differences in instruction and OTL_u. More of the same high quality of instruction (Level 5) over two years and lower OTL_u (Level 3 down from Level 4) in the second year resulted in little change in MiC Teacher Redling's classes. In contrast, more of the same low levels of instruction (Level 2) and OTL_u (Level 2) in Conventional Teacher Friedman's classes led to substantially lower student performance. For both groups, these results occurred after substantial gains in CA performance in the previous year.

Table 2-21
Level of School Capacity, Prior Achievement, and Classroom Achievement for Grade 7 Students in 1999-2000, by District and by Teacher

District/ Teacher ID	Level of School Capacity	PA		CA	
		(N)	Mean	*(N)*	Mean
MiC					
District 1					
(None)					
District 2					
69	3	30	289.94	29	287.49
MiC (Conventional)					
District 1					
13	2	8	240.62	7	263.74
74	2	41	258.01	43	291.06
90	3	24	271.50	29	285.95
District 2					
52	3	14	211.78	16	230.81
94		15	223.10	16	238.40
Conventional					
District 1					
77	2	3	200.81	12	232.39
District 2					
82	3	16	296.92	17	274.52

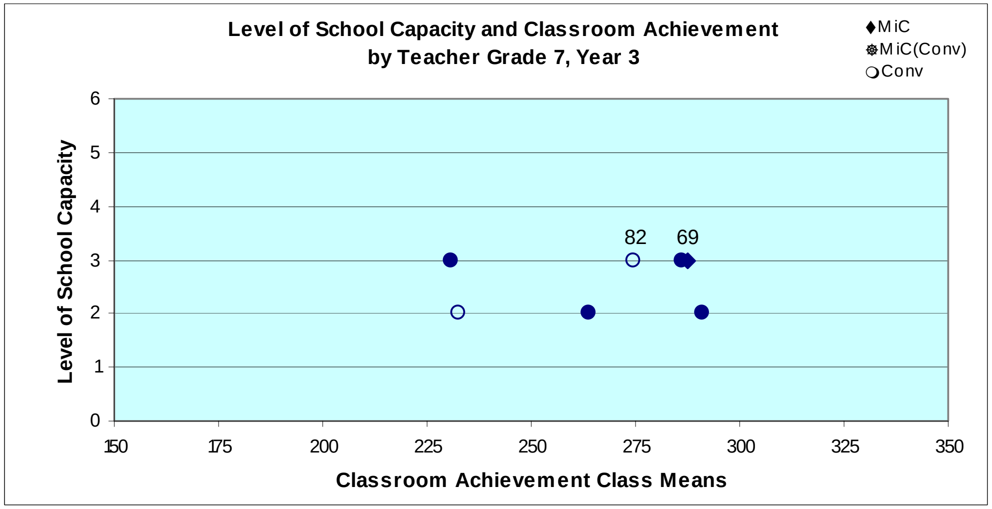


Figure 2-21. Level of school capacity for Grade 7 teachers in Year 3 compared with classroom achievement at the end of the year

Grade 8

In Grade 8, one cluster of students is of interest. Dillard, Gallardo, and Reichers, MiC Teachers 45, 62, and 87, and Wolfe, Conventional Teacher 86,[5] taught students who had mean PA scores between 250 and 270, and students experienced different levels of the quality of instruction (see Table 2-22 and Figure 2-22a). As instruction unfolded during the school year, Dillard attempted to teach for conceptual understanding. Student explanations were focused on procedures rather than on elaboration of reasoning, and multiple strategies were generally not elicited. Evidence of student learning was limited to correct answers. Feedback was indirectly responsive to student needs. Student–student conversation was limited. Dillard looped with his classes from Grade 6 to Grade 7 to Grade 8, teaching the same core of students for three years. These students experienced similar levels of instruction over the three-year period.

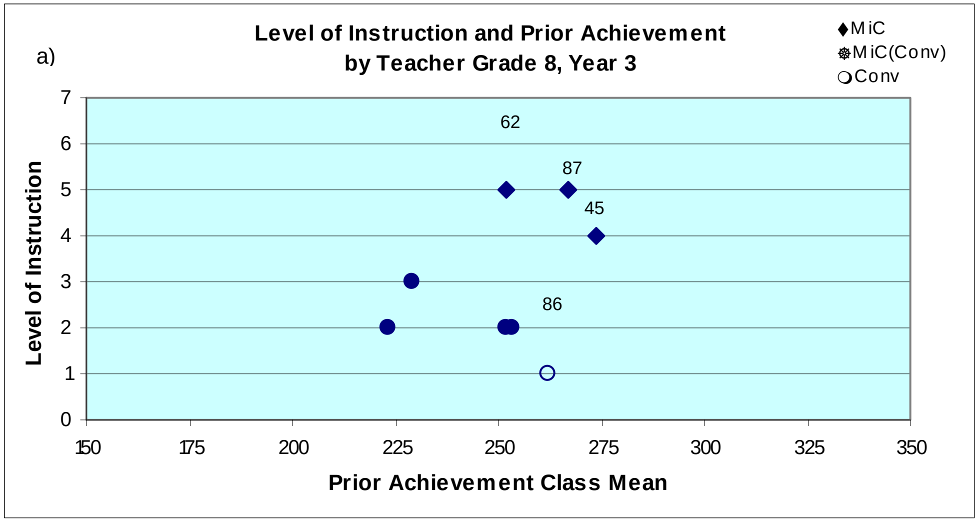
Gallardo, MiC Teacher 62, worked toward conceptual understanding. He shared the mathematical work with his students, posed questions that encouraged students to explain their thinking, and helped students make connections among mathematical ideas.

[5] In Monograph 6, Chapter 1, Wolfe, Conventional Teacher 86, was classified as an outlier in Year 3 in the cross-tabulation of the levels of instruction and opportunity to learn. Therefore, she was not included in the analysis of student performance by treatment group. However, we decided to include her classes in the analysis of comparable groups discussed in this chapter.

He was effective at eliciting student responses and orchestrating substantive whole class discussions. Student interactions were used to promote making sense of tasks, responses to tasks, and mathematical conventions. Feedback was ongoing and was offered by both teacher and students. Feedback addressed skills, procedures, and concepts. Gallardo taught the same students in both Grades 7 and 8. In Grade 7, students experienced instruction at Level 4, and in Grade 8, Level 6.

Table 2-22
Level of Instruction, Prior Achievement, and Classroom Achievement for Grade 8 Students in 1999-2000, by District and by Teacher

District/ Teacher ID	Level of Instruction	PA		CA	
		(N)	Mean	*(N)*	Mean
MiC					
District 1					
87	5	42	266.75	51	268.35
District 2					
45	4	13	273.79	14	285.12
62	5	20	252.17	11	289.32
MiC (Conventional)					
District 1					
75	2	30	253.18	34	257.58
91	2	11	251.84	23	245.33
District 2					
70	3	49	229.20	47	212.82
95	2	13	223.31	12	254.84
Conventional					
District 1					
86	1	42	262.14	28	250.81
92	4	3	217.56	3	267.45
District 2					
(None)					



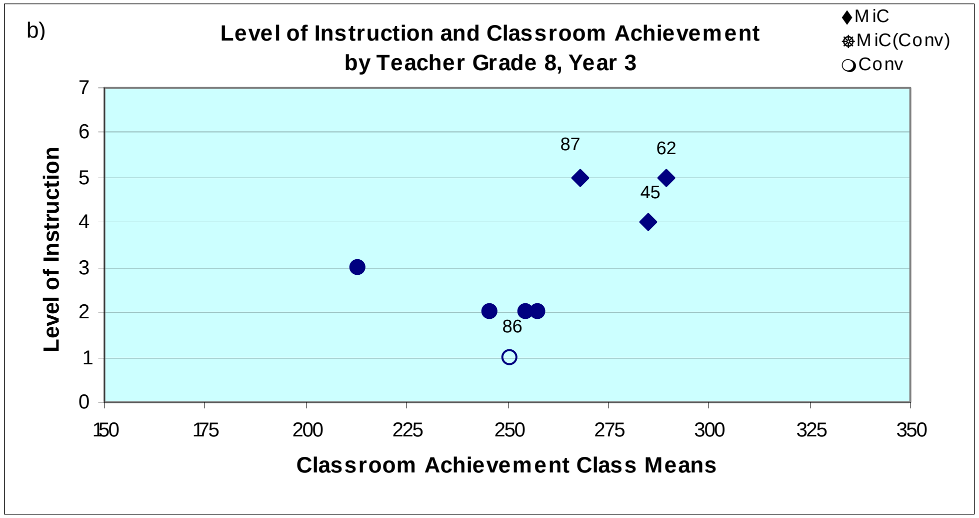


Figure 2-22. Level of instruction for Grade 8 teachers in Year 3 compared with a) prior achievement; b) classroom achievement at the end of the year

Reichers, Teacher 87, shared the mathematical work with her students, posed questions that encouraged students to explain their thinking, and helped students make connections among mathematical ideas. Reichers sought student explanations in addition to procedural understanding as evidence of student learning. A blend of clear and mathematically sound feedback addressed skills, procedures, and concepts. Students in Reichers' classes were in classes taught by Muldoon and Bartlett, MiC (Conventional) Teachers 74 and 88, the previous year, during which they experienced large gains in performance despite instruction and OTL_u at lower levels.

Wolfe, Conventional Teacher 86, presented particular procedures, and students practiced them in a rote fashion. Inquiry during lessons was limited to lower order thinking, and little attention was given to conceptual understanding. Procedural understanding was sought as evidence of student learning, and greater attention was given to student homework and classwork for instructional decision-making. Feedback was indirectly responsive to student needs, it was limited to checking for accurate answers, and it seldom addressed student misconceptions. Note that Wolfe's students were in classes taught by other study teachers during the previous years in which they experienced instruction at Level 1. That is, students in Wolfe's classes experienced a lower quality of instruction for three years than students in classes of MiC and MiC (Conventional) teachers.

With respect to CA, the group with the largest gain in performance experienced high quality of instruction; students with MiC Teacher Gallardo, Teacher 62, gaining 37 points (see Table 2-22 and Figure 2-22b). Also noteworthy is the loss of 11 points for students in Wolfe's, Conventional Teacher 86's classes. However, the gain in performance in Reichers' classes was minimal. This reflects the trends seen in longitudinal data for the three-year cohort beginning in Grade 6 reported in Monograph 6, Chapter 4. Substantial gains in performance were noted from Grade 6 to Grade 7, while minimal gains were noted from Grade 7 to Grade 8. Confounding factors likely affected Dillard's, MiC Teacher 45's, classes. The gain of 11 points for students is confounded with the notable lower ratings in observation data when he began to use a traditional high school algebra textbook midway through the second semester. Even so, students in Dillard's classes experienced a high level of OTL_u (see Table 2-23 and Figure 2-23). Dillard taught four units, two each in algebra and geometry, and supplemented MiC with conventional high school algebra textbook and a resource that provided manipulatives for solving equations. He presented lessons that at times emphasized conceptual understanding, but the main focus of lessons was on building students' procedural understanding. Observed student conjectures consisted mainly of making connections between a new problem and problems previously seen. Connections among mathematical ideas and connections between mathematics and students' lives were not discussed in detail. Dillard looped with his classes throughout their middle-school years. Because he looped with his students, he taught different MiC units each year. Students experienced similar levels of instruction and OTL_u over the three-year period. However, students recovered from a 3-point decline in mean CA performance in Year 2 to an 11-point gain in mean performance in Year 3.

Table 2-23

Level of Opportunity to Learn with Understanding, Prior Achievement, and Classroom Achievement for Grade 8 Students in 1999-2000, by District and by Teacher

District/ Teacher ID	Level of OTLu	PA *(N)*	PA Mean	CA *(N)*	CA Mean
MiC					
District 1					
87	4	42	266.75	51	268.35
District 2					
45	4	13	273.79	14	285.12
62	4	20	252.17	11	289.32
MiC (Conventional)					
District 1					
75	3	30	253.18	34	257.58
91	2	11	251.84	23	245.33
District 2					
70	2	49	229.20	47	212.82
95	2	13	223.31	12	254.84
Conventional					
District 1					
86	1	42	262.14	28	250.81
92	3	3	217.56	3	267.45
District 2					
(None)					

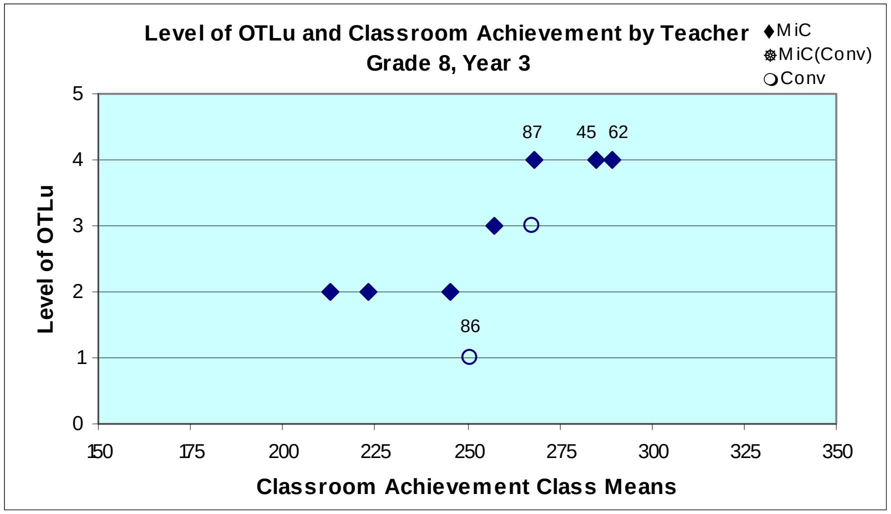


Figure 2-23. Level of opportunity to learn with understanding for Grade 8 teachers in Year 3 compared with classroom achievement at the end of the year

Students in Gallardo's, Teacher 62's, classes experienced high OTL_u during the school year. Gallardo taught an integrated curriculum, teaching three MiC units (one sixth-grade number) in two content strands. He supplemented the curriculum with practice of skills for the district standardized tests. Lessons promoted linking procedural and conceptual knowledge. Observed student conjectures consisted mainly of making connections between a new problem and problems previously seen. Connections among mathematical ideas were generally promoted, but connections between mathematics and students' lives were not discussed in detail. Gallardo taught the same students in both Grades 7 and 8. The quality of instruction (Levels 4 and 6) and moderate to high OTL_u (Levels 3 and 4) students experienced likely affected significant gains in student performance.

Students in Reichers', Teacher 87's, classes experienced high OTL_u. Teacher 87 presented a comprehensive, integrated curriculum, teaching five units (two seventh-grade units to support the eighth-grade units) in three content strands. MiC units were supplemented with a resource that provided manipulatives for solving equations. At times, procedural understanding was frequently given a central focus of instruction. Observed student conjectures consisted of investigating the truthfulness of particular statements, while at other times, conjectures stemmed from making connections between a new problem and problems previously seen. Connections among mathematical ideas were explored, but connections between mathematics and students' lives were not discussed in detail. As noted earlier, students in Reichers' classes were in classes taught by Muldoon and Bartlett, MiC (Conventional) Teachers 74 and 88, the previous year, during which they experienced large gains in performance despite instruction and OTL_u at lower levels

(Levels 2 and 3). However, the gain in performance in Reichers' classes was minimal. This reflects the trends seen in longitudinal data for the three-year cohort beginning in Grade 6 reported in Monograph 6, Chapter 4. Substantial gains in performance were noted from Grade 6 to Grade 7, while minimal gains were noted from Grade 7 to Grade 8.

Students in Wolfe's, Teacher 86's, classes experienced a low level of OTL_u. Curricular content spanned a vast content plane with little or no depth. The content was presented as disparate pieces of knowledge heavily laden with vocabulary and prescribed algorithms. However, lack of teacher preparation and/or student participation undermined the intent of the curriculum. Few questions fostered conceptual understanding, and making conjectures was not encouraged. Connections between mathematics and students' lives were not discussed in lessons. As noted earlier, Wolfe's students were in classes taught by other study teachers during the previous years in which experienced similar instruction (Level 1) and OTL_u (Levels 1 and 2). That is, students in Wolfe's classes experienced a low quality of instruction and OTL_u for three years. This likely influenced the 11-point decline in CA performance.

Table 2-24 and Figure 2-24 show data about the teachers' perceptions of school capacity. Dillard and Gallardo, Teachers 45 and 62, taught in different schools that had similar enrollments and lengths of class periods. Both teachers perceived high school capacity. Both felt that the principal and teachers had visions of teaching and learning mathematics that were aligned on some ideas, but were incompatible on others. They received strong administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. Professional development opportunities were available at their schools, and monetary support was available for other opportunities. Both met regularly with other teachers for collaboration, and common planning time was available for mathematics teachers on their teaching teams. Both felt they had a high level of influence in planning and teaching mathematics, educational policy such as curriculum and discipline, and the content of professional development programs. They believed that district and state standardized testing programs had little influence over curriculum and instruction because they felt that their current practices were sufficient in helping students prepare for these tests. With respect to CA scores, students with both teachers showed increases in performance. However, the mean performance for students in Gallardo's classes was over three times the increase shown by students in Dillard's classes, even though both teachers perceived high school capacity. Therefore, instruction and OTL_u likely had more influence on student performance.

Reichers, Teacher 87, transferred from Fernwood Middle School to Von Humboldt Middle School this year, the same school as Muldoon and Bartlett, MiC (Conventional) Teachers 74 and 88. The school capacity she perceived was lower this year than in the past year. She had looked forward to working with the other eighth-grade study teacher at Von Humboldt, but this collaboration did not take place. In lieu of formal meetings for mathematics teachers and common planning time, teachers informally met regularly for collaboration. Reichers felt that the principal and teachers had visions of teaching and learning mathematics that were aligned on some ideas, but were incompatible on others. She felt that she received weak administrative support in terms of clearly communicated expectations, support for selecting instructional materials, changes in instructional practice, and changes in policy. Teachers were not encouraged to observe exemplary teaching, but all teachers were encouraged to network with teachers in the district. Reichers felt that she had an average level of influence in planning and teaching mathematics, and limited influence over educational policies related to curriculum, content of professional development programs, and discipline. She felt that faculty and staff were not very committed to academic excellence and that teachers somewhat supported one another in their efforts to improve instruction. District and state standardized tests influenced instruction, as she added practice on computation. School capacity may have affected the minimal increase in performance for Reichers' students in Year 3.

Table 2-24
Level of School Capacity, Prior Achievement, and Classroom Achievement for
Grade 8 Students in 1999-2000, by District and by Teacher

District/ Teacher ID	Level of School Capacity	PA (N)	PA Mean	CA (N)	CA Mean
MiC					
87	1	42	266.75	51	268.35
District 2					
45	4	13	273.79	14	285.12
62	5	20	252.17	11	289.32
MiC (Conventional)					
District 1					
75	2	30	253.18	34	257.58
91	3	11	251.84	23	245.33
District 2					
70	3	49	229.20	47	212.82
95	4	13	223.31	12	254.84
Conventional					
District 1					
86	3*	42	262.14	28	250.81
92	2	3	217.56	3	267.45
District 2					
(None)					

* This is the level of school capacity from Year 2. Because the teacher did not complete teacher questionnaires in Year 3, the level of school capacity for Year 3 could not be determined.

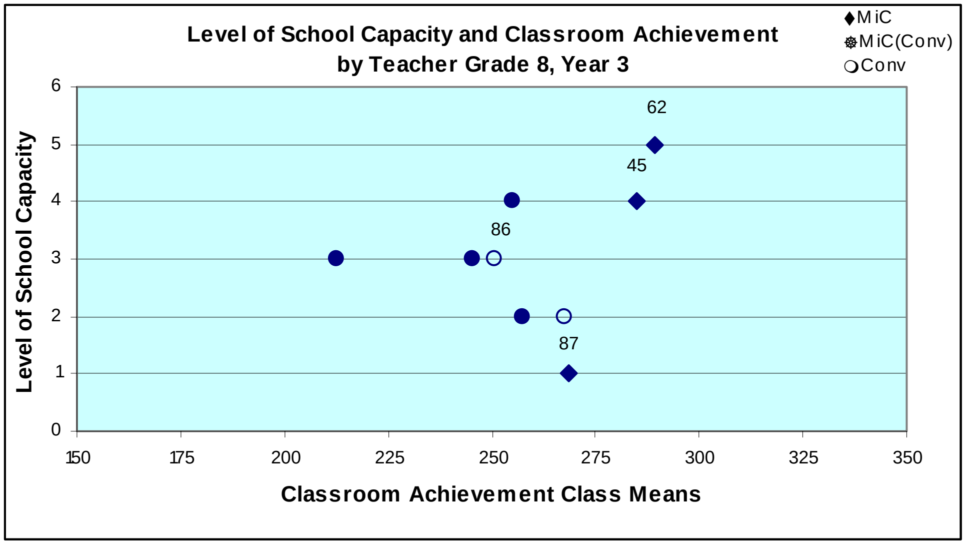


Figure 2-24. Level of school capacity for Grade 8 teachers in Year 3 compared with classroom achievement at the end of the year

Wolfe, Conventional Teacher 86, taught in the same magnet school for the performing arts as she did in the previous year. The school capacity she perceived seemed not to influence student performance, which was more likely attributed to low levels of instruction and OTL_u that these students experienced in three years of mathematics instruction at this school.

Summary of Year 3 Results

In Year 3, student performance seemed not to be influenced by teachers' perceptions of school capacity. Rather, student performance was likely attributable to differences in the quality of instruction and OTL_u they experienced over time. For MiC teachers in the examined clusters, performance generally increased when students experienced a high quality of instruction and moderate or high levels of OTL_u over time. In contrast, for conventional teachers, performance declined when students experienced low levels of instruction and OTL_u over time. The results were evident whether students had the same teachers in consecutive grade levels or different teachers throughout their middle-school years.

Conclusion

The analyses presented in this chapter suggest that for groups of students who began with comparable prior achievement, investigations of instruction, opportunity to learn mathematics with understanding, and school capacity provided important insights. In the clusters were examined, the results suggest that over time students who studied MiC and experienced high quality of instruction and opportunity to learn with understanding showed gains in performance. In Grade 6, student performance remained the same or was negatively affected when teachers implemented MiC in isolation in their schools and when students transitioned from elementary to middle schools. However, when teachers collaborated in meaningful discussions about mathematics curriculum, instruction, and assessment, the influence of these factors was mitigated and student performance increased. In Grade 7, student performance was positively affected when MiC (Conventional) teachers taught MiC throughout the school year, even when instruction was reflective of good conventional pedagogy. On the other hand, student performance stayed the same or declined considerably when MiC (Conventional) teachers substantially supplemented MiC with skill practice or when they replaced MiC with a conventional textbook. In contrast, students who studied conventional curricula and experienced low quality of instruction and opportunity to learn with understanding showed declines in performance.

CHAPTER 3: OTHER RESULTS

Thomas A. Romberg, Mary C. Shafer, and Lorene Folgert

In Chapters 1 and 2 of this monograph, the classroom achievement index (CA) was the dependent variable used to indicate differences in the three treatment groups in this study. However, data were collected on four other dependent variables: the mean percentile on the standardized test administered by the district, a set of anchor items on the External Assessment System designed for the study, a set of attitude scales, and the *Collis-Romberg Mathematical Problem Solving Profiles* (Collis & Romberg, 1992). In this chapter, the differences between the three treatment groups in Districts 1 and 2 on these variables are summarized.

Differences in Standardized Test Percentiles

In this analysis, the treatment group's mean standardized test percentile at the end of the school year is the dependent variable. As part of the regular standardized testing programs in each district, the *TerraNova* (CTB/McGraw-Hill, 1997) and the *Stanford Mathematics Achievement Test* (Harcourt Brace Educational Measurement, 1997) were administered in Districts 1 and 2, respectively. National percentile rankings were used for purposes of comparison across groups (see Romberg, Webb, Folgert, Shafer, 2004 for complete data sets.) As was done for CA scores in Chapter 1, the group's mean percentile rankings from the test given the prior year were used as a covariate in this analysis. The post and pre mean percentiles for the three treatment groups are shown in Table 3-1. The MiC group had higher pre percentile rankings than the other two groups. Note also that the post percentile rankings are similar to the pre percentile rankings for all groups and are slightly lower.

Table 3-1.
Post and Pre STS Percentile Means and Standard Deviations for the Three Treatment Groups

STS Post Level of treatment	N	Mean	Std Dev	STS Pre Level of treatment	N	Mean	Std Dev
MiC	28	51.06	15.41	MiC	28	52.21	16.50
MiC (Conventional)	28	45.18	13.33	MiC (Conventional)	28	48.31	15.87
Conventional	23	47.89	19.60	Conventional	23	48.92	19.39

The post least squares mean percentiles for the three treatment groups based on the adjustment due to differences in the pre percentiles are shown in Table 3-2. The order of the means is the same but the differences between the means are reduced.

Table 3-2.
Least Squares Means for the Three Treatment Groups

Level of treatment	N	LS Mean
MiC	28	49.17
MiC (Conventional)	28	46.44
Conventional	23	48.66

Next analysis of covariance (ANCOVA) was carried out using a SAS program (SAS Institute, 2000). The results are shown in Table 3-3.

Table 3-3.
ANCOVA for the Three Treatment Groups with Post STS Percentile Means as the Dependent Variable and Pre STS Percentile Means as the Covariate

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	15053.82	5017.94	73.91	<.0001
Error	75	5092.08	67.89		
Corrected Total	78	20145.90			

R-Square	Coeff Var	Root MSE	sts Mean
0.75	17.15	8.24	48.05

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMENT	2	115.51	57.76	0.85	0.43
PA MEAN	1	14569.19	14569.19	214.59	<.0001

This analysis shows that there is no overall treatment effect for the three treatment groups on the post STS percentile means. However, not surprisingly, the pre STS percentile means strongly predict the post STS percentile means (accounting for 75% of the variance). When the contrasts between the three treatment groups are examined (see Table 3-4), there are no significant differences between the means of the MiC, MiC (Conventional) and Conventional groups.[6] Also, there are moderate effect size differences between the MiC and MiC (Conventional) groups.

[6] Given the power of the statistical tests in this study p-values <.10 are considered to imply important differences, and p-values <.05 are statistically significant differences. Effect sizes (Cohen's d) > 0.50 are considered substantial, and effect sizes > 0.30 are considered moderate. Also, an effect size correlation > 0.30 is considered strong (Cohen, 1988).

Table 3-4.
Contrasts in the Post STS Percentile Means of the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	2.73	2.21	1.23	0.22	0.33	0.16
MiC vs Conventional	.52	2.33	0.22	0.82	0.06	0.03
MiC (Conventional) vs Conventional	-2.21	2.32	-0.96	0.34	-0.27	-0.13

Differences in Performance on the Anchor Items

The second set of variables is the treatment group's mean score on each of the twenty anchor items in the External Assessment System (EAS). In order to examine growth in mathematical knowledge and skills over time, a set of the same items was included on each grade-specific EAS assessment. These items are referred to as anchor items. (See Chapter 1 in Monograph #4 for more details about the construction of the EAS.) The least-square means for each of the items were submitted to an analysis of covariance (ANCOVA) with prior achievement (PA) as the covariate. There were important or significant differences in performance between the three treatment groups on eight of the 20 items (2 number items, 3 algebra items, 1 geometry item, and 2 statistics items). The items and the results of the analysis are shown below.

Number Item: Tourists on Bus

Two groups of tourists each have 60 people. If $\frac{3}{4}$ of the first group and $\frac{2}{3}$ of the second group board buses to travel to a museum, how many more people in the first group board buses than in the second group?

A) 2 B) 4 C) 5 D) 40 E) 45

Figure 3-1. External Assessment System Number Item, Tourists on Bus.

The least-squares means for the three treatment groups on this item are shown in Table 3-5 and the contrasts between the three groups are in Table 3-6.

Table 3-5.
Least Squares Means for the Tourists on Bus Item by the Three Treatment Groups

Level of treatment	LS Mean
MiC	0.48
MiC (Conventional)	0.43
Conventional	0.36

Table 3-6.
Contrasts in the Means of the Tourists on Bus Item by the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	0.05	0.04	1.25	0.22	0.34	0.17
MiC vs Conventional	0.11	0.04	2.93	0.00	0.83	0.38
MiC (Conventional) vs Conventional	0.07	0.04	1.75	0.08	0.49	0.24

The mean performance of the MiC group is significantly higher than the Conventional group's performance with a substantial effect size. The MiC (Conventional) group's performance is importantly higher than the Conventional group's performance.

Number Item: Tip Calculation

Of the following, which is the closest approximation of a 15 percent tip on a restaurant check of $24.99?

A) $2.50

B) $3.00

C) $3.75

D) $4.50

E) $5.00

Figure 3-2. External Assessment System Number Item, Tip Calculation.

The least-squares means for the three treatment groups on this item are shown in Table 3-7 and the contrasts between the three groups are in Table 3-8.

Table 3-7.
Least Squares Means for the Tip Calculation Item by the Three Treatment Groups

Level of treatment	LS Mean
MiC	0.13
MiC (Conventional)	0.10
Conventional	0.06

Table 3-8.
Contrasts in the Means of the Tip Calculation Item by the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	0.03	0.02	1.67	0.10	0.45	0.21
MiC vs Conventional	0.07	0.02	3.74	0.00	1.06	0.47
MiC (Conventional) vs Conventional	0.04	0.02	2.16	0.03	0.61	0.29

The mean performance of the MiC group is significantly higher than the Conventional group's performance and importantly higher than the MiC (Conventional) group's performance. Also, the MiC (Conventional) group's mean performance is significantly higher than the Conventional group's performance.

Algebra Item: 4m

If m represents a positive number, which of these is equivalent to $m + m + m + m$?

A) $m + 4$

B) $4m$

C) m^4

D) $4(m + 1)$

Figure 3-3. External Assessment System Algebra Item, 4m

The least-squares means for the three treatment groups on this item are shown in Table 3-9 and the contrasts between the three groups are in Table 3-10.

Table 3-9.
Least Squares Means for the 4m Item by the Three Treatment Groups

Level of treatment	LS Mean
MiC	0.29
MiC (Conventional)	0.23
Conventional	0.32

Table 3-10.
Contrasts in the Means of the 4m Item by the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	0.06	0.04	1.45	0.15	0.39	0.19
MiC vs Conventional	-0.03	0.04	-0.67	0.51	-0.19	0.09
MiC (Conventional) vs Conventional	-0.09	0.04	-2.05	0.04	-0.58	0.28

The mean performance of the MiC group is not different from the Conventional group's performance. Also, the Conventional group's mean performance is significantly higher that the MiC (Conventional) group's mean.

Algebra Item: Points on a Line

A straight line on a graph passes through the points (3, 2) and (4, 4). Which of these points also lies on the line?

A) (1, 1)
B) (2, 4)
C) (5, 6)
D) (6, 3)
E) (6, 5)

Figure 3-4. External Assessment System Algebra Item, Points on a Line.

The least-squares means for the three treatment groups on this item are shown in Table 3-11 and the contrasts between the three groups are in Table 3-12.

Table 3-11.
Least Squares Means for the Points on a Line Item by the Three Treatment Groups

Level of treatment	LS Mean
MiC	0.29
MiC (Conventional)	0.25
Conventional	0.22

Table 3-12.
Contrasts in the Means of the Points on a Line Item by the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	0.04	0.03	1.73	0.09	0.47	0.23
MiC vs Conventional	0.07	0.03	2.71	0.01	0.77	0.36
MiC (Conventional) vs Conventional	0.03	0.03	1.07	0.29	0.30	0.15

The mean performance of the MiC group is significantly higher than the Conventional group's performance with a substantial effect size, and importantly higher than the MiC (Conventional) group's performance.

Algebra Item: Dropped Ball

A rubber ball rebounds to half the height it drops. If the ball is dropped from a rooftop 18 m above the ground, what is the total distance traveled by the time it hits the ground the third time?

A) 31.5 m B) 40.5 m C) 45 m D) 63 m

Figure 3-5. External Assessment System Algebra item, Dropped Ball.

The least-squares means for the three treatment groups on this item are shown in Table 3-13 and the contrasts between the three groups are in Table 3-14.

Table 3-13.
Least Squares Means for the Dropped Ball Item by the Three Treatment Groups

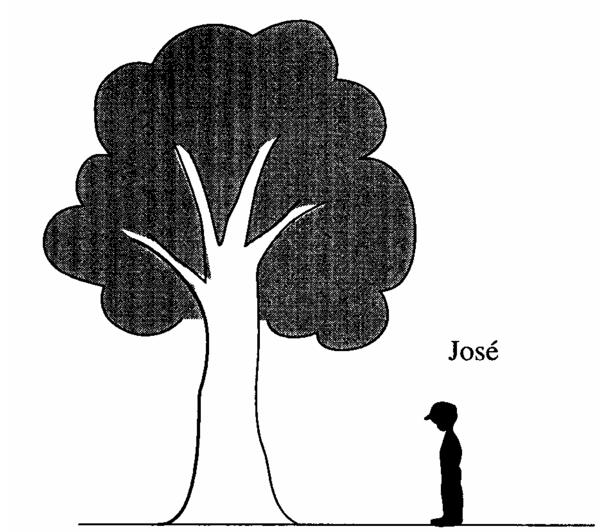
Level of treatment	LS Mean
MiC	0.30
MiC (Conventional)	0.25
Conventional	0.28

Table 3-14.
Contrasts in the Means of the Dropped Ball Item by the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	0.05	0.03	1.66	0.10	0.45	0.22
MiC vs Conventional	0.03	0.03	0.81	0.42	0.23	0.11
MiC (Conventional) vs Conventional	-0.02	0.03	-0.78	0.44	-0.22	0.11

The mean performance of the MiC group is importantly higher than the MiC (Conventional) group's performance.

Geometry Item: Jose's Tree



José is 1.5 m tall. About how tall is the tree?

A) 4 m

B) 6 m

C) 8 m

D) 10 m

Figure 3-6. External Assessment System Geometry Item, Jose's Tree.

The least-squares means for the three treatment groups on this item are shown in Table 3-15 and the contrasts between the three groups are in Table 3-16.

Table 3-15.
Least Squares Means for the Jose's Tree Item by the Three Treatment Groups

Level of treatment	LS Mean
MiC	0.49
MiC (Conventional)	0.42
Conventional	0.38

Table 3-16.
Contrasts in the Means of the Jose's Tree Item by the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	0.07	0.04	1.61	0.11	0.43	0.21
MiC vs Conventional	0.11	0.04	2.63	0.01	0.74	0.35
MiC (Conventional) vs Conventional	0.05	0.04	1.10	0.27	0.31	0.15

The mean performance of the MiC group is significantly higher than the Conventional group's performance with a substantial effect size.

Statistics Item: Blue Pen

A drawer contains 28 pens; some white, some blue, some red, and some gray. If the probability of selecting a blue pen is $\frac{2}{7}$, how many blue pens are in the drawer?

A) 4
B) 6
C) 8
D) 10
E) 20

Figure 3-7. External Assessment System Statistics Item, Blue Pen.

The least-squares means for the three treatment groups on this item are shown in Table 3-17 and the contrasts between the three groups are in Table 3-18.

Table 3-17.
Least Squares Means for the Blue Pen Item by the Three Treatment Groups

Level of treatment	LS Mean
MiC	0.43
MiC(Conventional)	0.37
Conventional	0.44

Table 3-18.
Contrasts in the Means of the Blue Pen Item by the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	0.06	0.04	1.72	0.09	0.46	0.23
MiC vs Conventional	-0.01	0.04	-0.35	0.73	-0.10	0.05
MiC (Conventional) vs Conventional	-0.07	0.04	-1.99	0.05	-0.56	0.27

The mean performance of the MiC group is not different than the Conventional group's performance, and both are importantly higher than the MiC (Conventional) group's performance with moderate to substantial effect size differences.

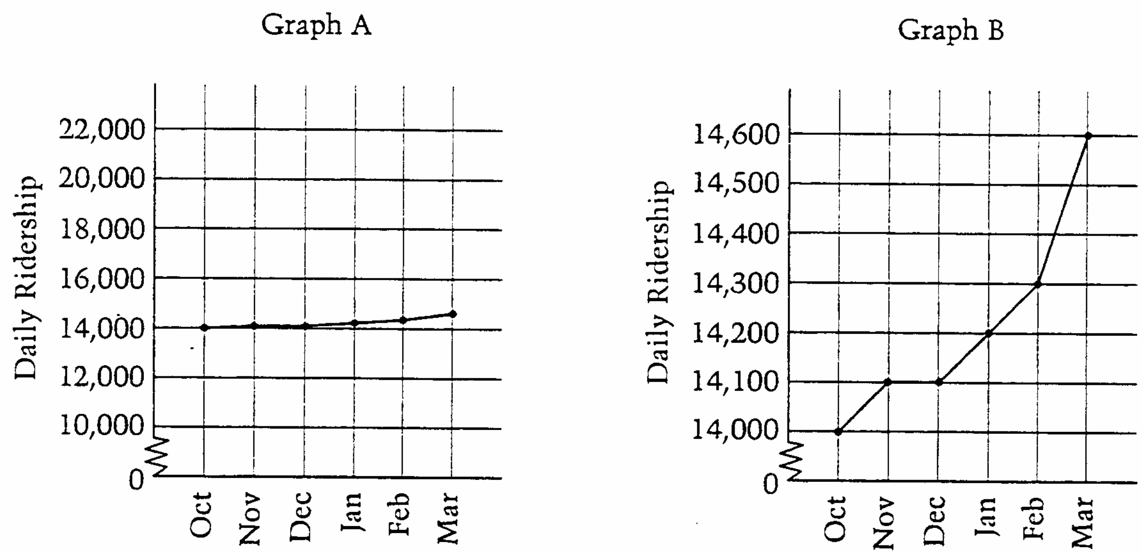
Statistics Item: Metro Rail

METRO RAIL COMPANY

Month	Daily Ridership
October	14,000
November	14,100
December	14,100
January	14,200
February	14,300
March	14,600

The data in the table above has been correctly represented by both graphs shown below.

Figure 3-7a. External Assessment System Statistics Item, Metro Rail.



Which graph would be best to help convince others that the Metro Rail Company made a lot more money from ticket sales in March than in October?

Explain your reason for making this selection

Why might people who thought that there was little difference between October and March ticket sales consider the graph you chose to be misleading?

Figure 3-7b. External Assessment System Statistics Item, Metro Rail.

The least-squares means for the three treatment groups on this item are shown in Table 3-19 and the contrasts between the three groups are in Table 3-20.

Table 3-19.
Least Squares Means for the Metro Rail Item by the Three Treatment Groups

Level of treatment	LS Mean
MiC	0.32
MiC(Conventional)	0.30
Conventional	0.24

Table 3-20.
Contrasts in the Means of the Metro Rail Item by the Three Treatment Groups

Parameter	Estimate	Standard Error	t Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	0.02	0.04	0.54	0.59	0.15	0.07
MiC vs Conventional	0.07	0.04	1.70	0.09	0.48	0.23
MiC (Conventional) vs Conventional	0.05	0.04	1.19	0.24	0.34	0.17

The mean performance of the MiC group is not different than the MiC (Conventional) group's performance and is importantly higher than the Conventional group's performance.

In summary, there were no differences in mean performance between the three treatment groups on 12 of the 20 anchor items included in the EAS. On the eight items where differences were found, the MiC group performed significantly better than the Conventional group on four items, importantly better than the Conventional group on one item, and importantly better than the MiC (Conventional) group on the other four items. The Conventional group performed significantly better than the MiC (Conventional) group on one item, and importantly better on one item. Finally, the MiC (Conventional) group performed significantly better than the Conventional group on one item and importantly better on one item.

Differences in Change in Attitudes

The third set of variables is the change in attitudes for the treatment groups. The Student Attitude Inventory (SAI) was designed to characterize the attitudes of middle-school students toward mathematics and toward themselves as learners of mathematics, taking reform-oriented goals into consideration (Shafer, Wagner, & Davis, 1997; see Monograph #2, Chapter 1 for more details about the SAI). At the beginning of the study, there were no real differences between students in the combined MiC and MiC (Conventional) group and the Conventional group on the first five subscales of the SAI—effort to succeed in mathematics, interest in and excitement about mathematics, confidence in learning mathematics, usefulness of mathematics, and communication of mathematical ideas. No significant differences (based on independent samples t-test) in the means were found between the two groups on these five subscales. The results on each subscale for all classes were similar and very positive, and the standard deviations were minimal. In the analysis of the changes in student attitudes, the least-square means for each of the five subscales were submitted to an analysis of covariance (ANCOVA) with preceding achievement (PA) as the covariate. PA did not contribute to the variance in any of these subscales, and there were no differences between treatment groups on these subscales. (Complete data sets can be found in Romberg, Shafer, LeMire, & Folgert, 2005.)

The general perception subscale of the SAI contained 16 statements that measured attitudes related to calculator use, the nature of mathematics, the learning of mathematics, and connections of mathematics to other school subjects. As with the other subscales, at the beginning of the study, the results on the general perception subscale were similar and very positive for all classes, and little variance was evident in class means. In two instances, significant differences (based on independent samples t-test for the combined MiC and MiC (Conventional) group and the Conventional group) were noted in favor of the MiC group. On Statement 16, “It’s okay if I solve a math problem differently than my classmates do,” the MiC group agreed more strongly than the Conventional group. On Statement 44, “When my teacher asks a question, I will get it right if I had memorized the correct rule or fact,” the MiC group disagreed more strongly than the Conventional group. In contrast, on Statement 3, “I feel sure that I’m able to learn new ideas in math class,” the Conventional group agreed more strongly than the MiC group.

In the analysis of changes in attitudes, the least-square means for each of the 16 statements in the general perception subscale were submitted to an analysis of covariance (ANCOVA) with preceding achievement (PA) as the covariate. PA contributed significantly to students’ change in attitudes only on Statements 6 and 16, and an important difference was found on Statement 4. The results for Statement 6, “If I use a calculator to solve a problem, I can be sure it will always give me the right answer,” are shown in Table 3-21.

Table 3-21.
ANCOVA for the Three Treatment Groups with Statement 6 Means as the Dependent Variable and Preceding Achievement as the Covariate

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.82	0.27	3.44	0.02
Error	62	4.91	0.08		
Corrected Total	65	5.73			

R-Square	Coeff Var	Root MSE	SAIGENPERC6CHG Mean
0.14	220.18	0.28	0.13

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMENT	2	0.12	0.06	0.77	0.47
PA MEAN	1	0.73	0.73	9.18	0.00

However, there were no differences between treatment groups on Statement 6 as shown in Tables 3-22 and 3-23.

Table 3-22.
Least-Squares Means for Statement 6 by Treatment Group

Treatment	LS Mean
MiC	0.13
MiC (Conventional)	0.08
Conventional	0.19

Table 3-23.
Contrasts in Means for Statement 6 by Treatment Group

Parameter	Estimate	Standard Error	t-Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	0.05	0.08	0.66	0.51	0.18	0.09
MiC vs Conventional	-0.06	0.09	-0.65	0.52	-0.18	0.09
MiC (Conventional) vs Conventional	-0.11	0.09	-1.24	0.22	-0.35	0.17

PA contributed also significantly to students' change in attitudes on Statement 16, "It's okay if I solve a math problem differently than my classmates do." The results are shown in Table 3-24.

Table 3-24.
ANCOVA for the Three Treatment Groups with Statement 16 Means as the Dependent Variable and Preceding Achievement as the Covariate

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.37	0.12	3.30	0.02
Error	62	2.31	0.04		
Corrected Total	65	2.68			

R-Square	Coeff Var	Root MSE	SAIGENPERC16CHG Mean
0.14	1168.39	0.19	0.02

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMENT	2	0.23	0.12	3.10	0.05
PA MEAN	1	0.14	0.14	3.66	0.06

The least-squares means for the three treatment groups on this item are shown in Table 3-25, and the contrasts between the groups are shown in Table 3-26.

Table 3-25.
Least-Squares Means for Statement 16 by Treatment Group

Treatment	LS Mean
MiC	0.06
MiC (Conventional)	-0.06
Conventional	0.07

Table 3-26.
Contrasts in Means for Statement 16 by Treatment Group

Parameter	Estimate	Standard Error	t-Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	0.12	0.06	2.16	0.03	0.58	0.28
MiC vs Conventional	-0.01	0.06	-0.16	0.87	0.05	0.02
MiC (Conventional) vs Conventional	-0.13	0.06	-2.09	0.04	-0.59	0.28

For Statement 16, the change in the MiC group's attitudes and the Conventional group's attitudes were significantly higher than the MiC (Conventional) group. However, there were no differences between the MiC and Conventional groups. This contrasts with the results at the beginning of the study that the combined MiC and MiC (Conventional) group agreed more strongly with the statement than the Conventional group. An example of the results over time is shown in Table 3-26.

Table 3-26.
Change in Response to Statement 16, Grade 5, District 1, 1997-1998

School-Class *(N)*	**Statement 16**				
	Prior		**End of Year**		**Gain**
	(N)	**Mean***	***(N)***	**Mean**	**Score****
—MiC—					
Beethoven-LaSalle	32	1.06	30	1.07	0.00
Beethoven-Linne	13	1.54	10	1.20	0.34
Dewey-Mitchell	54	1.28	52	1.08	0.20
—MiC (Conventional)—					
Banneker-Greene	17	1.24	17	1.06	0.18
Beethoven-Kipling	23	1.26	24	1.33	-0.07
Dewey-Hamilton	20	1.05	18	1.28	-0.23
—Conventional—					
Dewey-Kershaw	22	1.18	15	1.13	0.05
River Forest-Fulton	30	1.27	30	1.27	0.00

* Each item was judged on a scale of 1-4 (1 = Very True; 2 = True; 3 = Not True; 4 = Not True at All).

**Positive scores indicate more positive attitudes and negative scores indicate more negative attitudes

In this example, the MiC group and the Conventional group were significantly more positive than the MiC (Conventional) group that it was acceptable to solve mathematics problems differently than other students. In comparison, the MiC (Conventional) group tended to be more negative over time with respect to this statement.

PA contributed to an important difference in students' change in attitudes with respect to Statement 4, "In mathematics, you can discover new ways of solving problems that the teacher or your classmates may not have thought of." The results are shown in Table 3-27.

Table 3-27.
ANCOVA for the Three Treatment Groups with Statement 4 Means as the Dependent Variable and Preceding Achievement as the Covariate

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.66	0.22	2.27	0.09
Error	62	6.04	0.10		
Corrected Total	65	6.71			

	R-Square	Coeff Var	Root MSE	SAIGENPERC4CHG Mean
	0.10	-700.88	0.31	-0.04

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TREATMENT	2	0.28	0.14	1.43	0.25
PA MEAN	1	0.37	0.37	3.77	0.06

However, there were no differences between treatment groups as shown in Tables 3-28 and 3-29.

Table 3-28.
Least-Squares Means for Statement 4 by Treatment Group

Treatment	LS Mean
MiC	0.00
MiC (Conventional)	-0.13
Conventional	0.00

Table 3-29.
Contrasts in Means for Statement 4 by Treatment Group

Parameter	Estimate	Standard Error	t-Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	0.14	0.09	1.54	0.13	0.41	0.20
MiC vs Conventional	0.00	0.10	0.03	0.98	0.01	0.00
MiC (Conventional) vs Conventional	-0.13	0.10	-1.34	0.18	-0.38	0.19

Differences between the treatment groups were also found with respect to Statement 39, "Each new math topic I study is not related to ones I have learned before." The least-squares means for the three treatment groups on this item are shown in Table 3-30, and the contrasts between the groups are shown in Table 3-31.

Table 3-30.
Least-Squares Means for Statement 39 by Treatment Group

Treatment	LS Mean
MiC	0.05
MiC (Conventional)	0.26
Conventional	0.12

Table 3-31.
Contrasts in Means for Statement 39 by Treatment Group

Parameter	Estimate	Standard Error	t-Value	Pr > |t|	Cohen's d	Effect Size r
MiC vs MiC (Conventional)	-0.20	0.11	-1.86	0.07	-0.50	0.24
MiC vs Conventional	-0.07	0.12	-0.57	0.57	-0.16	0.09
MiC (Conventional) vs Conventional	0.13	0.12	1.10	0.27	0.31	0.15

For Statement 39, there was an important difference between the MiC and MiC (Conventional) groups' attitudes in favor of the MiC group. There were no differences between the MiC and Conventional groups or between the MiC (Conventional) and Conventional groups.

Three other findings are of interest in the overall data set. First, attitudes of study students tended to be more negative as they progressed from fifth grade through middle school, which is consistent with previous research (see Monograph #2, Chapter 1). Second, some MiC students who experienced high quality instruction and opportunity to learn with understanding tended to be more negative with respect to the effort needed to succeed in mathematics and their confidence in learning mathematics. These perceptions may be a result of the high expectations these teachers held for their students. Third, MiC (Conventional) students who experienced a low quality of instruction and opportunity to learn with understanding tended to be more negative than other groups, particularly with respect to confidence in learning mathematics, interest in and excitement about mathematics, and usefulness of mathematics.

In summary, preceding achievement did not contribute to the variance in any of the first five subscales, and there were no differences between treatment groups on these subscales. Of the 16 statements in the general perceptions subscale, preceding achievement contributed significantly to students' change in attitudes on two statements, but differences in treatment were found on only one statement for which the change in the MiC group's attitudes and the Conventional group's attitudes were significantly higher than the MiC (Conventional) group. Prior achievement contributed to an important difference in students' change in attitudes on one statement, but no differences in treatment were found. Only one other important difference was noted between the treatment groups, and the difference favored the MiC group over the MiC (Conventional) group.

Differences in Change in Mathematical Reasoning

The fourth dependent variable involves the changes in the *Collis-Romberg Mathematical Problem Solving Profiles* (Collis & Romberg, 1992) for the three treatment groups. Form A of the *Collis-Romberg Profiles* was used as a pre-test in Year 1. In the final spring of each student's participation in the study, Form B was administered as a post-test in the study. (See Chapter 1 in Monograph #2 for a discussion of this instrument and its use in the study.) The *Collis-Romberg Profiles* is a set of mathematical superitems designed to provide information about students' qualitatively different levels of reasoning ability. A "superitem" is a set of test questions based on a common situation or stem (Cureton, 1965). The *Collis-Romberg Profiles* contain five mathematical problem solving situations and four questions for each situation which were based on Collis and Biggs' (1979) SOLO taxonomy used to classify the structure of observed learning outcomes. Each question was designed to require more sophisticated use of the information from the stem in order to obtain a correct result. This increase in sophistication should parallel the increasing complexity of structure noted in the SOLO categories. A correct response to each question would be indicative of an ability to respond to the information in the stem at least at the level reflected in the SOLO structure of the particular question. For each category, the question was as follows:

- Uni-structural (U) — Use of <u>one obvious</u> piece of information coming directly from the stem.
- Multi-structural (M) — Use of two or more discrete closures directly related to <u>separate pieces</u> of information contained in the stem.
- Relational (R) — Use of two or more closures directly related to an integrated understanding of the information in the stem.
- Extended Abstract (E) — Use of an abstract general principle or hypothesis, which is derived from or suggested by the information in the stem.

In each superitem, the correct achievement of Question 1 would indicate an ability to respond to the problem concerned at least the uni-structural level. Similarly, success on Question 2 corresponds to an ability to respond at multi-structural level, and so on.

For this longitudinal analysis, it was impractical to separate the students who had been in MiC and MiC (Conventional) classes as was done in the previous analyses in this monograph because many students were in both types of treatments over the two years in the study. Also, for this longitudinal analysis, the student is the unit of analysis rather than teacher/student group. Furthermore, while the information about achievement on Form A of *Collis-Romberg Profiles* was available for each of the seven cohort groups discussed in Monograph #6 Chapter 4, because Form B was only given at the end of the study, data were available for just three cohort groups AD (Grade 6 through Grade 7), BE (Grade 7 through Grade 8), and C (Grade 7 through Grade 8).

Cohort AD

Cohort AD included students who completed both study assessments (External Assessment System [EAS] and Problem Solving Assessment [PSA]) in both Grades 6 and 7. The percent of students in the combined MiC and MiC (Conventional) group and the Conventional group whose profiles indicated a particular level of reasoning is shown in Table 3-32**.** Students in the MiC group show a reasonable shift in level of reasoning from Form A to Form B. There is a considerable reduction of percent at the uni-structural level with a corresponding increase in percent at the multi-structural level. On the other hand, only a few students in the Conventional group show a very small shift in the same direction as the MiC students and a lowering of the percent of students in the pre-structural level of reasoning.

Table 3-32.
Student Performance Level on the Collis-Romberg Mathematical Problem-Solving Profiles-A and B for Students in Cohort AD by Treatment

Cohort AD	**Level of Student Performance**						**Not available**
	(N)	**Prestructural**	**Unistructural**	**Multistructural**	**Relational**	**Extended Abstract**	
			Mathematics in Context				
CR-A	109	15%	55%	17%	1%	0%	12%
CR-B	109	13%	25%	44%	0%	0%	18%
			Conventional				
CR-A	16	25%	44%	0%	0%	0%	31%
CR-B	16	6%	38%	6%	6%	0%	44%

Cohort BE

Cohort BE included students who completed both study assessments (EAS and PSA) in both Grades 7 and 8. The percent of students in the combined MiC and MiC (Conventional) group and the Conventional group whose profiles indicated a particular level of reasoning is shown in Table 3-33**.** Students in the MiC group show a reasonable shift in level of reasoning from Form A to Form B. There is a reduction of percent at the uni-structural level with a substantial corresponding increase in percent at the multi-structural level. Students in the Conventional group had a reasonable number of students at the multi-structural on Form A. On the other hand, because of lack of cooperation of teachers in the Conventional group, no data exist for these students on Form B, making comparisons impossible.

Table 3-33.
Student Performance Level on the Collis-Romberg Mathematical Problem-Solving Profiles-A and B for Students in Cohort BE by Treatment

Cohort BE	Level of Student Performance						Not available
	(N)	**Prestructural**	**Unistructural**	**Multistructural**	**Relational**	**Extended Abstract**	
Mathematics in Context							
CR-A	137	17%	46%	7%	0%	0%	31%
CR-B	137	11%	31%	31%	1%	0%	26%
Conventional							
CR-A	27	11%	41%	37%	11%	0%	0%
CR-B	27	-	-	-	-	-	100%

Cohort C

Cohort C included students who completed both study assessments (EAS and PSA) in both Grades 7 and 8. The percent of students in the combined MiC and MiC (Conventional) group and the Conventional group whose profiles indicated a particular level of reasoning is shown in Table 3-34. Students in the MiC group show a reasonable shift in level of reasoning from Form A to Form B. There is a considerable reduction of percent at the uni-structural level with a corresponding increase in percent at the multi-structural level. Students in the Conventional group show a similar shift in the same direction as the MiC students.

Table 3-34.
Student Performance Level on the Collis-Romberg Mathematical Problem-Solving Profiles-A and B for Students in Cohort C by Treatment

Cohort C	Level of Student Performance						Not available
	(N)	Prestructural	Unistructural	Multistructural	Relational	Extended Abstract	
Mathematics in Context							
CR-A	67	16%	78%	4%	1%	0%	0%
CR-B	67	13%	49%	34%	0%	0%	3%
Conventional							
CR-A	34	18%	65%	12%	3%	0%	3%
CR-B	34	15%	44%	35%	0%	0%	6%

In summary, for MiC groups in all three cohorts, a similar shift in the pattern of a considerable percent reduction in the number of students at the uni-structural level of reasoning and a corresponding increase in students at the multi-structural levels. For the Conventional students in Cohort C, a similar shift is apparent. In Cohort AD, there is a very small shift in the same percentages. Unfortunately, no data for the Conventional students in Group BE exist. Also, we were surprised at the percent of both MiC and Conventional students classified as pre-structural in all three cohorts. This implies that 10-20% of these students were unable to answer the majority of first questions for each stem. Similarly, the nearly 0% of students at the relational level at the end of the study is worrisome, for it is assumed that students should be at that level if they are to take a formal algebra course.

Conclusion

In this chapter, the differences between the three treatment groups in Districts 1 and 2 on four other dependent variables were summarized. Analysis of differences in standardized test percentile scores showed that the prior test mean scores strongly predict the post mean scores, and no significant differences were noted between the three treatment groups. With respect to the set of twenty anchor items on the External Assessment System designed for the study, differences between the groups were evident on only eight items. The performance of the MiC group was significantly higher than that of the Conventional group on four items. Analysis of differences in student attitudes revealed no significant differences between the three treatment groups on five subscales. In the general perceptions subscale, only one change in attitudes was significant. The MiC group was significantly more positive than the MiC (Conventional) group with respect to one statement: “It’s okay if I solve a math problem differently than my classmates do.” The analysis of change on the *Collis-Romberg Mathematical Problem Solving Profiles* indicated that a considerable shift from the uni-structural to the multi-structural level occurred for three longitudinal cohorts of MiC students and for the one available Conventional cohort, but few students advanced to the relational level. Thus, across these four dependent variables, few differences were noted between the treatment groups.

References

Charles, R. I., et al. (1998). *Middle school mathematics.* Menlo Park, CA: Scott Foresman/Addison Wesley.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

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

Collis, K. F., & Biggs, J. B. (1979). *Classroom examples of cognitive development phenomena:The SOLO taxonomy*. Report prepared at the conclusion of an Educational Research and Development Committee funded project, University of Tasmania.

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

CTB/McGraw-Hill (1997). *TerraNova.* Monterey, CA: Author.

Cureton, E. F. (1965). Reliability and validity: Basic assumptions and experimental designs. *Educational and Psychological Measurement, 25,* 326-346.

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.

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

Harcourt Brace Educational Measurement. (1997). *Stanford mathematics achievement test (SAT), 9th edition.* San Antonio, TX: Harcourt Brace.

National Center for Research in Mathematical Sciences Education & Freudenthal Institute (Eds.). (1997-98). *Mathematics in context*. Chicago: Encyclopaedia Britannica.

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

Newmann, F. M., King, M. B., Youngs, P. (Draft, May, 2000). Professional development that addresses school capacity: Lessons from urban elementary schools. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

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

Romberg, T. A. (1997). Mathematics in context: Impact on teachers. In E. Fennema & B. S. Nelson (Eds.), *Teachers in transition* (pp. 357–380). Mahwah, NJ: Erlbaum.

Romberg, T. A., Shafer, M. C., LeMire, S., & Folgert, L. (2005) *Covariant and regression analyses of Mathematics in Context Longitudinal/Cross-Sectional Study data.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Tech. Rep No. 54). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

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

Romberg, T. A., Webb, D. C., & Folgert, L. (2004) *Standardized test national percents for the eight grade-level-by-year studies in 1997-1998, 1998-1999, 1999-2000.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Tech. Rep No. 46). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Romberg, T. A., Webb, D. C., Folgert, L., & Shafer, M. C. (2003a) *Classroom achievement for Mathematics in Context classrooms for 1997-1998.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Tech. Rep No. 33). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Romberg, T. A., Webb, D. C., Folgert, L. & Shafer, M. C. (2003b) *Classroom achievement for Mathematics in Context classrooms for 1998-1999.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Tech. Rep No. 34). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Romberg, T. A., Webb, D. C., Folgert, L. & Shafer, M. C. (2003c) *Classroom achievement for Mathematics in Context classrooms for 1999-2000.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Tech. Rep No. 35). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Romberg, T. A., Webb, D. C., Folgert, L. & Shafer, M. C. (2003d) *Classroom achievement for conventional classrooms for 1997-1998.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Tech. Rep No. 36). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Romberg, T. A., Webb, D. C., Folgert, L. & Shafer, M. C. (2003e) *Classroom achievement for conventional classrooms for 1998-1999.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Tech. Rep No. 37). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Romberg, T. A., Webb, D. C., Folgert, L. & Shafer, M. C. (2003f) *Classroom achievement for conventional classrooms for 1999-2000.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Tech. Rep No. 38). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

SAS Institute. (2000). *SAS, Ninth Edition.* Cary, NC: Author.

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

Shafer, M. C., Folgert, L., & Kwako, J. (2004a) *Opportunity to learn with understanding for 1998-1999.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Tech. Rep No. 42). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C., Folgert, L., & Kwako, J. (2004b) *Opportunity to learn with understanding for 1999-2000.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Tech. Rep No. 43). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C., Folgert, L., Wagner, L., & Kwako, J. (2004) *Opportunity to learn with understanding for 1997-1998.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Tech. Rep No. 41). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C., Folgert, L., Webb, D., & Kwako, J. (2003a) *School capacity for 1998-1999.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Tech. Rep No. 28). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C., Folgert, L., Webb, D. C., & Kwako, J. (2003b) *School capacity for 1999-2000*. (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 29). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C., Folgert, L., Webb, D. C., Kwako, J., Lee, C., & Wagner, L. (2003) *School capacity for 1997-1998.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 27). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C., Marten, B., Folgert, L., & Kwako, J. (2003a) *Instruction for1998-1999.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Tech. Rep No. 31). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C., Marten, B., Folgert, L., & Kwako, J. (2003b) *Instruction for 1999-2000.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Tech. Rep No. 32). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Shafer, M. C., Marten, B., Webb, D. C., Folgert, L., & Davis, J. (2003) *Instruction for 1997-1998.* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 30). 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.