Delaware Education R&D Research & Development CENTER

MATH PARTNERSHIP PROJECT:

ANNUAL EVALUATION REPORT – YEAR 4

December 2008

Report prepared for:
Mathematics and Science Education Resource Center

Recommended citation: Ackerman, C. M, & Uribe-Zarain, X. (2008). *Math partnership project evaluation* [Technical Report No. T2008.10.01]. Newark, DE: University of Delaware Education Research and Development Center.

Acknowledgements: The conduct of these evaluation activities would not have been possible without the cooperation of participating teachers, as well as the R&D Center staff members who conducted the observations.

Copyright © 2008 by the University of Delaware

TABLE OF CONTENTS

Abstract ii

Executive Summary iii

Introduction 1

Methodology 2

Results 3

Discussion 10

ABSTRACT

Through federal funding of a National Science Foundation grant, the Math and Science Education Resource Center has been implementing Delaware Secondary Mathematics Partnership Project, the goal of which is to help at-risk students by improving the instructional techniques of their teachers through a video-taping and critique process. Observations of teachers participating in the Math Partnership Project were conducted to determine how teaching techniques change over the school year. Staff from the Delaware Education Research and Development Center conducted 38 observations of math lessons during fall 2007 and 34 observations during spring 2008. Thirty teachers observed in spring were also observed in fall.

The observations conducted included three main components: The design and implementation of the lesson, mathematics content, and elements of classroom culture. Results indicate improvement in some areas of instruction from fall to spring, and a decrease in performance in other areas. The highest overall performance is the classroom culture component. The area needing the greatest improvement is mathematics content.

Executive Summary

Through federal funding of a National Science Foundation grant, the Math and Science Education Resource Center has been implementing Delaware Secondary Mathematics Partnership Project, the goal of which is to help at-risk students by improving the teaching techniques of their teachers.

Observations of teachers participating in the Math Partnership Project were conducted to determine how teaching techniques change over the school year. Staff from the Delaware Education Research and Development Center conducted 38 observations of math lessons during fall 2007 and 34 observations during spring 2008. Thirty teachers observed in spring were also observed in fall.

The observations conducted included three main components: The design and implementation of the lesson, mathematics content, and elements of classroom culture. The data gathered indicate improvement in some areas of instruction from fall to spring, and a decrease in performance in other areas. The highest overall performance is the classroom culture component. The area needing the greatest improvement is mathematics content. Specific details of strong and weak areas for each instructional component are provided below.[1] Performance that declined in the spring is indicated with an asterisk.

Design and Implementation

Strong performance areas:

- defined and communicated the mathematical purpose of the lesson*
- effectively engages students with important ideas related to the focus of the lesson
- provides adequate time and structure for "wrap-up"
- assessed students' level of understanding to accommodate their teaching*
- took into consideration prior experiences, how prepared the students were, and how teachers adjusted to different students learning styles

Weak performance areas:

- provided adequate time for students to engage in problem solving activities
- facilitated a collaborative approach to learning
- questioning strategies enhance the development of student conceptual understanding and sense-making

Mathematics Content

Strong performance areas:

- made the content challenging and accessible to students
- provided accurate content information during the observed lessons

Weak performance areas:

- balanced content between conceptual understanding and procedural fluency*
- elements of mathematical abstraction are included when appropriate to do so*

[1] Strong performance is defined as at least 75% of teachers performing "close to ideal" or "getting there," and weak performance was defined as at least 25% of teachers performing at the lowest level.

- made appropriate connections to other mathematics and/or real world content*

Classroom Culture

Strong performance areas:

- expected and value active participation of all students
- showed respect for ideas, questions, and contributions
- enhanced productivity through classroom procedures
- promoted the production of ideas and questions from students as they solved problems

Weak performance areas:

- intellectual rigor and/or constructive challenging of ideas are evident

The full report provides a detailed account of the year 4 evaluation results. Evaluators at the Delaware Education Research and Development Center (R&D Center) are available to answer questions regarding analyses presented in this report or to assist in their interpretation. R&D Center staff may be contacted via electronic mail at ud-rdc@udel.edu or by phone at (302) 831-4433

Introduction

This evaluation report, prepared by the Delaware Education Research and Development Center, includes a description of the performance of a group of mathematics teachers who participated in the Secondary Mathematics Partnership Project in the school year 2007-2008 in the state of Delaware. This is the fourth year of the evaluation.

This report includes four sections: The first section describes briefly the project and gives an overview of its main domains. The second section includes the methodology of the evaluation. The results of the fall and spring observations are described in the third section. Finally, a summary is presented in the fourth section.

The Delaware Secondary Math Partnership Project targets at-risk math students in grades six to 10. High-school and middle-school teachers videotape each other during class instruction and then watch and critique their own and others' techniques. The main goal of the Delaware Secondary School Math Partnership is to help students through observation, considering which kind of instruction is reaching them and which is not. The three main components of the projects are:

1. The *Design and Implementation of the Lesson* which encompasses a range of factors including communication of purpose, effective allocation of time to critical lesson components, and effective questioning and formative assessment technique;

2. *Mathematics Content* which addresses both rigor and appropriateness of the mathematics, assessing level of challenge *and* accessibility. Elements of mathematical abstraction, connections within mathematics and between mathematics and the phenomena it represents are elements of interest;

3. Elements of *Classroom Culture* which include factors that are believed to enhance effective mathematics discourse including high expectations for all students and a privileging of mathematical argumentation.

Methodology

An observation protocol was developed by University educators from the Mathematics & Science Education Resource Center in conjunction with researchers from the Delaware Education Research and Development Center. The observation protocol called "Determining the Quality of Mathematics Instruction" was adopted as the main measure of teaching quality. The protocol consists of the three main components in which the Secondary Mathematics Partnership Project is interested: The design and implementation of the lesson, mathematics content, and classroom culture. The items or questions for each of the components are as follows:

The design and implementation of the lesson:
1. Teacher clearly defines and communicates a purpose of the lesson.
2. Teacher effectively engages students with important ideas.
3. Teacher provides adequate time and structure for investigation and exploration.
4. Teacher provides adequate time and structure for "wrap-up."
5. Teacher achieves a collaborative approach to learning.
6. Teacher enhances the development of student understanding.
7. Teacher assesses the students' level of understanding.
8. Teacher plans and/or adjusts instruction based on students' level of understanding.

Mathematics content:
1. The content is balanced between conceptual understanding and procedural fluency.
2. The content is challenging and accessible to the students.
3. Teacher provides content information that is accurate.
4. Elements of mathematical abstraction are included when appropriate to do so.
5. Appropriate connections are made to other mathematics and/or to real world content.

Classroom culture:
1. Active participation of ALL is expected and valued.
2. There is a climate of respect for students' ideas, questions, and contributions.
3. Teacher's classroom management style/strategies enhance productivity.
4. The classroom climate is encouraging to students
5. Intellectual rigor and/or the constructive challenge of ideas are evident.

Using the "Determining the Quality of Mathematics Instruction" protocol a group of observers was trained until they achieved an adequate inter-rater reliability. During observations, questions were answered using three principal descriptors, "close to ideal," "getting there," and "not even close." While these concepts are illustrated through examples within the context of each of the

separate indicators, it is possible to characterize them in more general terms. An indicator is rated as "close to ideal" if there is a good bit of strong supporting and little or no contradictory evidence. "Getting there" suggests a convergence on exemplary practice but also an incomplete realization thereof. Practices that are clearly at odds with the ideal within an indicator may still be present but no longer represent the norm. Teaching that is rated as "not even close," however, is consistently impoverished with little indication of progress toward the exemplary.

In October of 2007 and May of 2008, as part of the fourth year of the Secondary Mathematics Partnership Project evaluation, observers were sent into math classrooms to gather data about math instruction across the state. We gathered data from 38 teachers in October and 34 teachers in May, and 30 teachers were observed both times. The lessons observed occurred in sixth to eleventh grade classrooms, and ranged from 20 to 107 minutes in length (average = 65 minutes). There were between seven and 32 students in each classroom with an average of about 20 students. Observers looked for specific evidence regarding the three main components: lesson design and implementation, math content, and classroom culture.

Results

The results of the fall and spring observations are presented in this section. Percentages of teachers rated in each category: "close to ideal," "getting there," and "not even close" as well as instances where teachers were rated in the middle of the categories (e.g. in between "close to ideal" and "getting there") are represented in graphs throughout this section. The graphs portray fall and spring observations side by side for each question of the three components. In addition to figural displays of performance, tables summarizing performance are included for each survey component showing which items performed well or poorly in spring, and which items improved or worsened over time. It is desirable that the category "close to ideal" increases from the first to the second observation and that the category "not even close" would also decrease, however, there are instances

when performance can be considered good in the spring, despite being worse than fall performance. There are also instances when performance remains poor, but has shown improvement over time.

The design and implementation of the lesson

The first domain, pertaining to the design and implementation of the lesson, is described in Table 1 for spring and represented in Figure 1 for fall and spring. Five of the eight questions show good performance, defined as having at least 75% of teachers being rated as "getting there" or "close to ideal" in the spring (Questions 1, 2, 4, 7, 8). Three of these had at least 50% of teachers rated "close to ideal" (Questions 1, 2, 8). Two of the five questions showing good performance showed improvement from fall to spring, one showed poorer performance, one showed improvement at one end of the spectrum and poorer performance at the other end (polarization), and one showed no change.

Table 1. Performance summary for design and implementation questions

Performance	**Improving** (more "ideal" or less "not even close")	**Worsening** (less 'ideal' or more "not even close")	**Remaining the Same**
Good (at least **75%** 'ideal' or 'getting there' in spring)	2*, 4, 7	1*, 7	8*
Poor (at least **25%** "not even close" in spring)	5, 6, 3	-	-

* At least 50% "close to ideal"

Performance that was good in the spring and had shown improvement from the fall was evidenced by the success of the majority of teachers at engaging students with important ideas related to the focus of the lesson. Also, there were fewer teachers with a "not even close" rating and more than half of teachers rated "close to ideal" (Question 2). In addition, while fewer than half of teachers were rated "close to ideal" at providing adequate time and structure for "wrap-up" in

spring, there was a decrease from 30% to 14% for teachers rated "not even close" by spring (Question 4). However, similar to last year (Uribe-Zarain, 2007)[2], there is still room for improvement in this area, especially since "summing it up" represents an opportunity for the teacher to orchestrate the presentation of student ideas that have resulted from the exploration phase and, in so doing, draw out the important mathematical ideas from the lesson. This is where mathematical connections are often made and may be a final opportunity for the teacher to assess the impact of the day's lesson.

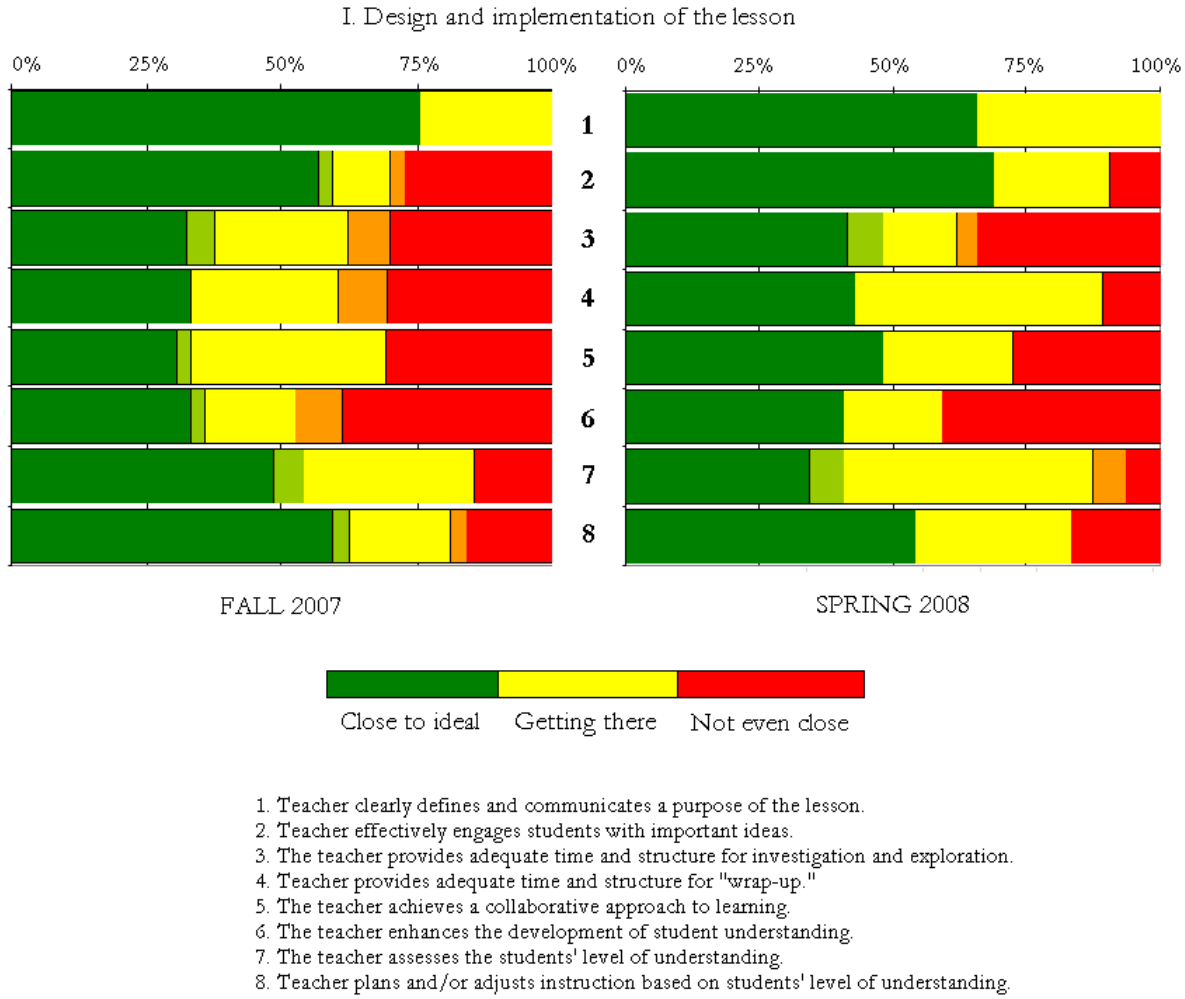

FIGURE 1. Design and implementation of the lesson FALL and SPRING

[2] Uribe-Zarain, X. (2007). *Math partnership project evaluation – Year 3.* [Technical Report No. T2007.27.01]. Newark, DE: University of Delaware Education Research and Development Center.

Like Question 4, Question 7 showed a high level of performance in spring with fewer teachers performing at the lowest level compared to fall performance. However, there was also a decrease in the percentage of teachers performing at the highest level at less than half by spring. This indicates that many teachers are "getting there," but have not mastered the ability to assess students' level of understanding. Therefore, teachers will find it difficult to accommodate students' prior experiences and appropriately adjust to their teaching in the future.

When defining and communicating the mathematical purpose of the lesson, all the teachers observed were "getting there" or "close to ideal" in spring, however, the percentage of teachers at the highest level decreased from the fall (Question 1). Additionally, there was no change in teacher performance regarding how they adjust their lessons to the students' level of understanding from fall to spring (Question 8). For both observations, more than half of the teachers were "close to ideal."

Three of the eight questions associated with lesson design and implementation show poor performance, defined as having at least 25% of teachers being rated as "not even close" in the spring (Questions 3, 5, 6), all of which showed improvement from the fall. Greater than one-fourth of teachers were "not even close" to stimulating investigation and exploration by providing adequate structure and enough time in their lessons in fall and spring, though there was a small increase in the percentage of teachers performing at the highest level in the spring (Question 3). Similar to Question 3, more than 25% of teachers had not shown a collaborative approach to learning in the spring, there was also an increase in the percentage of teachers performing at the "close to ideal" level (Question 5).

The last area of poor performance that showed some improvement over time, was enhancing the development of students' understanding (Question 6). However, there is still much room for improvement in this area. Research has found that higher order questions are correlated with increased student achievement, particularly for conceptual understanding (see Redfield &

Rousseau, 1981). Higher order questions offer opportunities to learn and think about the mathematical ideas in many different ways.

Mathematics content

Items regarding mathematics content of the lesson are represented in Figure 2 and Table 2. The questions fall into two categories: There were two questions with good performance in spring showing improvement from the fall, and three questions with poor performance in the spring showing worse ratings than the fall.

Nearly all, if not all, teachers are performing at the highest two levels regarding how challenging and accessible the content is to students (Question 2) and providing accurate mathematical content during instruction (Question 3) in spring. In fact, greater than three-quarters of teachers were rated "close to ideal" on their content accuracy. There are also more teachers performing at the highest level in spring.

Table 2. Performance summary for mathematics content

Performance	**Improving** (more "ideal" or less "not even close")	**Worsening** (less 'ideal' or more "not even close")	**Remaining the Same**
Good (at least **75%** 'ideal' or 'getting there' in spring)	2, 3*	-	-
Poor (at least **25%** "not even close" in spring)	-	1**, 4, 5**	-

* 75% "close to ideal" in spring
** At least 50% "not even close" in spring

The remaining three questions related to mathematics content showed very poor performance in the spring that was worse than fall performance. There is a clear need for improvement in these areas. First, more than half of the teachers have not master the practice of balancing content between conceptual understanding and procedural fluency in the spring, and are

emphasizing procedural skills (Question 1). Similarly poor performance was observed when nearly half of the teachers failed to include elements of mathematical abstraction when appropriate in the spring (Question 4). The final area of math content needing improvement is making appropriate connections to other mathematics and/or real world content (Question 5). Half of teachers observed performed at the lowest level in the spring, a greater percentage than the fall.

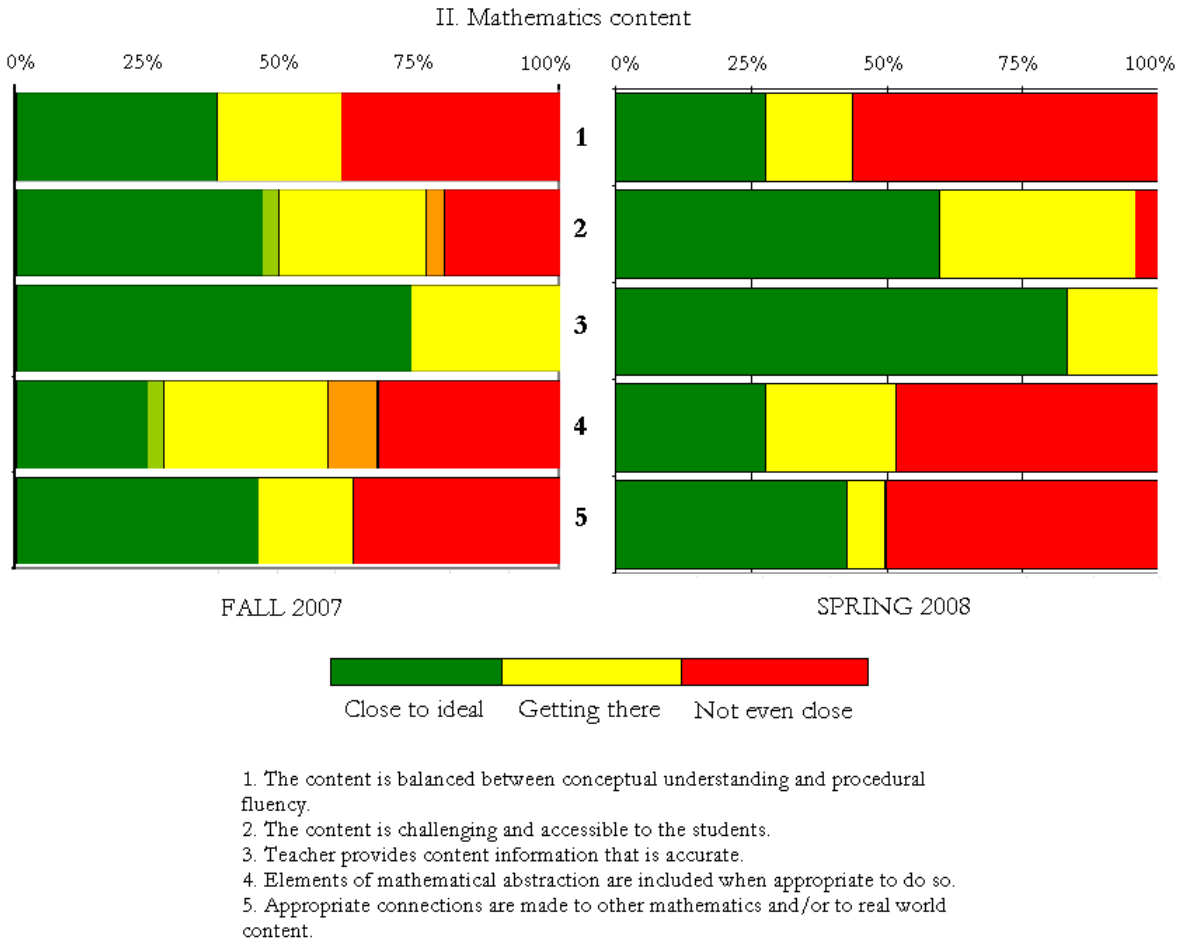

FIGURE 2. Mathematics content FALL and SPRING

Classroom culture

Of the three survey components, Classroom Culture is the strongest needing the least amount of improvement (see Figure 3 and Table 3). Four of the five questions show good performance in spring that improved since the fall.

Table 3. Performance summary for classroom culture

Performance	**Improving** (more "ideal" or less "not even close")	**Worsening** (less 'ideal' or more "not even close")	**Remaining the Same**
Good (at least **75%** "ideal" or "getting there" in spring)	1, 2*, 3*, 4	-	-
Poor (at least **25%** "not even close" in spring)	-	-	5

* 75% "close to ideal" in spring

It is apparent that active participation of all students is almost always expected or valued; a small percentage of the teachers were rated "not even close" on this statement in the spring and the percentage of teachers rated "close to ideal" increased from the fall (Question 1). Also, most lessons showed a climate of respect for students' ideas, questions and contributions with very few rated "not even close" in the spring (Question 2). The question showing the greatest improvement relates to classroom management style with the percentage of "close to ideal" ratings increasing 36% from fall to spring resulting in 97% of teachers at the top level of performance (Question 3). Teacher support for generating ideas and questions from students as they solved problems also showed good performance in the spring, with an increase in the percentage of teachers performing at "close to ideal" (Question 4).

Only one area in the *Classroom Culture* section showed poor performance. Greater than 25% of lessons were given the lowest rating for intellectual rigor or constructive challenge of ideas

(Question 5). This item also had the lowest percentage of teachers rated "close to ideal" across the entire observation protocol (25%). In other words, students' conjectures were explored and students were held to the standard of justification and proof in very few lessons in the fall and spring.

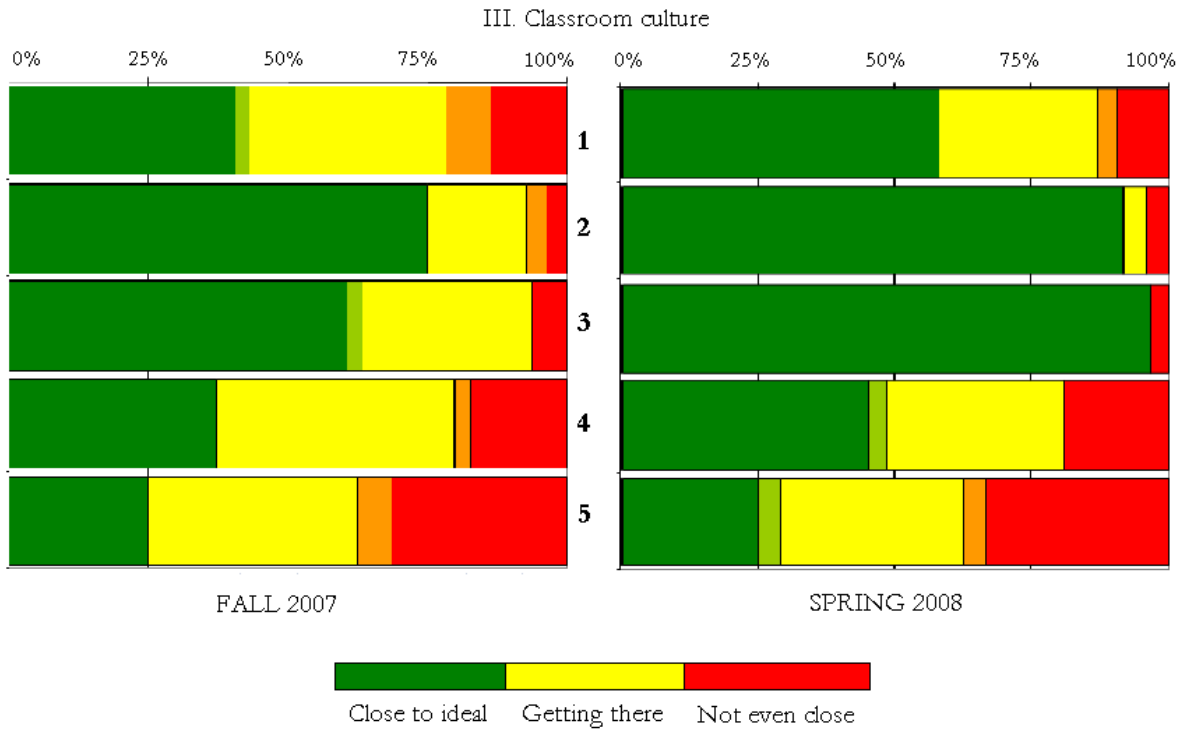

FIGURE 3. Classroom culture FALL and SPRING

Discussion

It is clear that the teachers participating in the Math Partnership program have mastered some of the program components, while others continue to elude them. Several aspects across the

three major components show a high level of performance and improvement over time for most teachers in the program.

- Design and implementation – Questions 2 & 4
- Mathematics content – Questions 2 & 3
- Classroom culture – Questions 1, 2, 3, & 4

However, there are several other aspects across components showing poor performance with a decrease over time. Clearly, these areas need to be addressed with participating teachers before improvement can be seen in the classroom.

- Design and implementation – Questions 1 & 7
- Mathematics content – Questions 1, 4, & 5

In addition, while no formal longitudinal comparison has been conducted, it appears that there are some consistent patterns of performance over time: Some components are consistently strong while others are weak, and some regularly show improvement in the spring while others show declining performance in the spring. A close examination of these patters and trends could help inform the program leadership as to how professional development should proceed in the future.

Because some aspects of lesson design and implementation, as well as mathematical content, appear to be more problematic than others, and more problematic than classroom culture as a whole, in-depth exploration of teachers' impressions of the program could shed light on the difficulties they experience. The added interview evaluation component for 2009 should provide insight into key issues in program implementation and could, as a longitudinal analysis, help guide professional development for the program in the future.

While this fourth year evaluation has operationalized "good" and "bad" performance, no benchmarks or guiding cutoffs have been described elsewhere thus far. Therefore, it is impossible

to make a final determination of the degree to which the Math Partnership Program is meeting its goals. For the final year's program evaluation, it may be useful to clarify this point.