# Seeing Math Elementary

# Final Evaluation Report

# September 2005

**A Report Prepared for**
**Concord Consortium**
**And**
**U.S. Department of Education**

edcentric

7003 Waxmyrtle Court
Wilmington, NC 28409-4852
910.796.0632

edcentric@edcentric.org

# TABLE OF CONTENTS

# EXECUTIVE SUMMARY

Seeing Math Elementary is a video-case based mathematics professional development program for elementary and middle school teachers created by the Concord Consortium and Teachscape. It is funded under the U.S. Department of Education's Ready To Teach Program (grant # R286A000006-03) as part of the Seeing Math Telecommunications Project. Seeing Math Elementary's professional development integrates web-based learning with face-to-face interaction. The overall goal of Seeing Math Elementary is "To improve student achievement by developing high quality, standards-based digital professional development to teachers and by developing high quality, standards-based digital classroom content."

Each facilitated course focuses on a concept from the National Council of Teachers of Mathematics (NCTM) standards that is typically considered difficult to teach or learn, such as fractions, division with remainders, calculating the area of a triangle, pre-algebra, or using data to make predictions.

In the past several years, the U.S. Department of Education has established new priorities for evaluation, with a strong emphasis on experimental and quasi-experimental research. This has resulted in Concord Consortium's development of a comprehensive research plan for the overall Seeing Math Elementary project beginning with the 2003 – 2004 school year and extending through the 2004 – 2005 school year. This final evaluation report for the project focuses on the research and evaluation efforts conducted by Edcentric, the external evaluator. Its primary emphasis is on the findings of the 2004 – 2005 school year, in which experimental research is conducted with both treatment and control groups. Findings from the 2003 – 2004 school year are conflated into this report to provide a complete picture of the overall research within the design and to assess potential student growth over time through a longitudinal study based on data collected both years.

The evaluation report focuses on the following two objectives of the Seeing Math Telecommunications Project produced by Concord Consortium:

- Evaluate the quality and usefulness of project materials and strategies.
- Research the effect and impact of the Seeing Math materials.

## QUALITY AND USEFULNESS OF PROJECT MATERIALS AND STRATEGIES:

The formative evaluation focused on the appeal, functionality, usefulness of the course materials and facilitation, and related changes in classroom practice related to participation.

During the 2004 – 2005 school year, fifty-four teachers and teacher leaders pilot tested and completed formative evaluations of these Seeing Math Elementary courses: *Division* *with Remainders, Pan Balance Equations, Broken Calculator, Calculating the Area of a Triangle, Using Data to Make Predictions, Formative Assessment in the Mathematics Classroom, Effective Questioning in the Mathematics Classroom,* and *2D and 3D Figures*. The courses met or exceeded participants' expectations and these elements of the courses – video case study, expert commentary, diving in activity, online activities, additional resources, and face to face meetings and discussion – were considered important by the majority of the respondents.

The 54 teachers taught 1,333 students during the year, with 1,089 in classes where they believed they used the new insights/strategies they learned concomitant with participation in Seeing Math Elementary. They provided examples of both teacher and student changes to support their beliefs that new understandings were engendered through their Seeing Math Elementary involvement.

## EFFECT AND IMPACT OF THE SEEING MATH ELEMENTARY MATERIALS:

The quantitative analysis, based on a quasi-experimental design with treatment and comparison groups, looked at teacher and student outcomes related to treatment teachers' enrollment in Seeing Math Elementary in the 2004 – 2005 school year. Longitudinal analysis compared student gains in the 2003 – 2004 and 2004 – 2005 school years for treatment group teachers, referencing the time frame in which teachers took Seeing Math Elementary courses: spring 2004, fall 2004, or spring 2005. Four school districts were involved in the study and their results are analyzed individually.

Assessments included teacher and student tests created specifically for the project. Teacher tests were open-ended measures developed to measure teachers' math content and pedagogy knowledge, taken before and after participation in a Seeing Math Elementary course. Student tests, 20-item multiple choice tests, measured math knowledge related to the National Council of Teachers of Mathematics standard that was directly related to the Seeing Math Elementary courses taken by participating teachers.

Treatment teachers performed significantly better in the content areas of Modeling/Formulating, Transforming/Manipulating, Inferring/Drawing Conclusions, and Communicating than did the control teachers. Additionally, they performed significantly better than the control teachers in Pedagogy ratings. Significant differences between gains for treatment and control teachers offer evidence that participation in Seeing Math Elementary can provide teachers with relevant new learnings in a number of different content and pedagogic areas.

Student findings were more difficult to interpret. Students in the one district following the Pre-Algebra strand had significant gains on total score and both sub-scores from pre- to post-test. However, this group did not have a control group so no gain score comparisons could be made across treatment and control groups.

Three districts were using Seeing Math Elementary courses in the Number and Operations strand. In one district, there were no significant student gains for either year of the study, and no differences in gains between treatment and control groups. In a second district, treatment students did post significant gains over the second year of the study, and the treatment students had significantly higher gains than did the control students. In the third district, treatment students had significant gains in both years in some of the test sub-scores. However, both mean scores and gains for the control students exceeded those for the treatment students for all scores.

The somewhat confusing findings led to recommendations for new directions in future research and evaluation. Because the research design focused broadly on the entire course, it was difficult to determine what, if any, aspect of the professional development intervention was directly related to the specific student gains. A series of small, intensive design studies that focus on specific aspects of the Seeing Math Elementary approach or specific learning objectives so that models for effective teacher professional development can be built from the findings was recommended.

A final recommendation was to focus future research and evaluation attention on the development, testing, and validation of new student assessment instruments for both mathematics content and process standards, in line with the content and process focus of the Seeing Math Elementary program.

## Introduction

Seeing Math Elementary is a video-case based mathematics professional development program for elementary and middle school teachers created by the Concord Consortium and Teachscape. It is funded under the U.S. Department of Education's Ready To Teach Program (grant # R286A000006-03). Seeing Math Elementary is part of the Seeing Math Telecommunications Project.

Seeing Math Elementary includes Web-based video case studies that provide mathematics professional development for elementary and middle school teachers. These case studies use both real-life video narratives and guided inquiry to present a math content strand that is aligned with National Council of Teachers of Mathematics (NCTM) standards. There is also a classroom pedagogy strand. Each Seeing Math Elementary case study focuses on specific math content that is widely recognized as difficult to teach. Courses are typically taught in a blended online and face-to-face format.

The Seeing Math Video Case Studies include these nine cases falling under each of the following four content standards of the NCTM:

- Number and Operations: *Division with Remainders*; *The Magnitude of Fractions; Broken Calculator*
- Pre-Algebra: *Pan Balance Equations; Patterns and Functions*
- Geometry: *2D and 3D Figures; Calculating the Area of a Triangle*
- Data Analysis and Probability: *Using Data to Make Predictions; Measures of Center*

In addition, two "cross-cut" courses use excerpts from the above courses and focus specifically on pedagogy issues:

- *Effective Questioning in the Mathematics Classroom*
- *Formative Assessment in the Mathematics Classroom*

## Overall Goals and Objectives

The overall goal of Seeing Math Elementary is stated in the overall Ready To Teach GPRA indicator: To improve student achievement by developing high quality, standards-based digital professional development to teachers and by developing high quality, standards-based digital classroom content.

This evaluation report focuses on the following two overall objectives of the Seeing Math Telecommunications Project:

- Evaluate the quality and usefulness of project materials and strategies.
- Research the effect and impact of the Seeing Math materials.

Edcentric, an evaluation consulting company, has conducted the external evaluation for Seeing Math Elementary. During the 2004 – 2005 school year, the formative evaluation focused on evaluating the quality and usefulness of Seeing Math Elementary materials, particularly those used within the fall pilot implementation. During earlier years, when training the new site facilitators was key, evaluation included assessing the training and support provided to sites implementing the program, as well as formative evaluation of specific Seeing Math Elementary courses and materials.

Edcentric developed online survey instruments and interview protocols and collected data from participants using these instruments. In addition, online discussion boards were mined for relevant information to provide a more complete picture of the materials and how they were used.

In the past several years, the U.S. Department of Education has established new priorities for evaluation, with a strong emphasis on experimental and quasi-experimental research. This has resulted in Concord Consortium's development of a comprehensive research plan for the overall Seeing Math Elementary project beginning with the 2003 – 2004 school year and extending through the 2004 – 2005 school year. This final evaluation report for the project focuses on the research and evaluation efforts conducted by Edcentric, the external evaluator, within the research framework created by Concord Consortium. Findings of the 2004 – 2005 school year, in which quasi-experimental research is conducted with both treatment and control groups, are reported. Quantitative findings from the 2003 – 2004 school year are conflated into this report to provide a complete picture of the overall research within the design and to assess potential student growth over time through a longitudinal study based on data collected both years.

Concord Consortium developed an overall research design and plan to be implemented with both Seeing Math Elementary and Ready To Teach Algebra/Seeing Math Secondary, using quasi-experimental strategies and quantitative data collection instruments developed under Concord Consortium's direction, in response to the Department of Education priorities for evaluation.

Concord Consortium was responsible for the research design as well as for determining quantitative instrument development and data collection efforts. Edcentric, however, was responsible for analysis of the quantitative data collected by Concord Consortium, using the Concord Consortium research plan and instruments created and scored by both Concord Consortium and Northwest Evaluation Association (NWEA).

The quantitative research and evaluation effort has focused on the effect and impact of the Seeing Math Elementary materials on both teachers and their students, concomitant with their involvement in Seeing Math Elementary. Possible student gains over the two years of the study are explored using test data from both the 2003 – 2004 school year and the 2004 – 2005 school year.  During this final year of the project, additional emphasis has been on a comparison of those students whose teachers actively participated in Seeing Math Elementary – the treatment group – with those of students whose teachers did not participate in Seeing Math Elementary – the control group.  These treatment and control groups were matched and comparable in terms of school district, grade level, and teaching experience of the participating teachers.

The four school districts serving as research sites for this effort include Champaign, Illinois; Dover, Massachusetts; Mohonasen, New York; and Swampscott, Massachusetts.

These were also the sites in which the formative evaluation of the courses involved in the fall 2004 pilot test occurred.  Edcentric was responsible for qualitative instrument development and data collection efforts for the formative evaluation efforts, and was in regular contact with teachers and facilitators to collect a wide array of survey data in a timely manner.

## FORMATIVE EVALUATION: EVALUATING THE QUALITY AND USEFULNESS OF SEEING MATH ELEMENTARY MATERIALS AND STRATEGIES

### METHODOLOGY

During the 2004 – 2005 school year, four school districts – Champaign, Illinois; Dover, Massachusetts; Mohonasen, New York; and Swampscott, Massachusetts – served as primary research and pilot test sites for Seeing Math Elementary (SME). During fall 2004, teachers in these four districts pilot tested the following Seeing Math Elementary courses with their site facilitators: *Division with Remainders, Pan Balance Equations*, and *Broken Calculator*.

Edcentric conducted the formative evaluation of this pilot test. The evaluation team developed two online surveys (Teacher Background Survey and end of course Teacher Reflection) to gather background demographic data on participants and to collect information related to participants' use of the course materials in the development of their own mathematical thinking and in their classrooms. In addition, online discussions were read and analyzed by Edcentric to aid in an understanding of participants' perceived value of the Seeing Math Elementary materials. Thirty-six teachers provided background information, and thirty completed end of course Teacher Reflections for this component of the formative evaluation.

Edcentric also developed a corresponding Teacher Background Survey to collect demographic and pedagogical information from those teachers serving as control group teachers in the research sites.

During spring 2005, Edcentric conducted formative evaluation of *Formative Assessment in the Mathematics Classroom* and *Effective Questioning in the Mathematics Classroom* in schools that were not research and pilot test sites during the current school year, although the school district was involved in earlier Seeing Math Elementary research studies.

Overall findings from these formative evaluation efforts for the above five courses are presented in this report, focusing on teacher assessment of course elements and their perceptions of teacher and student change. Analysis of aspects of the online discussion is included as well.

Edcentric also conducted formative evaluation of *Calculating the Area of a Triangle* and *Using Data to Make Predictions* through intensive one-on-one reviews with individual teachers. During summer 2005, *2D and 3D Figures* was tested in a summer professional development course in a school system that had not previously been involved in the project. Since the timing of these evaluations did not allow teachers to try them in their classrooms to indicate teacher or student change, they are not reported here.

The formative goal in data analysis was to use the above data sources to answer these guiding evaluation questions:

- Are the organization, pacing, content, clarity, usability, flexibility, and appeal of the Seeing Math modules appropriate and effective for the target population?
- What changes occur in classroom practice that are related to participation?
- What changes occur in students' classroom behavior that are related to teacher participation in Seeing Math Elementary?
- What issues related to implementation (i.e., online versus face to face meetings, role of facilitator) emerge during the project?

## QUALITATIVE DATA COLLECTION INSTRUMENTS

Edcentric used Perseus SurveySolutions software to create all the online surveys used in the qualitative evaluation effort for Seeing Math Elementary. The survey quantitative and narrative data were summarized and the survey's numerical data were also exported from SurveySolutions into SPSS format. With SPSS, formative survey data could be easily integrated with the quantitative data collected by Northwest Evaluation Association (NWEA) and Concord Consortium for ease of data analysis. Data from the teacher surveys were merged with student data and teacher pre- and post-tests so that multivariate analysis could be conducted. These findings are discussed in the quantitative Effect and Impact section of this report.

## FORMATIVE EVALUATION FINDINGS: QUALITY AND USEFULNESS OF PROJECT MATERIALS

During the fall and early winter of the 2004 – 2005 school year, Concord Consortium and Teachscape conducted a pilot test of three courses in the Seeing Math Elementary program: *Broken Calculator*, *Pan Balance Equations*, and *Division with Remainders*. These courses were offered in four school districts nationally: Champaign, Illinois; Dover, Massachusetts; Mohonasen, New York; and Swampscott, Massachusetts. Each district offered one or more of the courses in a blended environment, with a combination of online learning and discussion, and face-to-face meetings led by a local facilitator.

The evaluation team developed two online surveys, one to gather background demographic data on participants and one to collect information related to participants' use of the course materials in the development of their own mathematical thinking and in their classrooms. We administered these surveys to participating teachers, and then summarized and analyzed data from the completed surveys, with separate post-course analysis for each of the three courses in the fall 2004 pilot study and two courses in the spring 2005 pilot study.

## Course Review: Broken Calculator

Thirteen teachers (8 from Champaign, 5 from Swampscott) completed the end of course review for *Broken Calculator*. These teachers taught math to a total of 274 students, and indicated that 273 of these students were in classes where they used the insights/strategies/activities that they took from this Seeing Math Elementary course. While seven of the teachers believed their classes were representative of the overall race/ethnicity enrollment in their schools, six did not, as four were gifted/talented classes and the others were, for unstated reasons, not representative.

### Assessment of Course Elements

Teachers rated several elements of the course and the tables below indicate the findings.

**Please indicate how well you feel *Broken Calculator* accomplished each of the following objectives for you and your teaching practice.**

| | Not well | Somewhat well | Well | Very well |
|---|---|---|---|---|
| Presented concepts and ideas that are relevant to my instructional practice | 0 | 0 | 2 | 11 |
| Provided clearly stated objectives/professional development goals | 1 | 1 | 1 | 10 |
| Helped me to achieve the objectives/professional development goals | 0 | 0 | 5 | 8 |
| Prepared me to apply the ideas and concepts presented to my instructional practice | 0 | 0 | 5 | 8 |

**Please indicate how important you feel each of the following elements of *Broken Calculator* has been to your teaching practice.**

| | Not at all important | Somewhat important | Important | Extremely important |
|---|---|---|---|---|
| Face to face meetings and discussion | 0 | 1 | 0 | 12 |
| Diving in activity | 0 | 2 | 4 | 7 |
| Video case study | 1 | 1 | 5 | 6 |
| Online activities | 0 | 2 | 7 | 4 |
| Expert commentary | 0 | 2 | 8 | 3 |
| Additional resources | 0 | 3 | 7 | 3 |

Teachers found the course clear and relevant to their instructional practice and professional goals. They found the face to face meetings and discussion to be the most important part of the course, but the diving in activity, video case study, online activities, expert commentary, and additional resources were important or extremely important to most participants.

We were also interested in learning how easy participants found the project website and online course elements to use, and asked them the following:

**Please consider the design of the website. How easy did you find it to use when doing the following tasks?**

| | Very difficult | Somewhat difficult | Somewhat easy | Very easy |
|---|---|---|---|---|
| Posting to online discussions | 0 | 1 | 3 | 9 |
| Navigating within the Teachscape website | 0 | 1 | 6 | 6 |
| Playing videos | 0 | 1 | 6 | 6 |
| Finding help when I need it | 0 | 4 | 2 | 7 |

Teachers generally found the website to be "extremely teacher friendly." Several teachers had difficulty accessing the videos because of dial up access and/or memory problems on their computers, but most found using the CD alleviated that problem. One teacher found posting difficult at first, but with experience it became relatively easy.

## Teacher Change

We were particularly interested in learning what changes teachers perceived to come as a result of their involvement in the *Broken Calculator* course – added insights into mathematical content, changes in teaching activities, or changes in students' mathematical thinking, problem solving and achievement.

Teachers provided a number of specific examples of ways in which *Broken Calculator* helped them gain added insights into mathematical content:

- As I reflected on (struggled with) my own solutions to the broken calculator, I was aware of my own levels of thought and realized the value of such a learning activity in my classroom. The notion that there are many viable strategies that cause deeper thoughts about the math takes math instruction away from rote algorithms and toward critical thinking and application.
- It helped to see that algorithms are not the only way to solve math problems and there should be an alternative for students since they have different learning styles.
- I now teach my students to find alternative solutions to problems, have whole class discussions which has helped increase my own knowledge of the subject. We had a great discussion in class when we tried to solve one of the broken calculator problems. I could see that I had underestimated what my students could do.
- It motivated me to guide my students in explaining their thinking more often and also in allowing my students to vary their approaches to solving problems when applicable. For

example, when correcting homework or classwork I found myself asking students to come up to the board to explain their correct answer to the class rather than only reviewing incorrect answers.

They provided specific examples of changes in their teaching activities:

- I encourage my students to discuss their strategies more. I am aware of my own questioning and try to echo students, pose questions to clarify misunderstandings, and highlight different strategies. I have been reminded that best practice in math causes students to think and analyze using number sense rather than perform pages of algorithms like automatons.
- I changed the questioning strategies to be more like the ones used in the course. I extended the broken calculator problems to the whole class rather than just "enrichment" students.
- I now ask the children to find as many ways as they can to solve a problem. Before, I would take the first answer and assume that was the only one.

### Student Change

Changes in students' mathematical thinking, problem solving, and achievement teachers attributed to their involvement in *Broken Calculator*:

- [Student 1] only used repeated addition. [Student 2] sneaked paper and pencil instead of using the broken calculator. [Student 1] was exposed to other strategies… she needs more time to internalize and use them though. I'm sure that with more practice, she will have a breakthrough though. [Student 2] needed strict "watching" and "counseling" to convince him that paper and pencil wasn't the "quickest and easiest way." He eventually gave in to my demands that he think about his thinking and he was very successful with the broken calculator.
- I think the students who have more difficulty with math changed the most. Presenting different ways to solve math problems helped them find a way to solve math that made sense to them. For example, when I give the students a word problem and tell them to draw it out and use a method that works they feel like they have a choice and they start to solve it in a way that makes sense to them.
- My students have become more comfortable discussing different ways to approach a problem during our problem-solving sessions. I had two students who were good at giving answers, but not understanding "why" the answer was correct, or how to explore different strategies. It has been hard for them to think though the work and really demonstrate understanding to other members of the group. They were very attached to the teacher saying right or wrong. They were thrown by "why or how" questions. Now I see them working more comfortably with their classmates and respecting that some of the "slower" students are good and finding ways to solve problems.
- Many of my students enjoyed the challenge that the broken calculator problems brought. My students became more flexible with looking at ways to solve problems. They also began looking at how number and operations relate to each other.

Teachers provided a number of examples, a sample of which are provided above, of the ways in which both they and their students changed their approach to and understanding of mathematical concepts and problem solving brought out through the *Broken Calculator* experience.

## Course Review: *Pan Balance Equations*

Eight teachers from Dover, Massachusetts completed the end of course review for *Pan Balance Equations*. These teachers taught math to a total of 117 students, and indicated that 113 of these students were in classes where they used the insights/strategies/activities that they took from this Seeing Math Elementary course. The teachers believed their classes were representative of the overall race/ethnicity enrollment in their schools, although one indicated that she was a special educator.

### Assessment of Course Elements

Teachers rated several elements of the course and the tables below indicate the findings.

**Please indicate how well you feel *Pan Balance Equations* accomplished each of the following objectives for you and your teaching practice.**

| | Not well | Somewhat well | Well | Very well |
|---|---|---|---|---|
| Provided clearly stated objectives/professional development goals | 0 | 0 | 3 | 5 |
| Helped me to achieve the objectives/professional development goals | 0 | 0 | 3 | 5 |
| Presented concepts and ideas that are relevant to my instructional practice | 1 | 0 | 2 | 5 |
| Prepared me to apply the ideas and concepts presented to my instructional practice | 0 | 1 | 3 | 4 |

**Please indicate how important you feel each of the following elements of *Pan Balance Equations* has been to your teaching practice.**

| | Not at all important | Somewhat important | Important | Extremely important |
|---|---|---|---|---|
| Face to face meetings and discussion | 0 | 0 | 1 | 7 |
| Video case study | 0 | 0 | 3 | 5 |
| Diving in activity | 0 | 2 | 3 | 3 |
| Online activities | 0 | 1 | 5 | 2 |
| Expert commentary | 0 | 2 | 6 | 0 |
| Additional resources | 0 | 4 | 4 | 0 |

Teachers found the course clear and relevant to their instructional practice and professional goals. They found the face to face meetings and discussion to be the most important part of the course, but the video case study, diving in activity, online activities, and expert commentary were important or extremely important to most participants.

We were also interested in learning how easy participants found the website and online course elements to use, and asked them the following:

**Please consider the design of the website. How easy did you find it to use when doing the following tasks?**

| | Very difficult | Somewhat difficult | Somewhat easy | Very easy |
|---|---|---|---|---|
| Playing videos | 0 | 0 | 4 | 4 |
| Navigating within the Teachscape website | 0 | 2 | 4 | 2 |
| Posting to online discussions | 0 | 2 | 4 | 2 |
| Finding help when I need it | 0 | 2 | 5 | 1 |

Several teachers in this group had early problems with posts: "The posting on on-line discussions was a bit difficult and I think a better way could be found. Also, at times [the website] quit on my while I was in the midst of a lengthy discussion piece and I lost the whole piece which did not make me happy. I learned to save in 'word' but a save button would have been helpful."

## Teacher Change

We were particularly interested in learning what changes teachers perceived to come as a result of their involvement in the *Pan Balance Equations* course – added insights into mathematical content, changes in teaching activities, or changes in students' mathematical thinking, problem solving and achievement.

Teachers provided a number of specific examples of ways in which *Pan Balance Equations* helped them gain added insights into mathematical content:

- The videos helped me see the math unit through the students' eyes.
- I never realized pan balance equations could preteach algebra.
- The course really helped me understand the thinking behind order of operations.

They provided specific examples of changes in their teaching activities:

- I modeled my teaching after the videos in wait time, rephrasing the question and not giving immediate feedback as to the right or wrong of each oral answer.
- When teaching the "=" sign, I will teach my students it is called the "is the same as" sign. I hope this will help my students to think of equations as something that needs to be balanced. Too often, students see the "=" sign and feel they have to DO something – primarily add.

- I have included more activities for my students to explore equality.  Using pictures and then moving slowly to number sentences.

### Student Change

Changes in students' mathematical thinking, problem solving, and achievement teachers attributed to their involvement in *Pan Balance Equations*:

- The biggest change would be the way they understand the term "equals" and that it can appear in different places in an equation.
- Using pan balances and working through a series of activities that gradually become more difficult.  This helps to build confidence in students' thinking about algebra.

Some teachers are not teaching math this year, or have not yet gotten to a unit where they would introduce the concepts and ideas raised in *Pan Balance Equations*.

## Course Review:  *Division with Remainders*

Nine teachers from Mohonasen, New York completed the end of course review for *Division with Remainders*.   These teachers taught math to a total of 451 students, and indicated that 326 of these students were in classes where they used the insights/strategies/activities that they took from this Seeing Math course.  The teachers believed their classes were representative of the overall race/ethnicity enrollment in their schools, with three indicating they taught special needs students.

### Assessment of Course Elements

Teachers rated several elements of the course and the tables below indicate the findings.

**Please indicate how well you feel *Division with Remainders* accomplished each of the following objectives for you and your teaching practice.**

| | Not well | Somewhat well | Well | Very well |
|---|---|---|---|---|
| Presented concepts and ideas that are relevant to my instructional practice | 0 | 0 | 1 | 8 |
| Prepared me to apply the ideas and concepts presented to my instructional practice | 0 | 0 | 1 | 8 |
| Provided clearly stated objectives/professional development goals | 0 | 0 | 4 | 5 |
| Helped me to achieve the objectives/professional development goals | 0 | 0 | 4 | 5 |

**Please indicate how important you feel each of the following elements of *Division with Remainders* has been to your teaching practice.**

| | Not at all important | Somewhat important | Important | Extremely important |
|---|---|---|---|---|
| Face to face meetings and discussion | 0 | 1 | 0 | 8 |
| Video case study | 0 | 1 | 3 | 5 |
| Diving in activity | 0 | 1 | 3 | 5 |
| Online activities | 0 | 1 | 3 | 5 |
| Expert commentary | 0 | 1 | 6 | 2 |
| Additional resources | 0 | 4 | 4 | 1 |

Teachers found the course clear and relevant to their instructional practice and professional goals. They found the face to face meetings and discussion to be the most important part of the course, but the video case study, diving in activity, online activities, and expert commentary were important or extremely important to most participants.

We were also interested in learning how easy participants found the website and online course elements to use, and asked them the following:

**Please consider the design of the website. How easy did you find it to use when doing the following tasks?**

| | Very difficult | Somewhat difficult | Somewhat easy | Very easy |
|---|---|---|---|---|
| Playing videos | 0 | 0 | 2 | 7 |
| Finding help when I need it | 0 | 0 | 2 | 7 |
| Posting to online discussions | 0 | 0 | 2 | 7 |
| Navigating within the Teachscape website | 0 | 0 | 3 | 6 |

Most teachers found the site to be "very accessible and easy to use," although a couple found posting to be "a little confusing." One commented, "I did like the accessibility of the program. I wouldn't have been able to do this if I couldn't work on it at home."

## Teacher Change

We were particularly interested in learning what changes teachers perceived to come as a result of their involvement in the *Division with Remainders* course – added insights into mathematical content, changes in teaching activities, or changes in students' mathematical thinking, problem solving and achievement.

Teachers provided a number of specific examples of ways in which *Division with Remainders* helped them gain added insights into mathematical content:

- The course was effective in getting me to think more deeply about division outside the algorithm especially in reference to bringing my students to that same higher level of cognition.
- Seeing Math made me think about the different types of division problems such as groups of n or n groups. I also became aware of some of the difficulties students have with division. For the most part they understand and can apply the algorithms they have been taught.
- I never really considered the difference between "groups of" and "number of groups" when teaching division. There are instances when this can be important, especially as students begin tackling more challenging problems.

They provided specific examples of changes in their teaching activities:

- I incorporated a lesson including pre and post assessment on dealing with remainders that I probably would not have done had I not been involved in this course.
- I spend more time reflecting and refining my questioning techniques. I also make sure that I include time for student discussion about their reasoning and strategies.
- I use more reciprocal teaching in the classroom. The different questioning techniques used in the videos demonstrated the importance that students are able to communicate their math reasoning.

### Student Change

Changes in students' mathematical thinking, problem solving, and achievement teachers attributed to their involvement in *Division with Remainders*:

- I have seen a big improvement among my students' division algorithms. The classification matrix served as a wonderful tool in guiding my students through interpreting remainders.
- Students are more aware that there are a variety of ways to look at mathematical problems. They are also improving in their ability to explain the methods they sued to solve problems. They benefit greatly from the dialogue with other students.
- I think that my students had, for the most part, already mastered the computation aspects of division with remainders. This course helped me to help my students start to think about remainders differently; realizing that in a real life situation, you have to have a plan for the dividend you begin with. It is a rare situation where you can just decide to end with a remainder. My students have become more adept at discerning how to express a remainder, whether that is a decimal, a fraction, rounding up to the next whole number, or coming up with a plan for dealing with the remainder.

## Course Review: *Assessment in the Mathematics Classroom*

Nine teachers completed the end of course review for *Assessment in the Mathematics Classroom*. These teachers taught math to a total of 379 students, and indicated that 377 of these students were in classes where they used the insights/strategies/activities that they took from this Seeing Math course. The teachers believed their classes were representative of the overall race/ethnicity enrollment in their schools.

## Assessment of Course Elements

Teachers rated several elements of the course and the tables below indicate the findings.

**Please indicate how well you feel *Formative Assessment in the Mathematics Classroom* accomplished each of the following objectives for you and your teaching practice.**

| | Not well | Somewhat well | Well | Very well |
|---|---|---|---|---|
| Presented concepts and ideas that are relevant to my instructional practice | 1 | 2 | 2 | 4 |
| Provided clearly stated objectives/professional development goals | 0 | 2 | 3 | 4 |
| Helped me to achieve the objectives/professional development goals | 0 | 3 | 5 | 1 |
| Prepared me to apply the ideas and concepts presented to my instructional practice | 1 | 3 | 5 | 0 |

**Please indicate how important you feel each of the following elements of *Formative Assessment in the Mathematics Classroom* has been to your teaching practice.**

| | Not at all important | Somewhat important | Important | Extremely important | Not applicable |
|---|---|---|---|---|---|
| Video case study | 0 | 2 | 2 | 5 | 0 |
| Face to face meetings and discussion | 0 | 1 | 4 | 3 | 1 |
| Expert commentary | 0 | 1 | 4 | 3 | 1 |
| Additional resources | 0 | 1 | 4 | 1 | 3 |
| Online activities | 1 | 3 | 2 | 1 | 2 |
| Diving in activity | 0 | 1 | 2 | 0 | 6 |

Teachers found the course materials clear and relevant to their instructional practice and professional goals. They found the video case study, face-to-face meetings and discussion, and expert commentary to be the most important parts of the course. This group of teachers used selected components of the course materials in conjunction with an ongoing professional development initiative. As such, the diving in activity in particular was not used by all participants.

We were also interested in learning how easy participants found the website and online course elements to use, and asked them the following:

**Please consider the design of the website. How easy did you find it to use when doing the following tasks?**

| | Very difficult | Somewhat difficult | Somewhat easy | Very easy |
|---|---|---|---|---|
| Posting to online discussions | 0 | 0 | 2 | 7 |
| Playing videos | 0 | 0 | 2 | 7 |
| Navigating within the Teachscape website | 0 | 3 | 1 | 5 |
| Finding help when I need it | 0 | 1 | 4 | 4 |

Teachers generally found the website to be teacher friendly. Several teachers had difficulty accessing the videos because of dial up access and/or memory problems on their computers, but most found using the CD alleviated that problem.

## Teacher Change

We were particularly interested in learning what changes teachers perceived to come as a result of their involvement in the *Formative Assessment in the Mathematics Classroom* course – added insights into mathematical content, changes in teaching activities, or changes in students' mathematical thinking, problem solving and achievement.

Teachers affirmed that *Formative Assessment in the Mathematics Classroom* helped them gain added insights into assessing mathematical content knowledge in their students:

- It helped me to better understand and clarify students' understanding of mathematical content.
- It helped me to be more confident as a teacher so that my students began to learn more. I think that by changing my thinking of how math looks has had a positive impact on the learning of my students.
- My knowledge of formative assessments and how to use them once you have them has increased.
- [I saw] the importance of examining student work using formative assessments to lead future content.
- The videos helped me to be able to be a better evaluator of student learning on the spot and gave me ideas for excursions to solidify content for my students.

They provided specific examples of changes or impact on their teaching activities as a result of the course:

- Although I was already using formative assessment, it helped to "see" what this might look like in a classroom through looking at the videos.
- Questioning strategies, using more formative assessments.
- This class reinforced another class I took last summer. Both helped me improve my questioning technique.
- I didn't necessarily change any practices, but I tried to improve questioning skills as a result of the videos and readings.

- Questioning strategies and allowing students more time to reflect on their mathematical thinking.
- I experimented with a variety of ways to gather student data.
- Using formative assessments to plan activities and define specific instruction for individuals.

### Student Change

Changes in students' mathematical thinking, problem solving, and achievement teachers attributed to their involvement in *Formative Assessment in the Mathematics Classroom*:

- They are better reasoners, problem solvers, have a better disposition towards mathematics. [They] are more accountable because they are more aware of expectations. I try to share a rubric or create one each time there is an assignment/assessment so students aren't surprised.
- With improved questioning techniques, I'm better able to elicit true understanding and misconceptions in all areas.
- I think my students have increased their knowledge due to the use of strategy boards.
- Students are becoming more proficient at explaining their thinking with words.

## Course Review: *Questioning Strategies in the Mathematics Classroom*

Seven teacher leaders used the *Questioning Strategies in the Mathematics Classroom* course as part of their professional development training and preparation for becoming teacher leaders and coaches during the 2005 – 2006 school year. They completed an abbreviated end of course review for *Questioning Strategies in the Mathematics Classroom*.

### Assessment of Course Elements

Teachers rated several elements of the course and the tables below indicate the findings.

**Please indicate how well you feel *Questioning Strategies in the Mathematics Classroom* accomplished each of the following objectives for you and your teaching practice.**

| | Not well | Somewhat well | Well | Very well |
|---|---|---|---|---|
| Provided clearly stated objectives/professional development goals | 0 | 0 | 3 | 4 |
| Helped me to achieve the objectives/professional development goals | 0 | 0 | 3 | 4 |
| Presented concepts and ideas that are relevant to my instructional practice | 0 | 0 | 0 | 7 |

| | Not well | Somewhat well | Well | Very well |
|---|---|---|---|---|
| Prepared me to apply the ideas and concepts presented to my instructional practice | 0 | 0 | 1 | 6 |

**Please indicate how important you feel each of the following elements of *Questioning Strategies in the Mathematics Classroom* has been to your teaching practice.**

| | Not at all important | Somewhat important | Important | Extremely important | Not applicable |
|---|---|---|---|---|---|
| Video case study | 0 | 0 | 3 | 4 | 0 |
| Expert commentary | 0 | 0 | 6 | 1 | 0 |
| Online activities | 0 | 0 | 5 | 2 | 0 |
| Additional resources | 0 | 1 | 4 | 2 | 0 |
| Face to face meetings and discussion | 0 | 0 | 4 | 2 | 0 |

## Teacher Change

Participating teachers indicated the following changes in teaching activities as a result of using these materials from *Questioning Strategies in the Mathematics Classroom*:

- The experience made me more aware of my purposes for questioning and how important it is to have an intent when questioning students.
- I changed the way I think about my teachers. I think I will be able to use these materials in a very meaningful way with my teachers.
- I used the questioning techniques and rationale with the students and teachers I coached.
- I became more aware of modeling what I saw in the videos whether it be questioning or assessment. After I see something being done I have a better ideas of what I need to do myself.
- I changed teaching activities as a result of using these resources by purposely phrasing questions using the beginning, middle, and end of the problem solving process. Scaffolding was happening!
- I always like to see modeling – it refreshes in my mind what I need to do as a teacher and a coach.

## Student Change

They attributed changes in their students' mathematical thinking, problem solving, and achievement to their own changed practice based on involvement in Seeing Math:

- It has helped me to question better, therefore giving my students a chance to clarify themselves in discussions about problems and justifications for their solutions.
- These materials were instrumental in moving me forward in my philosophy of inquiry base I have been implementing the techniques shown in Seeing Math for some time now.

It helped to preview these newer courses (assessment and questioning) to see what was in them and how I might use them in my building.

- I feel I can push their thinking to the next level. It really makes the student think through the problem to a better understanding.
- I am constantly working to refine my practice while I am in the classroom. I think helping teachers to develop a safe learning environment in their classrooms can be attributed to Seeing Math.
- When students were able to reach understanding by the teacher's use of highlighting, cueing, echoing, and conceptual facilitation questions, I was amazed as to how comfortable I felt in moving on to a new concept. Still "workin" on this!!!
- Questioning helps me to gently encourage deeper thinking. I am more conscious of how to help children understand math and my ears perk up when I hear telling!

## FACE TO FACE MEETINGS, ONLINE DISCUSSION AND FACILITATION

The fall 2004 – winter 2005 pilot test of Seeing Math Elementary featured blended instruction, with both face to face meetings and online instruction. The make-up of the groups in each of the four school districts was somewhat different: in two districts, there was only one or, at most, two participants from the same school. In one district, all participating teachers were from the same school. In the final district, all participating elementary teachers were from the same elementary school, and participating middle school teachers were from the same middle school.

In each of the four groups, the facilitators structured their meetings somewhat differently – one had weekly face to face meetings, two had bi-weekly face to face meetings, and one group met face to face monthly. In Dover, for example, teachers were all in the same school, and they met face to face weekly. Champaign and Swampscott teachers met bi-weekly, and each teacher was, in general, the only teacher in his/her building to be taking a Seeing Math Elementary course. In Mohonasen, where teachers met face to face on a monthly basis, participating teachers were in the same elementary or middle school building, depending on their teaching grade level.

Teachers had, therefore, varying amounts of contact with the other teachers in their Seeing Math class. Despite the contact difference, twenty-seven of the thirty respondents indicated they found the face to face meetings and discussion to be "extremely important" to their teaching practice. In contrast, eleven of the respondents considered the online activities to be "extremely important." Only two of the eight Dover participants, who taught in the same school and met weekly, found the online activities to be "extremely important." In Mohonasen, where teachers met face to face only monthly, more than half of the teachers found the online activities to be "extremely important."

Amount of face to face time might impact the relative importance of the online activities for participants, as the two could be in competition or coexistence with each other. However, the Seeing Math Elementary facilitators approached their online activities in quite different ways. What appeared most striking in the online discussion were differences between the amount of detail and depth in the facilitator's questions, and the guidance and direction these questions offered. Some facilitators' questions or discussion starters were brief and very open-ended, with an expectation that any kind of response would be acceptable. Some questions were very specific and the direction they offered would lead participants into in-depth thinking and analysis, as well as substantive interaction with each others' writing.

These differences were then reflected in the teacher responses. Here are some examples of varying levels of facilitator guidance offered in the Seeing Math Elementary courses. We compared two assignments from two different districts for the same activity:

> [Facilitator 1]: Read "Questioning Strategies" as well as the first half of "Overview, Rationale and Background" up to and including the section, "What Can Students Do with a Broken Calculator?" Continue making entries in your online journal, recording your insights and composing posts for online discussion.
>
> After making notes go to Week Three Discussions and post at least twice in the Discussions section under "What can students do with a broken calculator?"

And:

> [Facilitator 2]: Read "Questioning Strategies" as well as the first half of "Overview, Rationale and Background" up to and including the section, "What Can Students Do with a Broken Calculator?" Reflection on these readings can be made in your online journal to record your insights for later post to online discussions.
>
> Additionally, you are to read Number Sense-Making by Judith Sowder and Bonnie Schappelle. It can be found in your folder of articles handed out at the first meeting.
>
> After completing your readings, make an **initial post** on your reaction to the readings. **Revisit later in the week with at least two posts** to push the collective reflections on the readings.

Teacher responses to Facilitators 2's guiding questions were more thoughtful and in-depth than those to Facilitator 1's question in this example. Guiding questions on the part of the facilitators led to longer responses from teachers, and more substantive reflection on their peers' comments. These variations point out the salience of the facilitator in establishing and maintaining the overall learning climate and expectations within the online discussions.

In Seeing Math Elementary, the online component was only one part of the entire experience, however. Some facilitators may have chosen to emphasize discussion and sharing in the face to face component rather than through the online activities.

Participants did find the face to face component to be the most important for them, and the venue in which many participants found they had the greatest opportunity to share and learn from their colleagues.

## FORMATIVE EVALUATION SUMMARY

During the 2004 – 2005 school year, fifty-four teachers and teacher leaders completed end of course surveys for the eight courses tested during the year. Teachers rated elements of the course and found the face to face meetings and discussions, diving in activity, video case study, online activities, and expert commentary to be important or very important elements of the courses. The courses met their expectations overall and almost all participants felt they would recommend a Seeing Math Elementary course to a colleague. These teachers taught a total of 1,333 students. They indicated that 1,089 of those students were in classes where they used the insights, strategies, and/or activities they took from their Seeing Math Elementary course.

During the 2003 – 2004 school year, forty-two teachers completed full end of course surveys and sixteen short focused surveys for the five Seeing Math Elementary courses tested during the year (*Magnitude of Fractions, Patterns and Functions, Broken Calculator*, *Division with Remainders,* and *Measures of Center*).

Thirty-six of the forty-two teachers involved in the pilot testing of Seeing Math Elementary in spring 2004 believed they had changed their teaching practices as a result of their involvement with Seeing Math, and they provided specific examples in support of their beliefs.

Teachers reported that some 700 of their students were in classes where they used insights/strategies, or activities that they took from their Seeing Math Elementary course. Some teachers indicated they saw changes in their students' mathematical thinking, problem solving, or achievement and provided specific examples to bolster their beliefs, while most believed that their students in the coming school year would be even better able to reap the benefits of involvement. The 2003 – 2004 school year study did not include control groups against which to compare the treatment effects.

## PARTICIPATING TREATMENT AND CONTROL TEACHERS

Thirty-eight treatment teachers from the Seeing Math Elementary research sites completed online demographic background surveys during the 2004 – 2005 school year. These treatment teachers were active participants in Seeing Math Elementary professional development coursework. These teachers, mostly white and female, were primarily experienced, certified elementary teachers. They enjoyed teaching mathematics, but most of them did not consider themselves to be "master" math teachers. They had taken advanced coursework, with two-thirds having completed masters degrees and additional coursework. Their most recent work with math and math teaching, however, had involved in-service teacher professional development rather than graduate credit courses. These teachers were at least somewhat familiar with NCTM standards.

We were also interested in learning what objectives these teachers emphasized in their mathematics classes, and some of their practices in their math classrooms. These teachers, in general, emphasized learning math concepts, algorithms, computational skills, problem solving, and mathematical reasoning while increasing their students' interest in math. They used a variety of pedagogic techniques. They did not, however, expand mathematics connections to real world business applications or emphasize the history and nature of mathematics.

In addition, 17 control teachers, who were not participants in Seeing Math Elementary, also completed similar online demographic background surveys. We were interested in compiling a detailed demographic profile of participating teachers, both treatment and control. Facilitators at the participating sites were asked to select control teachers within their school districts for the participating treatment teachers, based on grade level and years of teaching experience. Completion of demographic profiles allowed us to determine the comparability of treatment and control teachers, however, to a much finer level of detail.

In comparing the demographic profiles of the treatment and control teachers, we found that the groups were comparable based on teacher survey responses. The teachers were predominantly white, female certified elementary teachers with up to ten years of teaching experience. They enjoyed teaching mathematics. Teachers, both treatment and control, were predominantly white, female, certified elementary teachers with up to 10 years of teaching experience. These teachers indicated they enjoyed teaching mathematics. More treatment teachers than control teachers had engaged in recent activities related to math and the teaching of math, as would be expected given the treatment teachers' active involvement in the Seeing Math Elementary program. Demographically, the two groups are quite comparable.

We were also interested in learning what pedagogic objectives these teachers emphasized in their mathematics classes, and some of their practices in their math classrooms. Teachers, whether treatment or control teachers, emphasized learning math concepts, algorithms, computational skills, problem solving, and mathematical reasoning while increasing their students' interest in math. They used a variety of student-centered pedagogic techniques.

The similarity of the demographic profiles and pedagogic emphases of both the treatment and control group teachers indicate that these two groups of teachers were comparable, and therefore appropriate to use as comparison groups in the empirical studies.

## QUANTITATIVE EVALUATION: EVALUATING THE EFFECT AND IMPACT OF SEEING MATH ELEMENTARY

### METHODOLOGY

In response to the U.S. Department of Education's stated interest in quantitative research, the Concord Consortium presented the external evaluator, Edcentric, with a quasi-experimental research design to assess the impact of the Seeing Math Elementary materials on teachers and students. In addition, the Concord Consortium organized the development and collection of all measures of teacher and student content learning. The charge to Edcentric was to take the quantitative performance data it was given and analyze these data to answer the initiative's central research questions:

- What *teacher gains* in math content and pedagogy can be attributed to their participation in Teacher Professional Development (TPD) designed by the Seeing Math Elementary Project?
- What *student gains* in math performance can be attributed to the participation of their teachers in Teacher Professional Development (TPD) designed by the Ready to Teach Project?

Additional questions, to be addressed if time and data permit, include:

- Are there differences in teacher gains by grade, school, academic background, experience, or implementation quality?
- Are there differences in student gains by grade, gender, ethnicity, school, teacher academic preparation, teacher experience, teacher gains, or implementation quality?
- Do student and teacher gains persist?

The table below schematizes the quantitative research design developed by Concord Consortium for the treatment and measurement sequence in Seeing Math Elementary research. **S-Pre-test** represents Student Pre-test; **S-Post-tes**t represents Student Post-test. **T** represents the Treatment teachers' participation in Seeing Math Elementary: Teacher pre-test, Seeing Math Elementary course, teacher post-test. **C** represents the Control or comparison teachers' participation in Seeing Math Elementary: Teacher pre-test and teacher post-test, without treatment.

**Seeing Math Elementary Treatment and Measurement Sequence**

| | Year 1, 2004 | Year 2, 2004-05 | |
|---|---|---|---|
| | Spring 04 | Fall 04 | Spring 05 |
| Cohort 1 | S-Pre-test **T** S-Post-test | S-Pre-test | S-Post-test |
| Cohort 2 | S-Pre-test S-Post-test | S-Pre-test **T** | **T** S-Post-test |
| Cohort 3 treatment | | S-Pre-test **T** | **T** S-Post-test |

| | Year 1, 2004 | Year 2, 2004-05 | |
|---|---|---|---|
| | Spring 04 | Fall 04 | Spring 05 |
| Cohort 3 control | | S-Pre-test **C** | **C** S-Post-test |

This quasi-experimental design allowed for cross-sectional comparisons (across cohorts within a given year) as well as longitudinal comparisons (within cohorts across consecutive years). This combination of comparisons was intended to provide a close inspection of the impact of the materials in the absence of random assignment to conditions.

## QUANTITATIVE DATA COLLECTION INSTRUMENTS – TEACHER

Teacher tests were developed to measure teachers' math content and pedagogy knowledge related to the corresponding Seeing Math Elementary course. Nine teacher tests were created, one for each of the nine content-based Seeing Math Elementary cases. These open-ended paper and pencil tests were created and validated by Concord Consortium's research group, with support from external math testing consultants, and with field testing of items and the corresponding scoring rubric. Teacher pre-tests were applied right before a course was offered, and teacher post-tests were applied upon completion of the course.

The teacher tests addressed the following content areas: Modeling/Formulating (MF), Transforming/Manipulating(TM), Inferring/Drawing Conclusions (IDC), and Communicating (C). These content area variables were given a weight and they were measured using an ordinal rating scale, which ranged from 0 (lowest) to 3 (highest) representing the proficiency level displayed in testing. The same ordinal rating scale was used in a pre- and post-test rating of a measure called Pedagogy, which was an assessment value based on the teacher's response to inquiries, their level of thought expressed, and their awareness of math content and student thinking. A total weighted score was calculated for each content area, MF, TM, IDC and C, in both pre-test and post-test modes. A Total Weighted Composite Score was developed for both pre- and post-test data and was a summative scale that combined the total weighted scores for the four content areas.

## FINDINGS FROM THE QUANTITATIVE ANALYSIS FOR TEACHER GAINS

A variety of statistical tests was employed to answer research questions related to possible teacher gains. A series of paired samples t-tests was performed with weighted sub-scores for the content areas of Modeling/Formulating (MF), Transformation/Manipulation (TM), Inferring/Drawing Conclusions (IDC), and Communicating (C), and then for the Total weighted composite score for treatment and control groups.  Table T1 below lists the pre- and post-test scores and related statistics for treatment group teachers.

**Table T1:  Post- Pre-test T-test results for Treatment Teachers**
**Paired Samples Statistics**

| | | Mean | N | Std. | Std. Error |
|---|---|---|---|---|---|
| Pair 1 | MF Pre- Total Weighted Score | 5.25 | 12 | 3.545 | 1.023 |
| | MF Post- Total Weighted Score | 7.00 | 12 | 3.384 | .977 |
| Pair 2 | TM Pre- Total Weighted Score | 4.94 | 18 | 3.115 | .734 |
| | TM Post- Total Weighted Score | 7.89 | 18 | 3.008 | .709 |
| Pair 3 | IDC Pre- Total Weighted Score | 3.56 | 18 | 2.357 | .556 |
| | IDC Post- Total Weighted Score | 4.89 | 18 | 1.745 | .411 |
| Pair 4 | C Pre- Total Weighted Score | 3.33 | 18 | 2.449 | .577 |
| | C Post- Total Weighted Score | 4.83 | 18 | 1.724 | .406 |

The results indicate that there were statistically significant gains ($p < 0.05$) in all content area sub-scores and the Total Weighted Composite Score for the Treatment teachers.

Table T2 below lists the pre- and post-test scores and related statistics for the Control teachers.`

**Table T2:  Post- Pre-test T-test results for Control Teachers**
**Paired Samples Statistics**

| | | Mean | N | Std. | Std. Error |
|---|---|---|---|---|---|
| Pair 1 | MF Pre- Total Weighted Score | 2.67* | 3 | .577 | .333 |
| | MF Post- Total Weighted Score | 2.67* | 3 | .577 | .333 |
| Pair 2 | TM Pre- Total Weighted Score | 8.00 | 3 | 4.000 | 2.309 |
| | TM Post- Total Weighted Score | 10.67 | 3 | 2.309 | 1.333 |
| Pair 3 | IDC Pre- Total Weighted Score | 3.33* | 3 | 1.155 | .667 |
| | IDC Post- Total Weighted Score | 3.33* | 3 | 1.155 | .667 |
| Pair 4 | C Pre- Total Weighted Score | 4.67 | 3 | 1.155 | .667 |
| | C Post- Total Weighted Score | 5.33 | 3 | 1.155 | .667 |

* The correlation and t cannot be computed because the standard error of the difference is 0.

There were no statistically significant gains in any content area sub-scores or in the Total Weighted Composite Score, for the Control teachers.

The Wilcoxon Signed Rank test was used to test for changes in the pedagogy rating for treatment and control groups. The results showed a significant ($p = 0.001$) positive increase in the pedagogy rating for treatment teachers from pre- to post-test. There was no significant positive increase in the pedagogy ratings from pre- to post-test for the control teachers.

A one-factor analysis of variance (ANOVA) was used to determine if there were differences in Total Weighted Composite Score gains when examined across school districts. The results indicated that the average weighted composite score gain for Champaign was higher than for the other districts, and the difference was statistically significant for Dover and Swampscott, but not for Mohonasen.

A one-factor analysis of variance (ANOVA) was used to determine if there were differences in Total Weighted Composite Score gains when examined across test topic. The results indicated there were statistically significant differences in average gains for treatment teachers who took the test for the *The Magnitude of Fractions* course. They demonstrated the highest gains in the Total Weighted Composite Score as well as the Transforming/Manipulating (TM) and Inferring/Drawing Conclusions (IDC) sub-scores.

Correlation analysis was done with both treatment and control teachers to see if there was a relationship between years of teaching mathematics and any of the test gain scores. Only one correlation proved to be significant and that was between the number of years experience teaching mathematics and the Pedagogy score gain ($r = 0.685$, $p = 0.01$) for the treatment teachers.

## SUMMARY OF TEACHER FINDINGS

The series of statistical tests performed on the teacher data provided a number of overall findings. First, the treatment teachers (N=18) had significant gains on the post-test for each of the content areas of Modeling/Formulating, Transforming/Manipulating, Inferring/Drawing Conclusions, and Communicating, as well as for the Total Weighted Composite score. In addition, their post-test rating for Pedagogy was significantly higher than their pre-test rating for Pedagogy. For treatment group teachers, participation in Seeing Math Elementary had a high correlation with enhanced performance on this test.

Control teachers, on the other hand, did not have statistically significant gains in any content area sub-scores or for the Total Weighted Composite score, nor did the control teachers demonstrate a significant positive rating in Pedagogy when comparing their pre- and post-test scores.

The significant differences between gains for treatment and control teachers provide evidence that participation in Seeing Math Elementary provides teachers with relevant new learnings in a number of different content and pedagogic areas. Treatment teachers performed significantly better in the content areas of Modeling/Formulating, Transforming/Manipulating, Inferring/Drawing Conclusions, and Communicating than did the control teachers. Additionally, they performed significantly better than the control teachers in Pedagogy ratings. These findings are both positive and promising.

Comparison between the test score performance of treatment and control teachers, however, must be regarded with some caution, as the control group sample size was very small. Only three of the eleven control teachers completed both the pre- and post-tests.

In terms of total gain scores, Champaign teachers, who had taken part in the *The Magnitude of Fractions* course, had higher gains than the other school districts, and the gains were statistically significant for all except Mohonasen, which was the only other school district using the *The Magnitude of Fractions* course.

Number of years of teaching experience appeared to have little or no correlation with teacher gain scores. For treatment teachers, only one correlation was significant and that was between number of years teaching and the pedagogy score gain.

## QUANTITATIVE DATA COLLECTION INSTRUMENTS – STUDENT

Northwest Evaluation Association (NWEA) constructed student pre- and post-tests, using a list of standards given to them by Concord Consortium that were based on the topics covered by the cases in each of four areas: Number and Operations, Pre-Algebra, Geometry, and Data Analysis and Probability. NWEA created four student tests, one for each area, with items that matched the standards Concord had requested. Items were selected to offer a grade-appropriate difficulty range for grades 3-6 in order to allow for the calculation of accurate growth scores. All items were multiple-choice and some allowed the use of a four-function calculator.

Each test was a 20-item multiple choice RIT (Rasch Unit) assessment test. The RIT tests were developed for this assessment as scaled proficiency instruments developed by Northwest Evaluation Association (NWEA). **Student tests** measure math knowledge related to the National Council of Teachers of Mathematics standard that is directly related to the Seeing Math Elementary courses taken by participating teachers. Each student test measures math content knowledge related to the two or three courses created by SME for each of four NCTM standards for grades 3-6. Four sets of 20-items paper-and-pencil tests were created by NWEA for each NCTM standard. Each test was built by selecting valid items from an existing test collection of around 5000 test items. Students took pre-tests before treatment teachers participated in Seeing Math Elementary professional development and post-tests near the end of the school year.

Student data were compiled and put into an SPSS spreadsheet for answering key research questions addressing student gains and their relationship to various teacher variables.

The resulting data file for the students included the following variables: Name, Date of birth, Sex, Ethnicity, Language, Test Number, Test Code, School, Grade, Subject, RIT Test Score (both pre and post), RIT and Goals 1-3 Test Score (both pre and post) for the Number and Operations test, RIT and Goals 1- 2 Test Score Gains (both pre and post) for the Pre-Algebra test. The teacher data were then merged with the student data in order to analyze the relationship, if any, of student and teacher gains.

For the Number Sense/Number and Operations tests, the Goal 1 sub-score corresponded to questions most closely associated with *The Magnitude of Fractions*, Goal 2 sub-score to *Division with Remainders*, and Goal 3 sub-score to *Broken Calculator*. For Pre-Algebra, the Goal 1 sub-score corresponded to questions most closely associated with *Patterns and Functions* and Goal 2, to *Pan Balance Equations*.

## FINDINGS FROM THE QUANTITATIVE ANALYSIS FOR STUDENT GAINS

In order to determine most accurately any longitudinal growth and track change over time, we compared test means over the two years of the study for the various cohort groups and treatment and comparison groups within each of the four school districts participating in the research. In this way, we could ascribe any student change to the specific Seeing Math Elementary interventions within the district. Dover teachers took Seeing Math Elementary courses within the Pre-Algebra strand, and their students took the Pre-Algebra student pre- and post-tests. Champaign, Mohonasen, and Swampscott teachers took Seeing Math Elementary courses within the Number and Operations strand, and their students took the Number Sense/Number and Operations student pre- and post-tests.

Data were compiled and analyzed for the longitudinal study to determine growth over time within each participating school district. The findings are presented as case studies of each of the four towns.

### Dover

Teachers in the Dover school district were involved in the Pre-Algebra strand of Seeing Math Elementary. Cohort 1 teachers took *Patterns and Functions* during spring 2004. Their students took the NWEA Pre-Algebra pre-test before the professional development course began in the spring, and the post-test at the end of the school year. Cohort 2 teachers took *Pan Balance Equations* during fall 2004. Their students took the NWEA

Pre-Algebra pre-test in fall 2004 and the post-test at the end of the 2004 – 2005 school year.

Mean data from participating teachers were developed from the raw student data sets with analysis at the individual student level, rather than using weighted aggregate class means. This allowed us to account for differing class sizes. A combination of T-tests and Analysis of Variance (ANOVA) were used to conduct the following analyses.

The following Table D1 shows the test means for Cohorts 1 and 2 across the 2 years of the study. Note that G1 and G2 represent goal 1 (*Patterns and Functions* sub-scores) and goal 2 (*Pan Balance Equations* sub-scores) respectively. Analyses of these means allowed us to determine gain scores over time.

**Table D1: Dover – Means For Pre- and Post-Tests By Cohort And Year**

| | | RIT Mean | G1 Mean | G2 Mean |
|---|---|---|---|---|
| Year 1 Pre-test (spring 2004) | Cohort 1 | 206.6 | 206.7 | 206.0 |
| Year 1 Post-test (spring 2004) | Cohort 1 | 206.6 | 209.2 | 206.2 |
| Year 2 Pre-test (fall 2004) | Cohort 1 | 201.9 | 204.0 | 201.4 |
| Year 2 Post-test (spring 2005) | Cohort 1 | 210.5 | 215.1 | 209.3 |

| | | RIT Mean | G1 Mean | G2 Mean |
|---|---|---|---|---|
| Year 1 Pre-test (spring 2004) | Cohort 2 | 212.7 | 212.1 | 214.0 |
| Year 1 Post-test (spring 2004) | Cohort 2 | 218.1 | 222.5 | 214.2 |
| Year 2 Pre-test (fall 2004) | Cohort 2 | 218.7 | 221.1 | 216.2 |
| Year 2 Post-test (spring 2005) | Cohort 2 | 225.9 | 227.5 | 223.9 |

Analysis of the various means for statistical significance (at the $p < 0.05$ level) in gains over the longitudinal time frame yielded the following findings:

- For year 1, there was a significant gain in goal 1 sub-scores for Cohort 1.
- For year 2, there was a significant gain in all test scores including the overall RIT and the goal sub-scores for Cohort 1. These gains were significant when viewed not only in year 2 but also across the 2-year time period for Cohort 1.
- With respect to Cohort 2, there was a significant gain in year 1 for the overall RIT and the goal 1 sub-score.

- For year 2 of the study, Cohort 2 gains were seen to be substantial for all scores with over a 6 point gain in all cases. These gains were all found to be statistically significant.
- In examining the data for Dover's Cohort 2 data over the 2 years, there were significant increases in the post-test scores across all tests, comparing year 2 to year 1 data.

During year 1, the Cohort 1 Dover teachers were involved in the *Patterns and Functions* Seeing Math Elementary course. The students of these teachers had significant gains in their sub-scores for goal 1 – those questions specifically related to the content standards for *Patterns and Functions*. At the same time, students of Cohort 2 teachers, who had not yet had completed any Seeing Math Elementary professional development, had significant gains in both overall RIT scores and goal 1 sub-scores.

During year 2, the students of Cohort 1 teachers continued to see gains, even though their teachers were not taking any Seeing Math Elementary courses during the 2004 –2005 school year. Students of Cohort 1 teachers had significant gains in both sub-scores (*Patterns and Functions* and *Pan Balance Equations*) as well as overall RIT scores.

During year 2, Cohort 2 teachers were involved in both the *Pan Balance Equations* (fall 2004) and *Broken Calculator* (spring 2005) courses. Their students took only the Pre-Algebra assessment tests. Students of Cohort 2 teachers had significant gains on all scores, both in year 2 and in comparing the year 2 to year 1 data.

These positive findings are promising indicators of the potential success of Seeing Math Elementary as an effective teacher professional development program within the Pre-Algebra content strand. For those Cohort 1 teachers who were involved in explicit Seeing Math Elementary training only in spring 2004, significant student gains continued into the 2004 – 2005 school year, and in fact increased from gains only in the goal 1 sub-score to gains in both the overall RIT scores and the goal 1 sub-score. The students of Cohort 2 teachers who were involved in explicit Seeing Math Elementary training during the 2004 – 2005 school year saw significant gains in both sub-scores and in overall RIT scores, whereas in 2003 – 2004, the students of Cohort 2 teachers saw significant gains in only in the goal 1 sub-score and overall RIT score.

These findings are presented with several caveats, however. First, the 2003 – 2004 time frame for Year 1 for possible student gains was only half a year, given the timing of pre- and post-tests. In contrast, the timing of the Year 2 testing allowed students (and their teachers) an entire school year to make gains from pre- to post-test. While the assumption is that students are comparable for year 1 and year 2, it is possible that they are, in fact, not. It is also possible that other unknown factors within the school district have influenced the overall achievement levels of participating students. Replication of this study with control or comparison groups would put these questions to rest and give even more credence to these promising findings.

## Champaign

Champaign teachers were involved in the Number and Operations strand of Seeing Math Elementary. Cohort 1 teachers took *Division with Remainders* during spring 2004. Their students took the NWEA Number Sense/Number and Operations pre-test before the professional development course began in the spring, and the post-test at the end of the school year. Two cohort 1 teachers took *Broken Calculator* in fall 2004. Cohort 3 teachers took *Broken Calculator* in fall 2004 or *The Magnitude of Fractions* in spring 2005. Cohort 3 teachers were matched with control teachers in their school districts for comparison during the 2004 – 2005 school year.

Mean data from participating teachers were developed from the raw student data sets with analysis at the individual student level rather than using weighted aggregate class means. This allowed us to account for differing class sizes. A combination of T-tests and Analysis of Variance (ANOVA) were used to conduct the following analyses.

The following Table C1 shows the test means for Cohort 1 across the 2 years of the study. Note that G1 represents Goal 1 (*The Magnitude of Fractions* sub-scores), G2 represents Goal 2 (*Division with Remainders* sub-scores), and G2 represents Goal 3 (*Broken Calculator* sub-scores).

**Table C1: Champaign – Means for Pre- and Post-Tests by Year – Cohort 1**

| | | RIT Mean | G1 Mean | G2 Mean | G3 Mean |
|---|---|---|---|---|---|
| Year 1 Pre-test (spring 2004) | Cohort 1 | 215.2 | 208.4 | 217.7 | 217.0 |
| Year 1 Post-test (spring 2004) | Cohort 1 | 212.1 | 210.3 | 209.5 | 218.9 |
| Year 2 Pre-test (fall 2004) | Cohort 1 | 213.5 | 206.0 | 215.4 | 217.2 |
| Year 2 Post-test (spring 2005) | Cohort 1 | 214.4 | 209.8 | 213.8 | 219.6 |

During year 2, control or comparison groups were added to the study. Comparison teachers within the district were selected and their students took Number Sense/Number and Operations pre- and post-tests within the same time frame as their Cohort 3 treatment counterparts.

Table C2 below shows the gain score results, or the difference between pre- and post-test scores, for the Cohort 3 treatment and control group students in Year 2.

**Table C2: Champaign – Mean Gain Score Results – Treatment and Control**

| | RIT Mean | G1 Mean | G2 Mean | G3 Mean |
|---|---|---|---|---|
| Treatment | 5.4 | 7.2 | 2.8 | 6.7 |
| Control | 6.3 | 8.6 | 2.3 | 7.8 |

Analysis of the various means for statistical significance (at the $p < 0.05$ level) in gains over the longitudinal time frame yielded the following findings:

- There were no significant gains in the overall RIT score for Champaign Cohort 1 comparing year 2 and year 1 performance. Results for individual goal scores indicate some gains, but they were not significant.
- Using gain scores in Year 2 for treatment versus control students, there were no significant differences in any scores for Cohort 3.

Although there were some gains in the individual goal sub-scores, they were not significant. These findings do not indicate any significant student gains over time for the Champaign students, based on the involvement of their teachers in the Seeing Math Elementary teacher professional development program.

## Mohonasen

Mohonasen teachers were involved in the Number and Operations strand of Seeing Math Elementary. Because of staff movement and changing assignments throughout the school system, only one of the participating treatment teachers from Mohonasen in the original Cohort 1 taught similar classes both years of the study. Therefore, the data set for the longitudinal study involved just this one teacher and his/her students, as well as a matching control or comparison teacher and his/her students.

Cohort 1 Mohonasen teachers took *The Magnitude of Fractions* during spring 2004. Their students took the NWEA Number Sense/Number and Operations pre-test before the professional development course began in the spring, and the post-test at the end of the school year. The Cohort 1 teacher who continued with a similar class in Year 2 (same grade level, heterogeneous group) took *Division with Remainders* in fall 2004 and *Broken Calculator* in spring 2005. This teacher's students, and the students of the matching control or comparison group teacher, took the NWEA Number Sense/Number and Operations pre-test in fall 2004 and the post-test at the end of the 2004 – 2005 school year.

Mean data from participating teachers were developed from the raw student data sets with analysis at the individual student level rather than using weighted aggregate class means. A combination of T-tests and Analysis of Variance (ANOVA) were used to conduct the following analyses.

The following Table M1 shows the test means for Cohort 1 across the 2 years of the study. Note that G1 represents Goal 1 (*The Magnitude of Fractions* sub-scores), G2 represents Goal 2 (*Division with Remainders* sub-scores), and G2 represents Goal 3 (*Broken Calculator* sub-scores).

**Table M1: Mohonasen – Means for Pre- and Post-Tests by Year – Cohort 1**

| | | RIT Mean | G1 Mean | G2 Mean | G3 Mean |
|---|---|---|---|---|---|
| Year 1 Pre-test (spring 2004) | Cohort 1 | 223.7 | 216.6 | 221.4 | 225.3 |
| Year 1 Post-test (spring 2004) | Cohort 1 | 221.4 | 215.3 | 219.0 | 226.5 |
| Year 2 Pre-test (fall 2004) | Cohort 1 | 222.7 | 214.6 | 223.4 | 223.9 |
| Year 2 Post-test (spring 2005) | Cohort 1 | 225.0 | 216.8 | 221.5 | 233.6 |

During Year 2, a control or comparison group was added to the study. A comparison teacher within the district was selected and his/her students took Number Sense/Number and Operations pre- and post-tests within the same time frame as the Cohort 1 treatment counterpart during Year 2.

Table M2 below shows the gain score results, or the difference between pre- and post-test scores, for the Cohort 1 treatment and control group students.

**Table M2: Mohonasen – Mean Gain Score Results – Treatment and Control**

| | RIT Mean | G1 Mean | G2 Mean | G3 Mean |
|---|---|---|---|---|
| Treatment | 1.8 | -.9 | -1.8 | 10.0 |
| Control | -2.3 | -.6 | -5.9 | 4.6 |

Analysis of the various means for statistical significance (at the $p < 0.05$ level) in gains over the longitudinal time frame yielded the following findings:

- Significant gains were seen in the overall RIT score for Mohonasen Cohort 1 comparing Year 2 to Year 1 performance. The results for the individual goal sub-scores indicate some gains, but with significance only for the goal 3 (*Broken Calculator*) sub-score. For this sub-score, there was approximately a 10 point gain comparing pre- and post-tests for Year 2 and with an 8 point gain comparing Year 2 to Year 1.
- Treatment students had significantly higher gain scores than control students for the overall RIT score and the goal 2 (*Division with Remainders*) and goal 3 (*Broken Calculator*) sub-scores.

These findings are promising with regard in particular to the goal 3 sub-score for *Broken Calculator*, especially as the treatment teacher completed the *Broken Calculator* course shortly before students took the Year 2 post-test, and that is the sub-score with significant Cohort 1gains, as well as for the treatment versus control group. Treatment students had significantly higher gain scores than control students for the overall RIT score and the goal 2 sub-score as well. At the same time, the findings need to be viewed with caution, as the treatment and comparison groups each involved a single teacher and his/her classes. It is possible that some other intervening classroom variable contributed to the differences between groups. Replication of this study with a larger sample of classes would put this question to rest and give even more credence to these promising findings. Unfortunately, the reality of life in schools, with changing enrollments and re-assignment of personnel to meet district needs, did not allow for that larger sample of classes in this district during the two years of this longitudinal study.

## Swampscott

Swampscott teachers were involved in the Number and Operations strand of Seeing Math Elementary. Cohort 1 teachers took *Broken Calculator* during spring 2004. Their students took the NWEA Number Sense/Number and Operations pre-test before the professional development course began in the spring, and the post-test at the end of the school year. Cohort 2 teachers took *Broken Calculator* in fall 2004.

Mean data from participating teachers were developed from the raw student data sets with analysis at the individual student level rather than using weighted aggregate class means. This allowed us to account for differing class sizes. A combination of T-tests and Analysis of Variance (ANOVA) were used to conduct the following analyses.

The following Table S1 shows the test means for Cohort 1 across the 2 years of the study. Note that G1 represents Goal 1 (*The Magnitude of Fractions* sub-scores), G2 represents Goal 2 (*Division with Remainders* sub-scores), and G2 represents Goal 3 (*Broken Calculator* sub-scores).

**Table S1: Swampscott – Means for Pre- and Post-Tests by Cohort and Year**

| | | RIT Mean | G1 Mean | G2 Mean | G3 Mean |
|---|---|---|---|---|---|
| Year 1 Pre-test (spring 2004) | Cohort 1 | 213.3 | 202.6 | 218.1 | 216.6 |
| Year 1 Post-test (spring 2004) | Cohort 1 | 214.3 | 208.7 | 216.5 | 218.7 |
| Year 2 Pre-test (fall 2004) | Cohort 1 | 214.7 | 206.1 | 217.2 | 218.2 |
| Year 2 Post-test (spring 2005) | Cohort 1 | 217.9 | 211.2 | 219.6 | 222.8 |

| | | RIT Mean | G1 Mean | G2 Mean | G3 Mean |
|---|---|---|---|---|---|
| Year 1 Pre-test (spring 2004) | Cohort 2 | 219.4 | 207.4 | 222.8 | 223.5 |
| Year 1 Post-test (spring 2004) | Cohort 2 | 215.8 | 216.3 | 217.5 | 215.6 |
| Year 2 Pre-test (fall 2004) | Cohort 2 | 205.6 | 200.6 | 209.1 | 209.7 |
| Year 2 Post-test (spring 2005) | Cohort 2 | 210.3 | 209.1 | 211.4 | 216.6 |

During year 2, control or comparison groups were added to the study. Comparison teachers within the district were selected and their students took Number Sense/Number and Operations pre- and post-tests within the same time frame as their treatment counterparts.

Table S2 below shows the gain score results, or the difference between pre- and post-test scores, for the treatment and control group students in Year 2.

**Table S2: Swampscott – Mean Gain Score Results – Treatment and Control**

| | RIT Mean | G1 Mean | G2 Mean | G3 Mean |
|---|---|---|---|---|
| Treatment | 4.0 | 5.3 | 2.4 | 5.5 |
| Control | 7.7 | 8.3 | 4.4 | 11.9 |

Analysis of the various means for statistical significance (at the $p < 0.05$ level) in gains over the longitudinal time frame yielded the following findings:

- Significant gains were found in the goal 1 sub-score (*The Magnitude of Fractions*) for Cohort 1 in both Year 1 and Year 2 performance.
- Cohort 2 students (whose teachers were not exposed to treatment until Year 2) had significant gains in goal 1 sub-scores (*The Magnitude of Fractions*) in Year 1. In Year 2, Cohort 2 students had significant gains in all scores except for goal 2 sub-scores (*Division with Remainders*).
- For treatment versus control students, no significantly higher scores were found for treatment students compared to control students, as the control students had higher mean scores and greater gain scores than the treatment students.
- Comparing the gain scores, or the extent to which performance changed from pre- to post-test, the gains for control students exceeded those for treatment students for all scores.

These findings indicate that both Cohort 1 and 2 treatment students had significant gains in their goal 1 sub-scores (*The Magnitude of Fractions*) for both Year 1 and 2. In addition, Cohort 2 students also had significant overall RIT scores and goal 3 sub-scores (*Broken Calculator*) in Year 2. Teachers in the Cohort 2 group took the *Broken Calculator* course in fall 2004. The Cohort 2 group had an entire year to improve their scores from pre- to post-test, whereas the Cohort 1 students were tested on gains over half a year, which could account in some measure for the additional gains among Cohort 2 students. Despite the involvement of teachers in the *Broken Calculator*, student performance among treatment students was strongest for the goal 1 sub-scores related to the *The Magnitude of Fractions* course.

The promising yet mixed findings for the Swampscott treatment students are tempered by the results for the control or comparison group, as the gains for the control students exceeded those for the treatment students for all scores.

## CROSS-SITE ANALYSIS

A variety of tests were employed to look for significant differences among several factors both within and across the four districts involved in the study:

- Student grade level
- Student ethnic group
- Sex of student
- Correlation of student gain scores with teacher gain scores and teacher experience teaching math

Student grade level: An analysis of variance was performed and the results indicated there were no significant differences (at the $p < 0.05$ level) in gain scores by grade level within any of the four districts in the study.

Student ethnic group: An analysis of variance was performed and the results indicated that for all scores and for all towns, topics and test subjects, there were no significant differences (at the $p < 0.05$ level) between any ethnic group gain scores.

Sex of student: Analysis of gain scores of students with respect to the variable sex was performed by an independent samples t-test. The results showed only two significant findings: In Mohonasen, males had higher mean gains ($p < 0.5$) than females on the RIT and goal 2 (*Division with Remainders*) sub-scores. In Dover, females had higher mean gains ($p < 0.5$) than males on the goal 1 (*Patterns and Functions*) sub-scores. All other gains were non-significant, although males in general had higher gains than females.

Correlation of student gain scores to teacher gains and teacher experience: A correlation analysis was used to examine the degree of relationship between student gain scores and gain scores of teachers. The very low correlation values, ranging from 0.021 to –0.19, indicate that student gains and teacher gains, and student gains and teachers' math teaching experience were unrelated.

## SUMMARY OF STUDENT FINDINGS

The Pre-Algebra strand students in Dover saw significant gains in all scores during the 2004 – 2005 school year, regardless of when their teachers took Seeing Math Elementary courses. The persistence of gains for Cohort 1 students, and the increase in gains for Cohort 2 students in Year 2 over Year 1 are promising as well. However, there were no comparison or control groups in Dover. Would these positive findings hold up if there were a control group with which to compare gain scores? Dover was the only site following the Pre-Algebra strand. Is there some other variable in Dover that is influencing student gains? Testing these materials again in a variety of sites and with control or comparison groups would answer these questions and allow us to say more definitively if Seeing Math Elementary can have a significant impact on student gains in pre-algebra.

It is difficult to interpret the findings in the Number and Operations strand. In Champaign, there were no significant student gains over time for Cohort 1 students, nor any significant differences between control and treatment students. In Mohonasen, treatment students significantly outperformed control students, with very high gains for the goal 3 (*Broken Calculator*) sub-scores in particular. Mohonasen had only one treatment-control teacher pair. Was there some other variable at the individual teacher level that influenced student gains?

In Swampscott, both Cohort 1 and 2 students had significant gains in the goal 1 (*The Magnitude of Fractions*) sub-score, even though Swampscott teachers took the *Broken Calculator* course. Was there some other variable in Swampscott that influenced student gains in the goal 1 sub-score? Swampscott control groups significantly outperformed treatment groups. Had control teachers prepared their students for testing?

In Champaign and Swampscott, control teachers performed as well as or better than treatment teachers, based on their pre- to post-test gain scores. It is possible that in these sites, the control teachers consciously or sub-consciously prepared their students for these tests with additional coaching. This John Henry effect is frequently seen among control groups who want to compete to out-perform the treatment group. In this age of teacher accountability, it is certainly understandable that teachers would not want to have any empirical evidence that their students had not improved in their mathematics skills over the course of a school year. Because teachers had access to the student tests before they administered them, it would be potentially tempting to "teach to the test."

## RECOMMENDATIONS AND FUTURE EVALUATION STRATEGIES

The questions raised by the somewhat confusing student findings lead us to a number of recommendations for future evaluation strategies to demonstrate the efficacy of Seeing Math Elementary for increasing elementary students' math proficiency.

First, studies should be replicated with control or comparison groups, within and across districts, to more clearly determine whether gains are due to involvement with Seeing Math Elementary, or some other variables. If possible, student pre- and post-tests should be administered by someone other than the classroom teacher to minimize the possibility of John Henry effects.

The research design itself cuts a broad swath by focusing on entire courses. If and when there are student gains, it is difficult to determine what, if any, aspect of the professional development intervention is directly related to the specific student gains. We recommend a series of small, intensive design studies that focus on specific aspects of the Seeing Math Elementary approach or specific learning objectives so that models for effective teacher professional development can be built from the findings.

We further recommend conducting focused action research projects to provide additional sources of demonstrable change in both student achievement and math content knowledge as well as teacher pedagogic change. These action research projects should take into account the context of the unique learning environments in which they occur, and offer additional quantitative student achievement data input to evaluating the effect and impact of Seeing Math Elementary to support and confirm the results of the overall research study. In turn, the action research projects developed by the teachers themselves would provide another window of learning into the teachers' own professional development as they take ownership of the research process.

We recommend focusing additional research and evaluation attention on the development, testing and validation of student assessment instruments for both mathematics content and process standards. Designing assessment instruments that embrace process as well as content standards and align closely with the Seeing Math Elementary approach, then testing and validating these instruments in real classrooms with real students, would provide an invaluable asset to the mathematics education community and a corollary component to the Seeing Math Elementary program rated so favorably by the teachers who participated in it.

## EVALUATION LESSONS LEARNED

There were multiple partners involved in the Seeing Math Elementary evaluation – with Concord Consortium developing the research design, Northwest Evaluation Association designing some of the data collection instruments under Concord Consortium's direction and model, Concord Consortium developing some of the data collection instruments and scoring conventions, Concord Consortium collecting the quantitative data, Concord Consortium and NWEA scoring and coding some of the quantitative data, and then handing off all the data to the external evaluator, Edcentric, to analyze. This is a design fraught with peril, no matter how well-intentioned each of the partners may be, as intensive coordination and communication are necessary to ensure a cohesive effort with a coherent evaluation product. This design may not be the most efficient or effective way to conduct a multi-faceted external evaluation.