Evaluation of the Greater Birmingham Mathematics Partnership (GBMP): A Comprehensive Approach to Evaluating Changes in Teachers' Mathematics Understanding

Rachel D. Cochran
Center for Educational Accountability
University of Alabama at Birmingham

Philosophy and History of the GBMP

Middle grades teachers are expected to teach rich and challenging mathematics in such varied areas as probability and statistics, geometry, proportional reasoning, and algebra—mathematics that most middle school teachers have not had adequate opportunities to learn. Over the course of many discussions among a gathering of a handful of mathematics teachers, curriculum experts, school administrators, and mathematics and education faculty members in the Birmingham area, it was determined that a common and critical need in middle school mathematics education in Birmingham was to deepen middle school mathematics teachers' knowledge and understanding of mathematics content. Thus began the Greater Birmingham Mathematics Partnership.

Strong community interest generated local funding in the spring of 2002 to support a pilot of this project. The first mathematics content course, *Patterns, Functions, and Algebraic Reasoning*, was offered for K-12 teachers in the summer and filled immediately. A second course was added and promptly filled. During 2002-2003, several school systems funded a second summer of content courses in 2003. Again, the summer courses were oversubscribed. These pilot efforts served to strongly confirm and establish the need for the extension and expansion of GBMP.

Nine Birmingham-area school systems now serve as part of the partnership as well as, Birmingham-Southern College (Division of Science and Mathematics and Division of Education), the University of Alabama at Birmingham (School of Engineering, School of Natural Sciences & Mathematics, and School of Education), and the Mathematics Education Collaborative (MEC). The goals of GBMP are: (I) to increase the effectiveness of middle school mathematics teachers within GBMP school systems; (II) to increase the leadership capacity of middle school mathematics teachers within GBMP school systems; (III) to unite the GBMP stakeholders (teachers, administrators, parents, IHE's and the public) in support of mathematics education programs that are high quality and effective; and (IV) to increase the mathematics achievement of all middle school students in GBMP schools and reduce discrepancies in disaggregated mathematics achievement data within these schools.

The fundamental outcome of this project is to support teachers' mathematics skills and understanding, which is projected to result in changes in students' mathematics skills and understanding. It is hypothesized by the partnership that deepening mathematics

teachers' content knowledge is a key component in improving classroom instruction in mathematics, thus improving the mathematics achievement of students.

GBMP's professional development curriculum for teachers involves a sequence of seven intensive mathematics content courses taught or co-taught by MEC staff each summer. The mathematics content consists of the "big mathematical ideas" of numerical reasoning, algebra, geometry, probability, and data analysis as identified in NCTM's *Principles and Standards for School Mathematics*. The focus is on developing conceptual understanding of the mathematics as well as the ability to put mathematical ideas and skills to work in solving complex and relevant problems. All courses attend to the process strands of problem solving, reasoning, making connections, and communicating.

The courses model a learning environment that optimizes the learning of quality mathematics and will meet a broad range of learner needs. They allow access for those teachers who fear and/or dislike mathematics, yet challenge all participants. The courses offer teachers opportunities to struggle with complex, rich, and expandable mathematical tasks with the potential of arriving at the development of concepts that are foundational to the field of mathematics.

The project's content courses and follow-up sessions are designed to develop a deep understanding of the mathematics that middle school students should know, as well as a deep understanding of how students learn mathematics, of optimal assessment practices, and of ways to create mathematics classrooms that are powerful learning communities. The GBMP equips teachers to teach in a way that makes courses challenging. Courses become challenging when a teacher uses in-depth content knowledge and inquiry-based pedagogical knowledge and skills to make them challenging. Evidence from MEC's previous NSF grant strongly suggests that when teachers experience all of this as mathematics learners themselves, they feel compelled to offer the same to their students.

Evaluation of the GBMP

The project's evaluation design is based on the realization that comprehensive mathematics education reform encompasses a set of interventions with short-term impacts that cumulatively influence the context within which mathematics is learned and experienced over the long term. The evaluation logic model being applied is a traditional context, input, process and outcome system that is supplemented by research studies based on principles and theories of learning and development.

The focus of this paper and a significant portion of the evaluation of the project is on measuring the effectiveness of the professional development provided through the GBMP summer courses in improving the mathematics understanding of middle grades teachers in the first course of the 7-course series.

One of the most challenging aspects of evaluating the effectiveness of professional development in mathematics education is finding or developing instruments to measure changes in teachers' content knowledge. Typically, content knowledge in mathematics is

viewed rather narrowly as an ability to arrive at an accurate answer to a mathematical problem. Very rarely is there an interest in examining the problem-solving process, including misconceptions and verification, as a pathway to understanding. This evaluation takes a more expansive view of mathematical content knowledge and uses a variety of data sources that provide a more complete picture of teachers as learners and teachers of mathematics than what could be gleaned from a multiple choice test.

The study included middle grades teachers enrolled in the first of a series of 9-day intensive summer courses titled Patterns, Functions, and Algebraic Reasoning. In that course, teachers engaged in inquiry-based mathematics investigations. The instruction in the course was designed to further teachers' conceptual understanding, procedural fluency, and strategic competence in mathematics while modeling the pedagogy touted to be effective at leading to mathematical understanding in students.

Assessment of teacher growth during the courses modeled best assessment practices that result in improved student performance. The teachers kept mathematics learning logs and used rubric-based scoring guides to assess the quality of their work and that of their colleagues. Teachers prepared portfolios of their work, completed a pre and post course performance assessment, and took an objective test before and after the course.

Methodology

Objective Assessment of Content Knowledge

The Learning Mathematics for Teaching (LMT) project at the University of Michigan has developed and made available several sets of items designed to assess content knowledge for teaching mathematics. Currently, items are available for elementary mathematics and middle school mathematics in the areas of (a) patterns, functions, and algebra; (b) geometry, and (c) numerical reasoning. The evaluation staff worked with Mathematics Education Collaborative (MEC) instructors to select items from the item pools that best matched the Patterns, Functions, and Algebraic Thinking course taught in Summer 2005. A few items were considered a very good match with course content, some items matched moderately well, and other items did not match at all. It was decided to use only items that fell into the first two categories and exclude those items that did not match. Although the match in some cases was only moderate, it was felt that the resulting scale should show improved content knowledge as a result of the course.

Using the item parameters provided by LMT staff, the resulting test (comprised of 36 items selected as described above) was analyzed. Although the LMT items are grouped into scales according to the areas listed above, each of which has been extensively analyzed, the LMT project permits use of selected items to build scales better suited to particular circumstances. They do recommend, however, that when building the scales, attention be paid to the amount of test information provided by the scale. It was recommended that any scale developed achieve a test information value of at least 2 for ability levels between -2 and 2 to ensure that the scale has reasonable reliability. The 36-

item scale developed jointly by MEC and the evaluation staff met this criterion and was, therefore, considered reliable enough to be used in the evaluation.

This scale, the Content Knowledge for Teaching Mathematics – Patterns (CKTM-Patterns) was used an objective measure of increased content knowledge that occurred as a result of participation in a Summer 2005 course. The CKTM-Patterns was administered to middle school teachers and preservice teachers as a pretest at the orientation session that preceded each of the course sessions. The same instrument was administered again on the last day of the course to the same individuals. The pretests and posttests for individuals were then paired so that individual progress could be examined. For further evidence of reliability, a Cronbach alpha was calculated to determine the internal consistency of the 36-item test. The test showed good internal consistency (alpha = .82).

Following the Year 1 administration, the CKTM was modified. The evaluation team met with the course instructors and project management team to review the items again for content validity. It was hypothesized by the team that the test was taking participants too long to complete. As well, some items were viewed by the instructors as unrelated to course content. As a result, 7 items were removed from the test. The test maintained its integrity, and the new set of 29 items still showed good internal consistency (alpha = .80)

After further deliberation with the project team and instructors, two open-ended items were selected from a group of items developed by Nanette Seago, Co-PI, Turing to the Evidence and Video Cases II Projects, WestEd. These two items required participants to generate an equation to describe a situation. The items were determined by the team and instructors to be an excellent fit with course content and were added to the CKTM. The Year 2 CKTM-Patterns contained 31 items.

Performance Assessment

During the summer courses, instructors administered a performance assessment at the beginning of the course and again at the end of the course. During Year 1, for courses in the first session, two different assessments were administered and could not, therefore, be compared. During the second and third session courses of Year 1, however, the same assessment was administered pre and post. Evaluators used the Oregon Department of Education Mathematics Problem Solving Official Scoring Guide as their rubric to assess four domains (conceptual understanding, processes & strategies, communication, and accuracy). Only those assessments for which both the pre-task and the post-task were available were scored.

Because of the inability to use any assessments from the first session and the lack of matching pre and post assessments, only 16 performance assessments were scored in Year 1. In Year 2, over half of the performance assessments have been scored to date (n=48), and the evaluation team anticipates having over 80 assessment tasks.

Although the rubric has been used successfully in Oregon, the evaluators took precautions to ensure that there was consensus on the performance assessment scores. Three scorers applied the rubric individually to each of the assessments then met to discuss their scores and to resolve any discrepancies. The individual results were in good agreement in most cases, and in all cases, the consensus scores were easily determined.

Course Portfolio

As part of the 2005 MEC summer courses each participant kept a portfolio of completed tasks, assignments, and reflections. The portfolios were assessed in the spring of 2006 using a rubric designed by the evaluation team that was based on the project definition of Challenging Courses and Curriculum. Three members of the evaluation team piloted the rubric on five portfolios, and in January 2006, the rubric was reviewed by the rest of the GBMP team. Slight modifications were made as a result of both the pilot exercise and input from the team.

Five portfolios from middle school teachers were randomly selected from each of the three summer course sessions, resulting in a total of fifteen portfolios to be scored. The same process of reaching consensus on scores that was applied to the performance assessments was used with the portfolios.

Behavioral Checklist

The CEA is in the process of developing a behavioral checklist for use in observing teachers as they participate in the GBMP courses. The checklist contains four dimensions—deepening mathematics understanding, productive disposition, inquiry and reflection, and communication. These four dimensions are the same dimensions that form the GBMP's definition of Challenging Courses and Curriculum. For each dimension, there are four observable behaviors that were determined to be evidence of each dimension by the evaluation team.

The checklist was piloted during the Summer 2006 course. Three course participants were selected randomly to be observed on three occasions throughout the 2-weeks.

RESULTS

Content Knowledge for Teaching Mathematics-Patterns

The CKTM-Patterns was used to monitor the content learning gains of teachers during the 2005 and 2006 summer classes. In Year 1, pretest and posttest paired results were available for 62 practicing teachers and 4 preservice teachers, eleven of whom were recruited to serve as Math Support Team (MST) Teachers after the summer course.

The baseline performance of MSTs was higher than that of other teachers and of the preservice teachers. On average, preservice teachers and MSTs passed one additional item on posttest than they passed on pretest. Although other participants started lower than MSTs, on average, they gained approximately two points between pretest and posttest.

In Year 1, raw scores for n = 66 teachers who had both pre-test and post-test scores were examined using the effect size index. The teachers' pre-test scores served as their own controls. The difference between the post-test score arithmetic mean and the pre-test score arithmetic mean formed the numerator, and the standard deviate of the pre-test score was used as the denominator. These statistics were: xbar_pre = 21.08, xbar_post = 23.20, and sigma_pre = 5.82.

The differences in means between posttest and pretest yielded an effect size of .34, which falls midway between the effect size index for "small" effects, **d** = 0.2, and the index for "medium" effects, **d** = 0.5, (Cohen, 1988). Educational research often considers effect sizes in excess of .33 to be practically meaningful.

Table 1: Year 1 CKTM Pre-Post Results

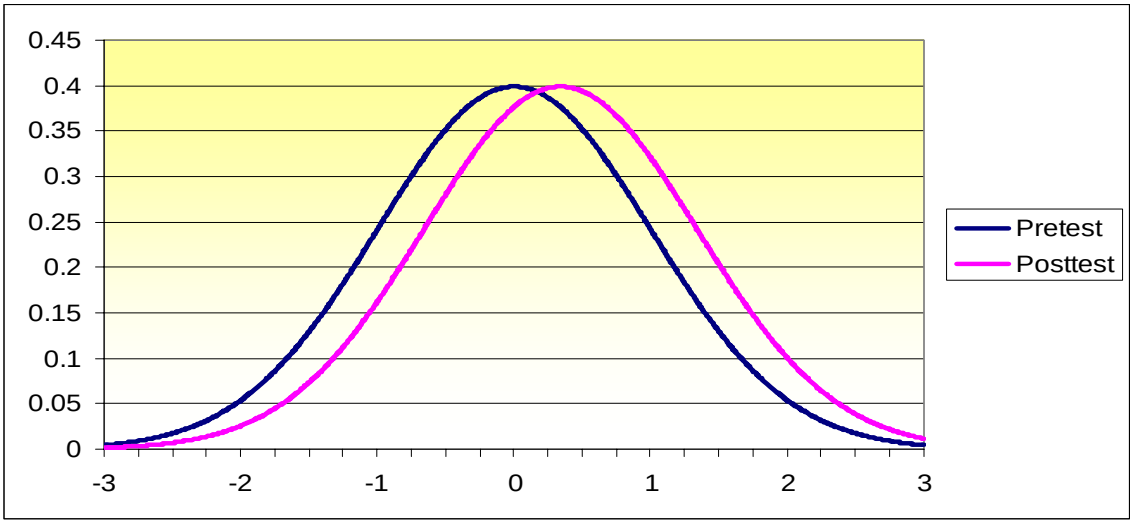
GROUP		Pretest	Posttest
Preservice	Yes *(N=4)*	19.75	20.75
	No	21.5	23.4
MST	Yes *(N=11)*	23.0	23.82
	No	21.05	23.1
ALL *(N=66)*		**21.08**	**23.2**

CKTM (36 items)

For additional interpretation, three different **U** statistics were tabled. These are based on the probabilities derived by the area under the standard normal curve. The $\mathbf{U_3}$ statistic has the property that it is equal to the cumulative probability under the normal curve from negative infinity to the calculated effect size index.

In Year 1, when **z** = 0.34, that probability is 0.6331. In other words, the upper-half of the post-test score population exceeds 63.31% of the pre-test score population.

Figure 1: Year 1 Pre-Post CKTM U Statistics



Several factors have been considered as potentially suppressing post scores on the CKTM. Project management and the evaluation team met to analyze items on which performance decreased from pre-test to post-test and found a number of problems. Some items were deemed to have more than one possible correct response. Others were problematic only for 5th grade teachers, most likely because the content of the items were not directly relevant to the content they teach.

It has been argued by project management and GBMP course instructors that the fit between the course content and the items on the CKTM is less than ideal. In an attempt to improve the fit of the items to the Patterns course, the evaluation team examined an additional set of items for potential use as an augmentation of the current pre-post measure. While not all of the items were accepted, the project management and evaluation team agreed to remove some items and add some new items for Year 2. Year 2 CKTM results are presented below.

Table 2: Year 2 CKTM Pre-Post Results

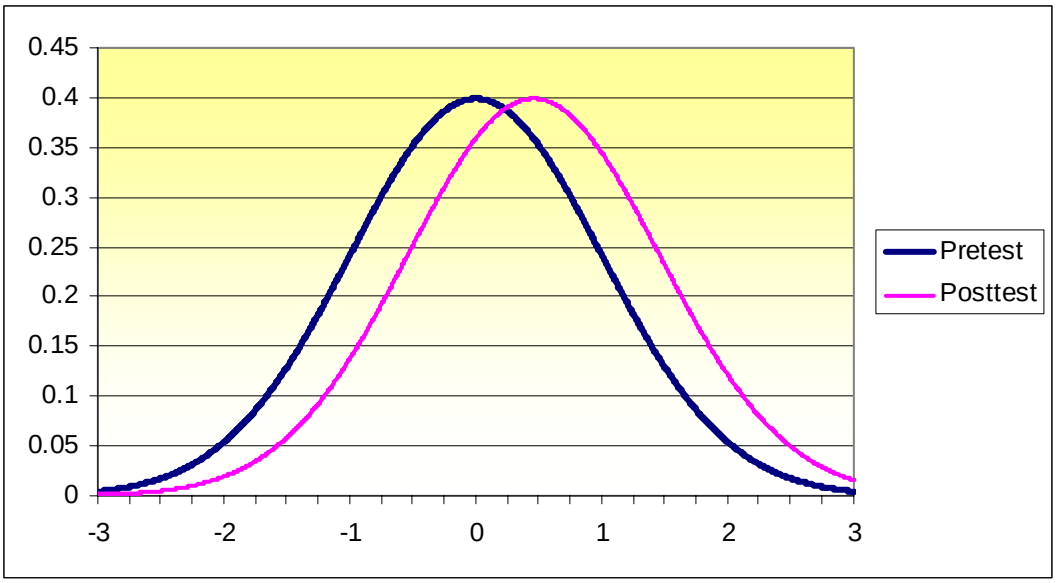
GROUP	Pretest	Posttest
All Participants *(N=98)*	16.04	18.78

CKTM (31 items)

In Year 2, raw scores for n=98 teachers who had both pre-test and post-test scores were examined. The difference between the pre-test score arithmetic mean (16.04) and the post-test score arithmetic mean (18.78) yielded an effect size of .46. This index falls at the high end of the effect size index for small effects, **d** = 0.2, and close to the index for medium effects, **d** = 0.5, (Cohen, 1988). When **z** = 0.46 that probability is 0.6772. In

other words, the upper-half of the post-test score population exceeds 67.72% of the pre-test score population. In other words, the upper-half of the post-test score population exceeds 67.72% of the pre-test score population.

Figure 2: Year 2 Pre-Post CKTM U Statistics



Because the last two items were such a good fit with course content and because they were open-ended items, they were looked at separately. For n=101 teachers, on the first of the two items, there was over a 30% increase in the number of teachers who responded correctly from the pre-test to the post-test. On the second item, there was close to a 40% increase in the number of teachers who responded correctly.

Table 3: CKTM Pre-Post Scores on Open Ended Items

CKTM Item Number	% Teachers with Correct Response on Pre-Test	% Teachers with Correct Response on Post-Test
30	29.7%	63.4%
31	33.7%	73.3%

One additional change noted in participants' responses to the last two CKTM items in Year 2 was an increase in the amount of work shown by the participants from pre-test to post-test. On the pre-test, some participants wrote nothing, others wrote only an equation with no explanation, and others scribbled a few notes as they tried to solve the problems. On the post test, most participants recorded their problem-solving process in enough detail so that an outside evaluator could follow their thinking from translating the problem from pictures to numbers to generating an equation to capture the pattern presented. It is believed by the partnership that this represents powerful evidence of change, and the evaluation team is currently in the process of developing a rubric to capture that change more systematically.

Performance Assessment

In Year 1, sixteen teachers were scored on a pre-post performance assessment task similar to the tasks used as part of the 2005 summer courses. These tasks were evaluated using the Oregon scoring rubric for performance assessments in mathematics. Results of the assessments are presented below.

Results revealed that even at pretest, accuracy was quite high, although accuracy still improved by posttest. Significant and impressive growth occurred in all other areas. The management team and evaluation team discussed whether these results were a function of "teaching to the rubric" and identified potential mechanisms for minimizing this threat to validity in subsequent summer course assessments.

Table 4: Year 1 Performance Assessment Pre-Post Score Frequencies

Score	1		2		3		4		5	
	N Pre	N Post	N Pre	N Post	N Pre	N Post	N Pre	N Post	N Pre	N Post
Processes and Strategies	5	0	6	0	3	1	2	7	0	8
Conceptual Understanding	3	0	8	0	2	2	2	5	1	9
Communication	4	0	6	0	4	2	1	6	1	8
Accuracy	4	1	N/A	N/A	N/A	N/A	10	2	2	13

The evaluation team is currently in the process of scoring the performance assessments from Year 2, the Summer 2006 Patterns courses. To date, over half of the performance assessment tasks have been scored. The table below reflects results for those 48 assessments.

Table 5: Year 2 Performance Assessment Pre-Post Score Frequencies

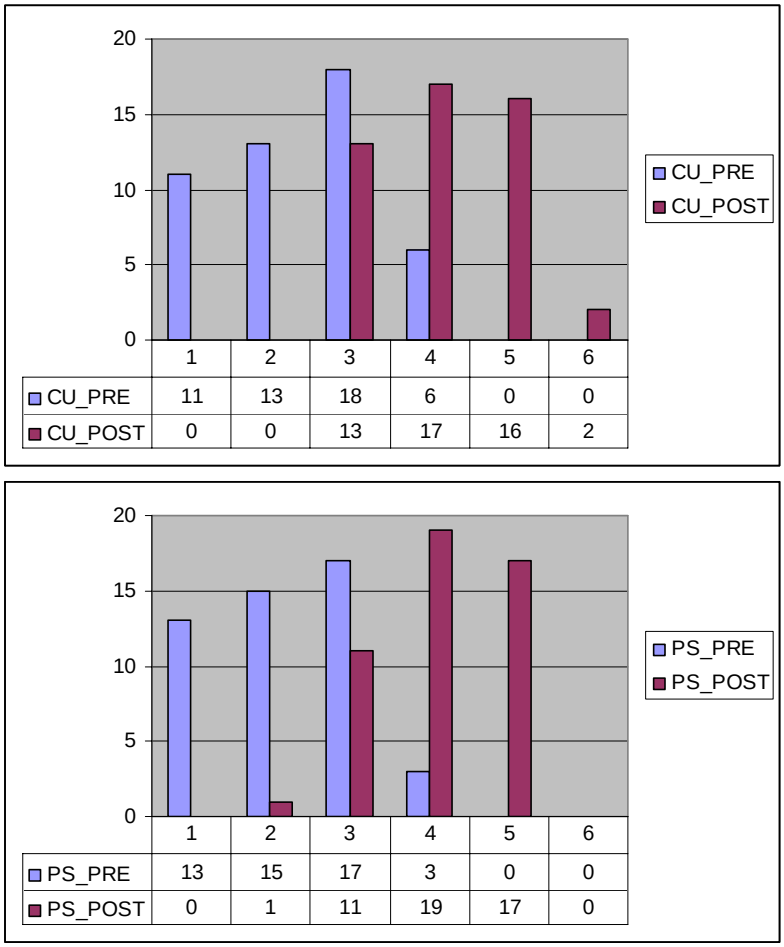
Score	1 – 1.5		2 – 2.5		3 – 3.5		4 – 4.5		5 – 5.5		6	
	N Pre	N Post	N Pre	N Post	N Pre	N Post	N Pre	N Post	N Pre	N Post	N Pre	N Post
Conceptual Understanding (CU)	11	0	13	0	18	13	6	17	0	16	0	2
Processes & Strategies (PS)	13	0	15	1	17	11	3	19	0	17	0	0
Communication (C)	15	0	19	1	11	11	2	20	1	12	0	4
Accuracy (ACC)	18	5	N/A	N/A	N/A	N/A	23	15	7	28	N/A	N/A

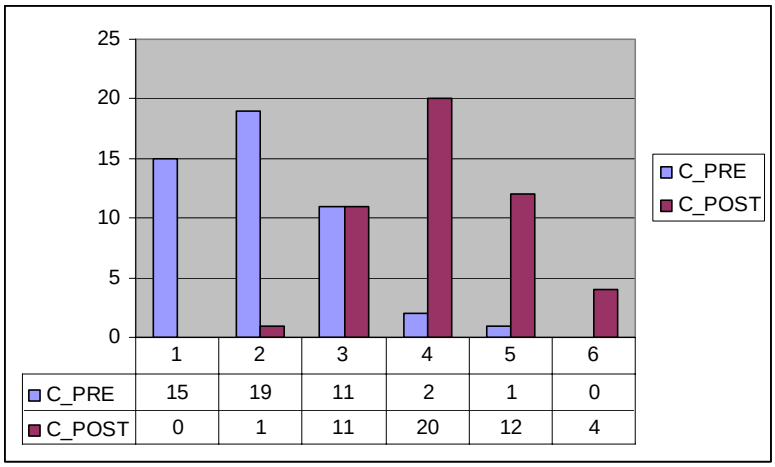
Table Year 2 Performance Assessment Pre-Post Median Scores

N=48	Conceptual	Processes and	Communication	Accuracy

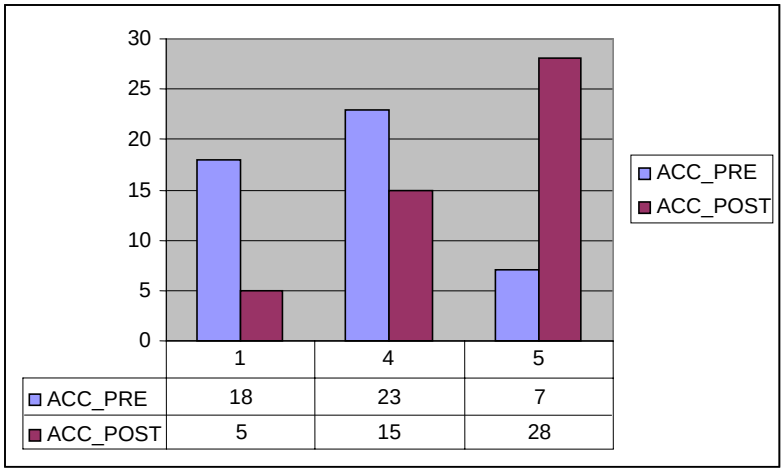
	Understanding (CU)		Strategies (PS)		(C)		(ACC)	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Median Score	2.750	4.000	2.000	4.000	2.000	4.000	4.000	5.000

Figure 3: Year 2 Performance Assessment Pre-Post Score Frequencies by Rubric Dimension





	1	2	3	4	5	6
C_PRE	15	19	11	2	1	0
C_POST	0	1	11	20	12	4



	1	4	5
ACC_PRE	18	23	7
ACC_POST	5	15	28

A Wilcoxon Signed Rank Test showed significant improvement on all four dimensions from pre to post administration. Participants showed marked improvement in the areas of conceptual understanding, processes and strategies, and communication, and some improvement on accuracy, where there was little room for improvement because of the limited range of rubric scores for that dimension.

Table 6: Year 2 Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
POST Conceptual Understanding - PRE Conceptual Understanding	Negative Ranks	2	7.25	14.50
	Positive Ranks	44	24.24	1066.50
	Ties	2		
	Total	48		
POST Processes & Strategies - PRE Processes & Strategies	Negative Ranks	1	9.50	9.50
	Positive Ranks	45	23.81	1071.50
	Ties	2		
	Total	48		
POST Communication - PRE Communication	Negative Ranks	1	1.00	1.00
	Positive Ranks	47	25.00	1175.00
	Ties	0		
	Total	48		
POST Accuracy - PRE Accuracy	Negative Ranks	3	12.50	37.50
	Positive Ranks	28	16.38	458.50
	Ties	17		
	Total	48		

Test Statistics(b)

	POST Conceptual Understanding - PRE Conceptual Understanding	POST Processes & Strategies - PRE Processes & Strategies	POST Communication - PRE Communication	POST Accuracy - PRE Accuracy
Z	-5.777(a)	-5.848(a)	-6.053(a)	-4.198(a)
Asymp. Sig. (2-tailed)	.000	.000	.000	.000

a Based on negative ranks.
b Wilcoxon Signed Ranks Test

Portfolios

Other evidence of teacher understanding can be derived from an analysis of portfolios generated during summer courses. Teachers participating in the 2005 summer course completed portfolios as part of their participation. Portfolios contained the following components:

- reflective pieces on the teacher as a learner of mathematics and a teacher of mathematics
- letter to someone addressing the "big ideas" from the course and changes anticipated to be made in the classroom
- pre-assessment task
- self-selected "most important piece of work" from the course
- scored task from the course
- teacher-selected task believed to reflect the teacher as a learner of mathematics
- assigned assessment task from the course.

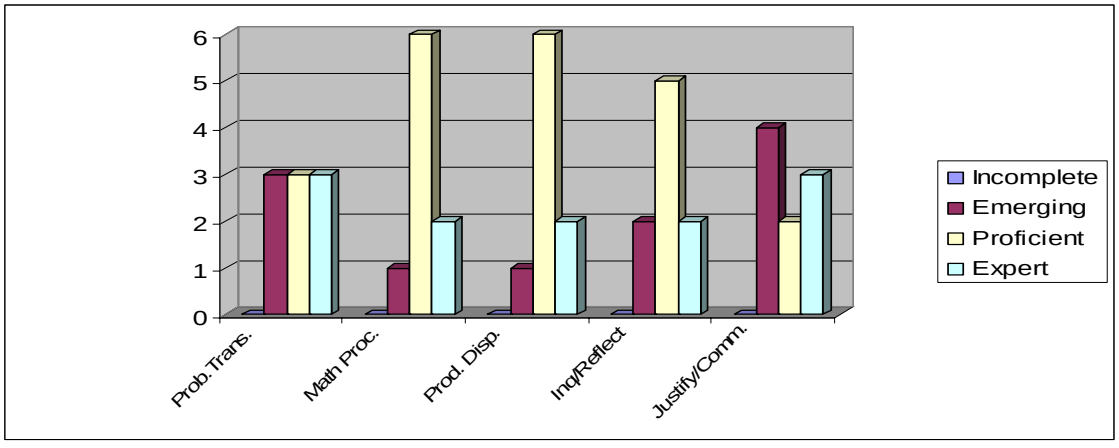
A sample of portfolios was rated holistically in terms of five features, all related to the challenging courses and curriculum dimensions delineated previously by the design team. The first two dimensions relate to the CCC feature of deepening knowledge of big ideas. The other dimensions directly relate to the remaining three components of challenging courses and curriculum. Results of consensus judgments among three raters are provided in the table below. These ratings indicate that more than half of the sample demonstrated performance that was at or above the proficient level on each dimension.

Year 2 portfolios have not yet been scored.

Table 8: Year 1 Portfolio Scores

	Mean	SD	N Incomplete	N Emerging	N Proficient	N Expert
Problem Translation	3.07	0.78	0	4	6	5
Mathematical Procedures	3.13	0.62	0	2	9	4
Productive Disposition	3.07	0.77	1	1	9	4
Inquiry and Reflection	2.93	0.85	1	3	7	4
Justification and Communication	2.8	0.91	1	5	5	4

Figure 4: Year 1 Portfolio Scores



Behavioral Checklist

The checklist was piloted in Year 2, in the Summer 2006 Patterns course. Three course participants were chosen at random to be observed three times over the 2-week course—once on the first day, once on the fourth day, and once on the eighth day. Observations took place when participants were working in groups or working on menu tasks with other participants.

For each of the three participants, meaningful change occurred on all four dimensions.

Table 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1000 1000

342 402 587 867 711 1272 862 332 543 543 711 867 361 480 361 689

711 711 711 711 711 711 711 711 711 711 402 402 867 867 867 617

964 776 762 724 830 683 650 811 837 546 555 771 637 948 847 850

733 850 782 710 682 812 764 1128 764 737 692 543 689 543 867 711

711 668 699 588 699 664 422 699 712 342 403 671 342 1058 712 686

699 699 497 593 456 712 649 979 669 651 597 711 543 711 867 1000

711 1000 332 711 587 1049 711 711 711 1777 710 543 1135 1000 692 1000

1000 332 332 587 587 711 711 1000 711 964 593 543 1068 1000 597 737

342 402 711 711 711 711 543 711 711 964 598 850 867 480 964 711

587 867 598 598 711 721 711 361 711 598 598 850 1182 1182 1182 617

776 776 776 776 776 776 1094 724 683 683 683 683 546 546 546 546

830 847 850 850 850 850 850 867 850 812 812 812 812 737 735 713

668 668 668 668 668 668 1018 588 664 664 664 664 342 342 342 342

679 712 686 686 686 686 686 867 686 712 712 712 712 651 699 651

]

endobj

11 0 obj

[

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1000 1000

352 394 459 818 636 1076 727 269 454 454 636 818 364 454 364 454

636 636 636 636 636 636 636 636 636 636 454 454 818 818 818 545

1000 684 686 698 771 632 575 775 751 421 455 693 557 843 748 787

603 787 695 684 616 732 684 989 685 615 685 454 454 454 818 636

636 601 623 521 623 596 352 623 633 274 344 592 274 973 633 607

623 623 427 521 394 633 592 818 592 592 525 635 454 635 818 1000

636 1000 269 636 459 818 636 636 636 1521 684 454 1070 1000 685 1000

1000 269 269 459 459 545 636 1000 636 977 521 454 981 1000 525 615

352 394 636 636 636 636 454 636 636 1000 545 645 818 454 1000 636

542 818 542 542 636 642 636 364 636 542 545 645 1000 1000 1000 545

684 684 684 684 684 684 984 698 632 632 632 632 421 421 421 421

775 748 787 787 787 787 787 818 787 732 732 732 732 615 605 620

601 601 601 601 601 601 955 521 596 596 596 596 274 274 274 274

612 633 607 607 607 607 607 818 607 633 633 633 633 592 623 592

]

endobj

14 0 obj

16 0 obj

<<

/Length 400

/Filter /FlateDecode

>>

stream

xレmRヒn、0□シ・ }・ ^ロ□c-
3ノ$メJ#E□nノ-シ・ □□□9-R□~m・ □EHィハクォォオルォルツ[R□‡廬$9摩□伝C・ トロIc□_k;k□醒チセ()ノO
Kア訛ィハ3・ ク・ iアウ・ □

f倚・ミ□ホ□ヨPTqｲ%"□ﾆﾀ(Kﾌ□F・ □ f

$・ﾅ・・ □ｲN2Aﾔ□

Discussion

[

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1000 1000

352 394 459 818 636 1076 727 269 454 454 636 818 364 454 364 454

636 636 636 636 636 636 636 636 636 636 636 454 454 818 818 818 545

1000 683 686 698 766 632 575 775 751 421 455 693 557 843 748 787

603 787 695 684 616 732 683 990 685 615 685 454 454 454 818 636

636 601 623 521 623 596 352 622 633 274 344 587 274 973 633 607

623 623 427 521 394 633 591 818 592 591 525 635 454 635 818 1000

636 1000 269 636 459 818 636 636 636 1519 684 454 1070 1000 685 1000

1000 269 269 459 459 545 636 1000 636 977 521 454 980 1000 525 615

352 394 636 636 636 636 454 636 636 1000 545 645 818 454 1000 636

542 818 542 542 636 642 636 364 636 542 545 645 1000 1000 1000 545

683 683 683 683 683 683 989 698 632 632 632 632 421 421 421 421

766 748 787 787 787 787 787 818 787 732 732 732 732 615 605 620

601 601 601 601 601 601 955 521 596 596 596 596 274 274 274 274

612 633 607 607 607 607 607 818 607 633 633 633 633 591 623 591

]

endobj

The evaluation team is using multiple measures to assess improved instructional practices as a result of improved teacher content knowledge. Classroom observations using the Reformed Teaching Observation Protocol have been conducted in middle school mathematics classes, and teachers have been asked to complete two different surveys that address this issue. As well, student data on state tests is being collected, and samples of student work are being analyzed.

Because the project is still in its early stages, some data, like student-level state testing data, is not available beyond what can be reported as baseline data. Other data, particularly survey data and observation data show promising results.

Observation Data

Middle school mathematics classes of teachers scheduled to participate in Summer 2005 courses were observed in Spring 2005, prior to their exposure to GBMP course content. Follow-up observations in the same teachers' classes as well as classes of other GBMP course participants are being conducted as this report is being written. Observations were conducted by two observers in each of 11 middle school classrooms during Spring 2005. Thirteen teachers were observed in Spring 2006 observations have been conducted in Spring 2006, four of which were classrooms of teachers observed in 2005.

The results indicated that while there was a considerable range of performance across the teachers in all subscales, the mean score was approximately 25-30% of possible points. These results suggested baseline classroom contexts that were largely teacher-directed, didactic, whole-group, and focused on delivery rather than inquiry and discourse.

The results for Spring 2006 are appear to show substantial improvement on each of the scales. The four observations that were follow-up observations (i.e., same teachers observed in 2006 that were observed in 2005) provide further evidence that classroom practice may be improving. Again, however, this small number of comparable observations limits the interpretability of the results. The table below shows the mean improvement on each scale for the three paired observations.

Survey Data

Multiple scales were developed by the evaluation team to measure the effectiveness of GBMP professional development, teacher knowledge and implementation of summer course content, aspects of proficient teaching of mathematics, and factors that impact implementation of inquiry-based instructional practices. Preliminary results from the 41 teachers who have completed surveys to date are presented in the following discussion.

MSTs and other teachers reported understanding inquiry-based instruction as well as GBMP course content and strategies but frequently reported needing planning time to support implementation in these areas. Teachers reported the greatest confidence with implementing inquiry based instruction. Teachers were least comfortable implementing student performance assessments and student portfolios.

Table 10: Effort of Implementation of Content and Pedagogy

AREA		No knowledge	Trying to learn so can feel comfortable enough to implement	Would apply if there was sufficient support	Understand and am applying but need planning time	Implement routinely with minimal effort	Not appropriate for the course I teach
Inquiry-based instruction	MST		27%		40%	33%	
	non-MST		19%		38%	43%	
GBMP course content and strategies	MST		27%	7%	40%	27%	
	non-MST		5%	30%	45%	15%	5%
Student performance assessments	MST	6%	25%	25%	19%	25%	
	non-MST		24%	38%	19%	19%	
Student portfolios	MST		63%	19%	6%	6%	6%
	non-MST	5%	29%	10%	29%	29%	

342 402 587 867 711 1272 862 332 543 543 711 867 361 480 361 689

711 711 711 711 711 711 711 711 711 711 402 402 867 867 867 617

964 776 762 724 830 683 650 811 837 546 555 771 637 948 847 850

733 850 782 710 682 812 764 1128 764 737 692 543 689 543 867 711

711 668 699 588 699 664 422 699 712 342 403 671 342 1058 712 686

699 699 497 593 456 712 649 979 669 651 597 711 543 711 867 1000

711 1000 332 711 587 1049 711 711 711 1777 710 543 1135 1000 692 1000

1000 332 332 587 587 711 711 1000 711 964 593 543 1068 1000 597 737

342 402 711 711 711 711 543 711 711 964 598 850 867 480 964 711

587 867 598 598 711 721 711 361 711 598 598 850 1182 1182 1182 617

776 776 776 776 776 776 1094 724 683 683 683 683 546 546 546 546

830 847 850 850 850 850 850 867 850 812 812 812 812 737 735 713

668 668 668 668 668 668 1018 588 664 664 664 664 342 342 342 342

679 712 686 686 686 686 686 867 686 712 712 712 712 651 699 651

]

endobj

11 0 obj

[

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1000 1000

352 394 459 818 636 1076 727 269 454 454 636 818 364 454 364 454

636 636 636 636 636 636 636 636 636 636 454 454 818 818 818 545

1000 684 686 698 771 632 575 775 751 421 455 693 557 843 748 787

603 787 695 684 616 732 684 989 685 615 685 454 454 454 818 636

636 601 623 521 623 596 352 623 633 274 344 592 274 973 633 607

623 623 427 521 394 633 592 818 592 592 525 635 454 635 818 1000

636 1000 269 636 459 818 636 636 636 1521 684 454 1070 1000 685 1000

1000 269 269 459 459 545 636 1000 636 977 521 454 981 1000 525 615

352 394 636 636 636 636 454 636 636 1000 545 645 818 454 1000 636

542 818 542 542 636 642 636 364 636 542 545 645 1000 1000 1000 545

684 684 684 684 684 684 984 698 632 632 632 632 421 421 421 421

775 748 787 787 787 787 787 818 787 732 732 732 732 615 605 620

601 601 601 601 601 601 955 521 596 596 596 596 274 274 274 274

612 633 607 607 607 607 607 818 607 633 633 633 633 592 623 592

]

endobj

14 0 obj

16 0 obj

<<

/Length 400

/Filter /FlateDecode

>>

stream

xレmRヒn、0□ｼ・ }・ ^ﾛ□c-
3ﾉ$ﾒJ#E□nﾉ-ｼ・ □□□9-R□~m・ □EHｲﾊｸｵｵｵﾙｵﾙｯ[R□‡廬$9摩□伝C・ﾄ□Ic□_k;k□醒ﾁｾ()ﾉO
Kｱ訛ｲﾊ3・ｸ・ iｱｳ・ □

342 402 587 867 711 1272 862 332 543 543 711 867 361 480 361 689

711 711 711 711 711 711 711 711 711 711 402 402 867 867 867 617

964 776 762 724 830 683 650 811 837 546 555 771 637 948 847 850

733 850 782 710 682 812 764 1128 764 737 692 543 689 543 867 711

711 668 699 588 699 664 422 699 712 342 403 671 342 1058 712 686

699 699 497 593 456 712 649 979 669 651 597 711 543 711 867 1000

711 1000 332 711 587 1049 711 711 711 1777 710 543 1135 1000 692 1000

1000 332 332 587 587 711 711 1000 711 964 593 543 1068 1000 597 737

342 402 711 711 711 711 543 711 711 964 598 850 867 480 964 711

587 867 598 598 711 721 711 361 711 598 598 850 1182 1182 1182 617

776 776 776 776 776 776 1094 724 683 683 683 683 546 546 546 546

830 847 850 850 850 850 850 867 850 812 812 812 812 737 735 713

668 668 668 668 668 668 1018 588 664 664 664 664 342 342 342 342

679 712 686 686 686 686 686 867 686 712 712 712 712 651 699 651

]

endobj

11 0 obj

[

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1000 1000

352 394 459 818 636 1076 727 269 454 454 636 818 364 454 364 454

636 636 636 636 636 636 636 636 636 636 454 454 818 818 818 545

1000 684 686 698 771 632 575 775 751 421 455 693 557 843 748 787

603 787 695 684 616 732 684 989 685 615 685 454 454 454 818 636

636 601 623 521 623 596 352 623 633 274 344 592 274 973 633 607

623 623 427 521 394 633 592 818 592 592 525 635 454 635 818 1000

636 1000 269 636 459 818 636 636 636 1521 684 454 1070 1000 685 1000

1000 269 269 459 459 545 636 1000 636 977 521 454 981 1000 525 615

352 394 636 636 636 636 454 636 636 1000 545 645 818 454 1000 636

542 818 542 542 636 642 636 364 636 542 545 645 1000 1000 1000 545

684 684 684 684 684 684 984 698 632 632 632 632 421 421 421 421

775 748 787 787 787 787 787 818 787 732 732 732 732 615 605 620

601 601 601 601 601 601 955 521 596 596 596 596 274 274 274 274

612 633 607 607 607 607 607 818 607 633 633 633 633 592 623 592

]

endobj

14 0 obj

16 0 obj

<<

/Length 400

/Filter /FlateDecode

>>

stream

xレmRヒn、0□シ・ }・ ^ロ□c-
3ノ$メJ#E□nノ-シ・ □□□9-R□~m・ □EHィハクォォルォルツ[R□‡廬$9摩□伝C・ トロIc□_k;k□醒チセ()ノO
Kア訛ィハ3・ ク・ iアウ・ □