Webb, D. C., Romberg, T. A., Ford, M. J., & Burrill, J. (2005). Teacher Collaboration: Focusing on Problems of Practice. In T. A. Romberg & T. P. Carpenter (Eds.), *Understanding Mathematics and Science Matters* (pp. 231-251). Mahwah, NJ: Lawrence Erlbaum Associates.

Chapter 10

Teacher Collaboration: Focusing on Problems of Practice

David C. Webb
Thomas A. Romberg
Michael J. Ford
Jack Burrill
University of Wisconsin–Madison

In 1997, a group of middle school teachers in the Prairie Creek Middle School contacted researchers from the National Center for the Improvement of Student Learning and Achievement in Mathematics and Science (NCISLA) and expressed an interest in improving their mathematics and science program. This chapter focuses on the resulting successful collaboration.

In September 1997, we held an initial meeting with the interested mathematics and science teachers. After negotiations with administration, teachers and researchers collaborated on plans for the spring semester. In spring 1998, we began to work directly with four teachers (three at Grade 6 and one at Grade 8), with each teacher implementing at least one new unit from the algebra strand of one of the NSF-funded mathematics curriculum projects, *Mathematics in Context* (MiC; National Center for Research in Mathematical Sciences Education & Freudenthal Institute, 1996–1998). That summer three additional teachers (two more from Grade 6 and one from Grade 7) joined the project, and we held a 2-day summer institute in preparation for the school year. Some of the teachers also chose to meet with us at other times to develop new units of instruction for their classrooms.

During the 1998–1999 school year, teachers again implemented at least one new unit, either MiC or Boxer-based units. (Boxer is a flexible software system that supports student exploration of mathematics or science.) We intended to focus on ideas of modeling and student thinking, but in order to address issues relevant to teachers' concerns, we focused interchangeably on teachers' pedagogical concerns and instructional development of mathematical content. During the 1999–2000 school year, five Grade 7 teachers (persuaded by the sole Grade 7 teacher from the previous year) joined the project, and meetings focused on the teaching of MiC algebra units and the development of Boxer-based motion units. Problem-solving assessments were administered to all participating students in September, January, and May in the second and third year of the collaborative.

During this 3-year period, the pre-algebra curriculum in Grades 6 and 7 changed from a tracked program with a fragmented collection of activities to a more coherent and comprehensive mathematics program for all students. Teachers drew on the instructional expertise of colleagues and the research perspectives of the university team to explore ways to "uplevel the mathematics standards" in the middle grades. As teachers discovered new instructional methods that worked with their students, they shared these experiences with their peers, and, in many cases, instructional adaptations spread through conversations, both formal and informal, focused on students' mathematical thinking. As technical knowledge and discoveries were shared, the instructional quality of participating teachers was influenced in powerful ways.

In the sections that follow, we discuss the background of this successful project, examine issues in teacher practice, and close by talking briefly about the resulting impact on student achievement.

BACKGROUND TO THE COLLABORATIVE: DISTRICT LEADERSHIP AND EMERGENCE OF TEACHER COMMUNITY

District Leadership

District administrators were anxious to implement the district's instructional standards, which teachers had crafted in 1997, but the middle-grades mathematics teachers had major concerns: Sixth-grade teachers wanted help building on the background of the many students who had used innovative curricula in elementary grades, and all the teachers wanted to be able to provide both quality experiences for "all students" (which meant the school's tracking of students for mathematics needed to be reconsidered) and appropriate pre-algebra experiences in Grades 6 and 7 in preparation for algebra in Grade 8.

The collaborative was initiated with the full support of the district administration (including the district Director of Instruction and the Learning Resource Coordinator) and the middle school principal. Administrators saw the district in the process of changing their mathematics program from a traditional arithmetic program with ability grouping, toward an integrated program with heterogeneous grouping. They believed in a long-term collaborative approach to change, expected teachers to work toward agreed-on goals over time, and trusted teachers to make informed, reasonable decisions about ways to proceed. The Learning Resource Coordinator directly assisted the teachers and kept them informed about activities and decisions. Administrators also believed that the collaborative would help teachers broaden their scope about mathematics and mathematics instruction, and that, as a consequence, teacher-leaders would emerge. These new leaders working with their colleagues would, in turn, contribute to a change in the culture of what it meant to be a professional teacher in the Prairie Creek schools.

The innovative activities of teachers in the collaborative were also supported by district administration through their hiring practices and communication of goals to teachers not in the collaborative. Several teachers remarked (verbatim), "This district hires teachers that are overachievers and expects teachers to figure' out the rest." This shared perception led to a shared expectation that teachers would continue to experiment with, identify, and work to achieve best practices in their classrooms.

This district-based supportive approach to change in mathematics instruction needs to be seen in terms of issues that were facing the school district. First, the district population had recently changed from a stable mixed rural and bedroom community to a rapidly growing suburban community with considerable social diversity. (This population growth resulted in the building of a new middle school, which opened after this collaborative project ended.) The need for teacher-leaders was seen as important to the development of coherence. Second, needed bond proposals were typically passed only after repeated rejections, and in this case, a small but vocal group of parents was not pleased with the changing demographics and the needed resources. In response, some members of the community created a charter elementary school using a back-to-basics curriculum. Third, this district ranked among the best in the state and the top in the region in mathematics on the Grade 8 state assessment. They saw maintaining that high rank in the state as the challenge and noted the need to implement instructional standards aligned with the state and national standards. Fourth, the district mathematics committee established a goal to provide all students with ideas and skills central to algebra by the time they finished the eighth grade. In the final year of the project, this "algebra for all" initiative was later modified due to the adoption of a 3-year, integrated approach to mathematics for Grades 8–10, intended to provide the equivalent of the traditional algebra, geometry, and advanced algebra courses by the end of Grade 10. This integrated math program was to be implemented starting with Grade 8 in fall 2000. Our work with teachers on pre-algebra instruction proved to be a critical step in facilitating high school curriculum decisions. Finally, although the district was focused on the mathematics program, the science program was undergoing a similar shift.

The district administrators, although perhaps unfamiliar with all the details of the collaborative project, saw that "the culture [of the math department] has changed," as a district administrator commented in an interview (April 18, 2000). Teachers took responsibility for developing a coherent mathematics program *collaboratively*, and they recognized the strong connections of the work at Grades 6 and 7 with the work both at the elementary grades and with the new integrated math program at Grades 8, 9, and 10.

We also note that the programs in both mathematics and science prior to this project could best be described as "fragmented." There were, for example, different learning goals in place for different students (approximately 25% of the students were taught algebra in Grade 8). The instructional programs of each teacher varied in content covered, text materials used, methodologies for determining achievement, and so forth.

Growth of Professional Community

The coherent community of teachers we saw developing was built by the teachers themselves from the bottom up. With no appointed leader taking steps to direct the activities of individual teachers, the community emerged as a result of teachers' efforts to reform classroom instruction. Primary among their reform efforts was the articulation of elements of effective instruction, which occurred as teachers were introduced to the MiC and Boxer materials and again when they later used the materials with students. These instructional materials directed the activities of both students and teachers in ways that emphasized collaboration and discussion rather than rote memorization of facts and individual deskwork. Teachers found that in order to manage and direct discussions, they had to focus on student thinking while keeping the instructional end points in mind, a task they admittedly found particularly challenging.

The collaborative managed monthly meetings of researchers and teachers that explicitly addressed salient elements of instruction. As teachers shared experiences and began to articulate issues that arose in their classrooms, they were provided a vocabulary with which to discuss classroom practice. Several teachers remarked that the articulation of these elements legitimized and "professionalized" the intuitive ideas that had guided their practice prior to their participation in the collaborative.

Teachers also began meeting informally in smaller groups to discuss their experiences teaching specific units. They found that this informal collaboration provided an important source of support in that teachers could provide each other useful feedback about specific issues and content. Also, because many teachers felt the switch to a new form of instruction "risky," their collaboration and "sharing" served to reassure them that their decisions were good ones.

This practice-centered informal communication also supported authority and leadership at the department level. As noted earlier, the middle school mathematics program was the focus of school reform throughout the collaborative. A goal of this reform was to provide students with ideas and skills central to algebra by the time they finished the eighth grade. Teachers who had rarely expressed opinions in such a forum began taking more active roles in the decision-making process. One teacher in particular organized all other teachers in one grade to teach at least partly from MiC units because she believed so strongly that the structure and content of the curriculum met their needs. We believe that the emergence of formal and informal leadership stemmed from teachers' developing sophistication about effective mathematics instruction.

Classroom observation reports and teacher interviews suggest that growth in teacher practice occurred within the context of partnerships, in which teachers voluntarily worked with another teacher or a researcher to explore a pre-identified area of study (e.g., algebraic representations, the study of motion, classroom assessment). Through these partnerships, teachers grew more innovative and overcame significant obstacles in achieving desired classroom practices. These partnerships provided a safe context for experimentation, focused attention to classroom practice, peer and researcher observation, and shared reflection. The challenges teachers faced in designing classrooms that promoted understanding were often defined according to their conceptions of mathematics, conceptions of student learning, and pedagogical abilities. These partnerships gave teachers a safe context in which to challenge these conceptions and reconstruct their classroom practice.

Although a main feature of this research project was the teacher–researcher collaborative, finding that most teachers valued collaboration with their peers is not remarkable. Yet it is worth noting that teachers did not always take advantage of opportunities for collaboration, even when schools supported such activities. During the last year of the project, for example, school administrators encouraged teachers to observe each other, but teachers chose not to do so. Structural issues also limited the collaboration. At the end of our work with this district, teachers continued to meet in informal ways, but the physical separation that occurred when the middle school district moved about half of the teachers to a new school decreased overall collaboration. The middle school also reorganized teachers into teams or "houses," designed such that all teachers within a house taught different subjects to the same students, thus structuring teacher collaboration around students, not around instructional content. Clearly, such a focus has its advantages, but the emergence of informal "teams" of teachers during the collaborative suggests that instructional content-centered collaboration is also helpful for teachers, particularly during reform efforts.

TEACHER FOCUS ON PROBLEMS OF PRACTICE: EVIDENCE FROM CASE STUDIES

During our work with teachers, the growth in their knowledge of student reasoning and mathematical content supported a shift in their instructional approach and practices away from mere topic coverage toward facilitation of "sense-making" by students. In the second and third year, two recurring themes emerged: the overcoming of the "risks" entailed when engaging in student-centered pedagogy and teacher interest in discussing and exploring methods to assess student understanding. To elaborate on these *problems of practice*, we draw here on case studies of participating teachers. In general, teacher participation in these studies required a 4-to-8-week partnering with a researcher, who administered additional interviews, completed up to four classroom observations per week, and supported teacher reflection on classroom events and examples of student work.

Emerging Issue: Working Toward Student-Centered Pedagogy

For learning with understanding to occur on a widespread basis, students need opportunities to (a) develop appropriate relationships, (b) extend and apply their mathematical knowledge, (c) reflect about their own mathematical experiences, (d) articulate what they know, and (e) make mathematical knowledge their own (Carpenter & Lehrer, 1999). To teach for understanding, teachers need to design learning environments in which students have these opportunities to make sense of content and come to "own" what they know and learn (Bransford, Brown, & Cocking, 1999). But even when teachers recognize the need to change classroom practices, they often encounter significant difficulty when they attempt to orchestrate learning opportunities that are more student centered (Fennema & Nelson, 1997; Fennema & Romberg, 1999). As the following case summaries demonstrate, however, even though these difficulties can be quite similar, the ways that teachers resolve them can be distinctively different.

The Case of Beth Resnick. Ms. Resnick had been a middle-grades teacher for almost 20 years. During the collaborative, she taught both mathematics and science. Her primary reason for participating in the collaborative was that she was dissatisfied with the science curriculum, which was organized by topic (e.g., "the solar system," "rocks and minerals," "weather"). According to Ms. Resnick, students did not have the opportunity to pose issues worth investigating or to make choices along the path of inquiry. The predetermined nature of the textbook (focused on sets of steps that students followed to generate conclusions from "experiments") removed any student motivation to engage in issues of interest or to develop deeper understanding of a textbook topic.

Participation in the collaborative led Ms. Resnick to shift from traditional, lecture-driven, direct instruction to pointed use of questioning, attention to student's ideas, and a demand that students articulate their own reasoning. In interviews, Ms. Resnick stated that this shift was not an easy one to make. Opening the lesson plan to the directions that students' ideas might take implied a certain amount of risk. She could not predict what students would say and often felt as though she had lost control of instruction. More generally, she found the new role as facilitator of discussions awkward, but found useful the sense of herself as a coach. With the new instructional style, however, Ms. Resnick found that students were learning more. By opening up discussions with the norm that all well-supported ideas were welcome, the variety of ideas presented increased greatly, students experienced more complexity, and what students articulated facilitated her identifying and helping students who were having trouble.

At the end of the collaborative, Ms. Resnick's goals were to develop better ways to encourage students to be aware of the reasons they believed what they did. She found it difficult to find ways to bring out in students a conscious awareness of their reasoning as well as to provide them with feedback that would direct them toward productive ways of thinking about a problem without giving them the answer. Ms. Resnick's work with the "physics of motion" unit illustrates some of the themes of this transition. In the interest of developing a unit that made clear the connection between motion and algebra, one of the authors of this chapter conducted weekly after-school Boxer enrichment workshops for interested sixth-grade students in the 1999–2000 school year. An adaptation of this sequence was then taught in the spring semester. During these sessions, students observed constantly accelerated motions, programmed these motions in Boxer, and drew representations of the motions with pencil and paper. Through their work, students faced important issues such as conceptualization of variables, connection of these variables to invented representations, quantification of these variables through measurement, and the separation of "signal" from "noise" in the resulting data.

Prior to participating in the development of the motion unit, Ms. Resnick had believed the best way to teach a physics concept such as acceleration was to introduce the concept directly and then to illustrate it with activities. In the motion unit, however, students were presented with motions with which they were already familiar: a ball dropped from a height of about seven feet; a ball rolling across a table and falling off to the floor; a shoved book that slid for a bit then slowed to a stop. Ms. Resnick demonstrated a particular motion for the students to observe and then asked them to draw the motion as well as they could with pencil and paper (see Fig. 10.1). The task to depict motion in a static medium required the invention of representational tools and was intended to focus students on ways of making their posited patterns explicit.

Initially, Ms. Resnick was afraid that the students would not see anything interesting in this task. After it was decided, for example, that the book slowed down, what else would there be to discuss? She was, therefore, surprised at the variety of ideas students had for representing the motion and at the symbol systems they developed. Presentation of students' representations led to further inquiry: Why is one symbol system better than another? Are all of the representations comparable to each other, or is there an element of incommensurability? What kinds of trade-offs are there from one representational technique to another? Is there a way to decide which is the best representation? Students ultimately had to grapple with the idea that what was considered "best" was relative to what they deemed important.

Ms. Resnick later traced her change in practice to the use of the reform curricula, which focused on student thinking, and the formal and informal collaboration with peers and researchers. She worked closely with the research team on the motion unit and also initiated a long-term collaborative relationship with Tracy Wilson, the other teacher who used the motion unit. This relationship developed to the point that they planned their math and science lessons together.

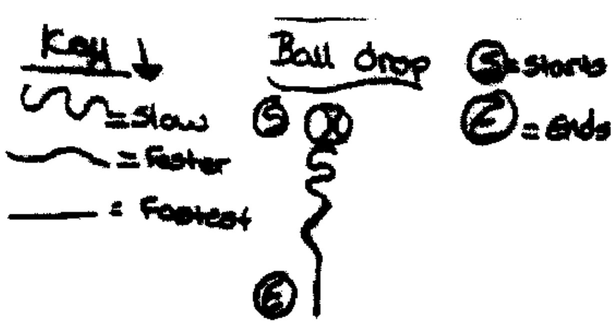

FIG. 10.1 Student representation of a dropped ball.

The Case of Jack Stoughton. Mr. Stoughton's most intense involvement with the collaborative was in the second year, when he and a researcher designed a 4-week Boxer-based geometry unit. This unit gave students the task of writing Boxer programs that would draw complex geometric shapes by iterating the rotation of simpler shapes. Mr. Stoughton stated that through the Boxer activities and subsequent class discussions, students learned relationships among important geometric concepts, such as interior angles, exterior angles, and regular polygons. In contrast to other geometry curricula, this unit engaged students, allowed them the freedom to develop their own tasks, and encouraged them to learn from each other as they worked in groups. As he shifted his focus to student thinking, the classroom environment also shifted, becoming more informal. Students had greater freedom to express their thoughts and approaches to problems, which in turn facilitated his assessment of their understanding.

Mr. Stoughton also thought it important that students not see math as something foreign to everyday life, but as something that had applications that could "empower" them in their day-to-day lives. He felt that listening to students was important because it made them feel valued as individuals. At the end of 2 years, he still felt he needed to strengthen his ability to elicit and understand his students' reasoning and to keep focused on the need to view each student as an individual with unique strengths and weaknesses. In the third year, the value he held for listening to students shifted to "helping students develop an awareness of how they and their peers are solving problems." He believed that excellent teaching resulted in students being willing to "take risks," and he made presentation of ideas to the group a class norm so that collaboration might lead to deeper thinking.

Over time, Mr. Stoughton recognized that the logistics of focusing on student thinking was difficult, particularly with a large number of students in the classroom. Although laudable, meeting each individual's instructional needs in a classroom of 25 students was, in practical terms, impossible. As a result, he shifted his goal from attempting to meet individual needs to identifying common problems, but acknowledged that defining such commonalities was a challenge. Similar to most teachers in the collaborative, he found the sharing of instructional techniques with other teachers essential and identified the lack of time for peer collaboration as the main challenge to realizing excellence in teaching.

Summary. As the cases of Ms. Resnick and Mr. Stoughton show, some teachers gave considerable effort to designing student-centered learning environments. Yet, as teachers reexamined what it meant to learn and understand mathematical ideas, they struggled to create learning environments consistent with their new assumptions. Practical considerations often resulted in trade-offs between meeting the needs of individual students and the practical constraints of teaching a classroom of students. To some degree, teachers were able to resolve the difficulties they encountered, although all agreed much work remained ahead. For each teacher, the collaborative positioned experimentation with practice as a worthy endeavor and promoted individual persistence in moving beyond initial uncertainty and discomfort. The collaborative also provided some degree of accountability, in that the move toward student-centered pedagogy was a collective endeavor.

Emerging Issue: Developing Principled Classroom Assessment Practices

Through the first two years of the collaborative, teachers' excitement with the new curricula (i.e., MiC and Boxer) was tempered by a growing sense that their methods for assessing student learning were inappropriate to the materials. In response to issues raised by teachers during monthly meetings in the second year and in the end-of-year teacher interviews, the research team decided to organize the meetings in the third year around two themes: (a) the identification of "core concepts" for assessing the development of algebraic understanding across all grades and (b) the discussion of alternative methods for designing, scoring, and grading formal assessments. To illustrate contrasting approaches resolving issues in classroom assessment, we focus on two teachers' efforts to develop a more meaningful approach to assessment.

The Case of Judy Koster. Ms. Koster had taught K–8 students for over 20 years. During the collaborative, she taught seventh-grade mathematics, language arts, and social studies. After hearing sixth-grade teachers talk about their experiences using MiC, she decided to participate in the collaborative. Dissatisfied with the dryness and repetition of teaching with a textbook, Ms. Koster wanted to experiment with a curriculum that challenged students *to reason* about math rather than simply to memorize computation methods. As she grew more familiar with MiC, she also realized that the way in which MiC spiraled algebraic concepts from unit to unit provided a developmental basis for instruction sequencing that she had tried to create in prior years, but had been unable to construct on her own.

To Ms. Koster, the greatest challenge was the design, selection, and scoring of performance assessment tasks, but her persistence in addressing assessment issues in the monthly meetings also reflected the concerns of her colleagues. Even though she had developed her own method of rubric-based scoring of student work, she found the scoring of more challenging assessment tasks problematic and experimented with various ways to

implement assessment that was fair, informative, and practical for both students and teachers. In contrast to the ubiquitous protocol for testing, in which students are not allowed to ask questions about the test problems, Ms. Koster decided to engage students in a brief reading and discussion of the problem context in order to mediate limitations in students' reading ability or lack of experience with a particular context. As evidenced in classroom observations and interviews during the third year, Ms. Koster was also less concerned with the *number* of problems covered or assessed than with the ways students articulated their understanding of mathematics. As Ms. Koster explained:

> It goes back to that whole explanation of, do you need 20 problems from a kid to understand that they understand it? Or can you put your energy and your focus in your discussion into one word problem? I think it's a tremendous shift in thinking and trade-off, but my, my main suggestion would be, put emphasis, focus, and energy on one or two problems. The exhausting part of this is you're on deck the whole time. But the trade-off is, if you're tuned in, I think you will have a more immediate sense of how the kids are doing. (Interview, May 6, 1999)

Ms. Koster noted that her participation in the collaborative made her excited about being "a student again"—solving problems and discussing with researchers and colleagues ways to use activities in the classroom. The discussions in the monthly meetings affirmed her philosophy of how mathematics should be taught and gave her the opportunity to relate her views to contemporary research in mathematics education. The collaborative also reinforced the value of setting aside time to collaborate with other teachers.

Ms. Koster and Deborah Harley, her "house" partner, collaborated in planning lessons, designing assessments, and scoring student work. During the 1999–2000 school year, they used the same formal assessments, administered tests the same way, and used a similar rubric to score student responses to assignments and tests. When using the end-of-unit assessments, both Ms. Koster and Ms. Harley decided to use the tasks as additional learning opportunities instead of as time-restricted tests of student knowledge. On the first day of a unit test, the entire class discussed ways to solve each problem, much as had been done throughout the unit. On the second day, students received a blank copy of the same assessment and were asked to write individual responses to each problem without the benefit of classroom discussion. In this way, Ms. Koster and Ms. Harley were able to informally assess student thinking during the pre-discussion and identify student misconceptions still present in students' work on "test day." We note that, although their approach gave the appearance of coaching students during formal assessment, it did not compromise student performance on the third-year algebra assessment we administered to all sixth- and seventh-grade mathematics classes (Webb et al., 2001).

The Case of Rebecca Mauston. Ms. Mauston had taught middle-grades mathematics, social studies, and language arts for 6 years. During the collaborative, she taught sixth-grade mathematics and social studies. When Ms. Mauston initially attended monthly meetings, she found that she enjoyed the opportunity to discuss the issue of practice with her colleagues, and she became more intrigued with MiC. She realized that the materials and activities of the collaborative paralleled much of what she was interested in accomplishing in her own classroom: creating a classroom community in which students respected each other and felt safe enough to share their thinking. In the final year of the collaborative, Ms. Mauston began to experiment with new methods of assessing, documenting, and grading growth in student knowledge over time. During several meetings with one of the authors of this chapter, Ms. Mauston outlined a more principled methodology for assessing student thinking.

Central to Ms. Mauston's change in practice was her development of a multilevel framework for assessing student reasoning in each unit of instruction. During her preinstructional planning, she established assessment goals for the unit, identified benchmark tasks (from tasks included in the unit and other available resources), and decided the appropriate social context for using these tasks. Her instructional goals were a combination of district content standards, learning objectives outlined in each unit, and personal interests. Ms. Mauston selected benchmark tasks on the basis of the accessibility of the task to a range of informal and formal student reasoning, the relevance of problem context, and her judgment of which tasks best represented her instructional goals. She used assessment tasks in a range of social contexts to create opportunities for dynamic assessment (Brown & Ferrara, 1999; Campione & Brown, 1987), in which student performance could be assessed with various degrees of instructional support, such as during whole-class discussion, group work, semistructured interviews, or during individual seatwork. As part of her new method of assessment planning, Ms. Mauston instituted student participation in "math chats" in order to take advantage of assessment opportunities available during students' mandatory tutorial period.

In assigning grades, instead of computing a percent-based grade, she rearticulated what students should demonstrate to earn a particular grade. Ms. Mauston organized learning objectives according to three levels of student understanding: basic skills, application, and analysis and extension. She then designed a rubric that assigned grades according to student performance on tasks at these levels. To earn a "C," for example, students needed to demonstrate most basic skills, good homework habits, and moderate participation. To earn a "B," students needed to demonstrate all basic skills, earn partial credit on most application problems, and maintain good homework habits and class participation. To earn an "A," students needed to extend their mathematical knowledge and demonstrate higher levels of reasoning on a higher level task.

In this process, Ms. Mauston made the meaning of grades and the ways they were earned explicit to herself and to her students. Classroom observations and interview data suggest that after Ms. Mauston introduced these assessment practices, her instruction and classroom discussions became more focused, and student motivation for learning mathematics increased. In a series of follow-up student interviews in the final year of the collaborative, Ms. Mauston's students described quizzes, tests, and "math chats" as opportunities to demonstrate their understanding rather than as a way to "get a good grade." By communicating curricular goals to students and designing a system that could be understood by students and flexible to their needs, Ms. Mauston encouraged students to assume greater responsibility for their own learning. As one sixth-grade student noted, "[In the new system] you actually have to show that you know the skill. It is not like in a multiple choice test, and you get it right, luckily, or something like that. You have to prove it to her that you know it and demonstrate it."

Ms. Mauston's assessment program caught the attention of the Director of Instruction, who asked her to participate in district-level meetings to write guidelines for application of district curriculum standards. In summer 2000, sixth-grade mathematics teachers in the Prairie Creek School District reviewed her methods for assessment planning and grading and agreed to develop similar assessment frameworks for use in their classrooms. District administrators encouraged these meetings and supported the planning needs of the sixth-grade mathematics teachers through paid time for summer and after-school monthly meetings. During the 2000–2001 school year, similar efforts at improving classroom assessment were promoted by the district through several staff development meetings to disseminate assessment guidelines for rubric-based scoring and grading and to allow other teachers to share other innovative assessment practices. In a relatively short period of time, the district attempted to institutionalize one teacher's innovation as a cross-disciplinary model for assessment.

Summary. We note that not all teachers in the collaborative adopted such assessment practices. Some colleagues questioned the validity of these methods and were uncomfortable with ways students could abuse assessment methods that included opportunities for students to collaborate with peers, share solution methods, and revise their work. On the other hand, teachers that continued to use conventional assessment practices were not able to resolve the disparity between their personal assessment of what students knew and the results they obtained from their more conventional routines for assessing, scoring, and grading student work. As Mr. Vandenberg, a seven-grade teacher, noted:

> I knew in my heart and my gut who was getting it and who wasn't. But in terms of, "Okay, how do I put that into a grade?"—that was a whole other challenge. We talked about it, and you have to give yourself permission probably [to] not grade as much. I mean things that you do grade, you have to get a rubric, or you have to get something down on paper so the kids really understand how they get the grade. (Interview, June 20, 2000)

IMPACT ON STUDENT ACHIEVEMENT

Student achievement data were gathered on all students taught by participating teachers during the second and third years of the collaborative. In the second year, 1998–1999, three different assessments for each of the three grade levels (specifically developed for this study) were administered at the beginning, middle, and end of the school year. In the first of three problem contexts, students were asked to continue a given pattern and provide an example of a similar one. The second problem context involved a skateboard race and addressed aspects of motion such as relative position, velocity, and acceleration. The third context was designed to elicit students' informal solutions to covariation problems that involved different combinations of menu orders, rental fees, and campground areas. Each context included questions to elicit student explanations of their solution strategies.

Many of the tasks used in the physics of motion and algebraic reasoning contexts were quite challenging and required students to interpret and reason with representations rarely addressed in the middle grades. Importantly, we found that students who engaged in the problem contexts were able to do so with a relatively high degree of success. The assessment results also suggested the following patterns of student achievement for the second year of the collaborative:

- Most incoming sixth-grade students were able to complete a given pattern.
- Seventh and eighth-grade students showed progressive growth in informal and preformal algebraic reasoning and were able to generalize the pattern described in a realistic problem context.
- Comparison of teachers' curricular decisions and student achievement data showed that student opportunity to learn significantly affected student performance in preformal algebraic reasoning and generalization.
- The cooperation and mutual support of the sixth-grade teachers led to a more coherent mathematics program and increased opportunities for students to engage in and improve their solving of complex, nonroutine problems. By the spring, sixth-grade students were beginning to perform at a level equal to the performance of eighth-grade students.

In the third year, 1999–2000, the teachers emphasized the study of student reasoning in algebra and sought to provide students with appropriate pre-algebra experiences in sixth and seventh grade. In alignment with this goal, the third-year assessment was designed to document student achievement in various aspects of algebra, including use of patterns, covariation, and understanding of exponential phenomena. The first section of the assessment included five multiple-choice questions selected from eighth-grade public-release items used in the Third International Mathematics and Science Study (TIMSS). The second section of the assessment included four contexts with several constructed-response prompts associated with each problem context. Three of these problem contexts were used on the previously described second-year assessments. Another problem context was taken from the Grade 7 Problem Solving Assessment (Dekker et al., 1997–1998) developed for the "Longitudinal/Cross-Sectional Study of the Impact of Mathematics in Context on Student Mathematical Performance." Each context included prompts to elicit student explanations of their solution strategies. (For a more detailed report of data from this assessment, see Webb et al., 2001.)

Overall, the student achievement results for the third year demonstrated substantial growth in students' algebraic reasoning. We attribute these results to teachers' emphasis of algebra content in sixth and seventh grades. More specifically, the assessment results showed the following patterns of student achievement:

- In the longitudinal analysis, we found that the differences in student performance from year to year were statistically significant. In two cases, the difference in student performance was greater than one standard deviation. However, there remains significant room for improvement on many of these items.
- On the multiple-choice tasks, we found an overall difference in performance between sixth- and seventh-grade students. Similar to achievement patterns found in the second-year data, we conjectured that student opportunity to learn and instructional emphasis significantly affected student performance on these items.

- In contrast to this difference, we found student performance on particular items in Grades 6 and 7 to be equivalent. In order to explain why performance on these items was equivalent when all other items saw performance differences across grades, we also cite student opportunity to learn.
- In Grade 6, we found a significant difference in performance between those students who used preformal algebraic tools (e.g., combination chart) and those who did not. Use of such tools, however, was not found in responses of the seventh-grade students who successfully responded to this item. Very few seventh graders organized their problem solving with such tools and instead demonstrated use of formal, symbolic approaches.
- We found variation in student performance across teachers that taught the same MiC units. Although some of these differences could be attributed to a general teacher effect, based on observations and informal conversations we had with teachers, we conjectured that teachers' differential instructional emphasis within the same units had an impact on student performance. For instance, some seventh-grade teachers explored exponential growth and decay phenomena in detail, whereas others skimmed the topic.

Students' opportunities to learn and teachers' instructional emphasis had a significant impact on student achievement. However, simply stating that "students learned what teachers decided to teach" misrepresents the relationship between teacher change, curriculum coherence, and student achievement that we observed in this study. Providing students opportunities to learn mathematics with understanding required an environment in which teachers had an opportunity to develop pedagogical knowledge, mathematical knowledge, and classroom practice. By the end of the third year, five of the six Grade 7 teachers were in their first year of using a new mathematics curriculum. Every teacher in the collaborative noted that his or her prior classroom practices were challenged in one form or another. We view the positive trend in student achievement data to be a set of assessment snapshots that confirm the improvement in student learning of mathematics that occurred during the course of the collaborative. The longitudinal growth in student performance in the third year is particularly significant given that most of the teachers for this student cohort were just beginning to grapple with the inherent challenges of student-centered instruction and more ambitious forms of classroom assessment. Over time, as teachers continue to improve their classroom practices and become more familiar with the way mathematics concepts are developed, we expect their students will demonstrate even greater performance on similar assessments.

CONCLUSION

The collaborative created a "safe" environment for teachers to share and discuss their problems of practice. As they made their practice public, they found that some issues, which they might have perceived as idiosyncratic, were often shared by many of their colleagues. Through the collaborative, "personal" teaching issues were often reflected back as *problems of practice*, and teachers, supported by the district and working jointly with researchers were able to experiment with changing their practices through reform approaches to instruction and assessment. Certainly, the growth of teachers as they participated in the collaborative varied considerably. Some, like Ms. Resnick and Ms. Mauston, experienced considerable change in their work as professional teachers. Some, like Ms. Koster, found that the instructional expectations reinforced their beliefs about excellent teaching. All found the teaching of reform units to be challenging, but all also saw the beneficial impact of such instruction on student learning. In the sections that follow, we summarize the effect the collaborative had on their instruction and on their implementation and use of assessment.

Collaborating to "Uplevel" Instruction

As teachers piloted the new units (MiC and Boxer-based), they deepened their understanding of student conceptions of math and science. As teachers continued to elicit student explanations of their strategies, sometimes taking the stance of someone completely ignorant of a mathematical topic to encourage student explanations, they found that their own understanding of the concepts they were teaching was enhanced by students' responses. As a result, teachers began to make more-informed decisions about which sections of a unit they should emphasize, skim over, or avoid. As they became more familiar with the content of the units, their instructional stance became more proactive, and they began to interject classroom discourse with more formal mathematical language in order to connect student conceptions to a shared language.

Teachers' experimental use of Boxer and MiC, for example, permitted a "safe zone" for teachers to explore content not previously used with students at these grade levels. As teachers piloted Boxer and MiC algebra units, they were intrigued by the enriched nature of students' mathematical communication and reasoning. Because teachers often experimented with the same instructional units, they were able to discuss not only the interaction of curriculum, instruction, and assessment in a shared context but also the merits of a particular activity and the range of student responses their instructional practices had elicited.

Teachers were quick to identify this peer collaboration as key to supporting their efforts to implement reform practices and in helping them resolve the practical problems that arose as they worked to reform their practice and to develop the skills they felt they needed. Sharing strategies and conceptions of the desired skills tended to "spread out the risk" the teachers were taking in implementing change. The reforms supported by the collaborative necessitated a modification of the traditional model of the classroom, in which "correct" information is transmitted from an expert to a roomful of novices, toward a more open-ended, student-centered model in which knowledge is dynamic, collaboration is involved in cognition, and students and teachers sometimes share leadership roles. The difficulty faced by most participants in implementing this model seemed rooted in two challenging issues:

- Planning out an instructional route that would enable them to simultaneously focus on student discussion, infer student thinking, and gather evidence on student understanding.
- Shifting from the comfortable role of class leader to the "risky" role of class facilitator.

Effecting these changes was difficult, but as teachers shifted to the focus on student thinking inherent in this reform model and experimented with new ways to scaffold student learning, many realized that this was what they had been trying to do for years. Teachers already knew from experience that students varied considerably in their reasoning about problems, but what students thought and the strategies they used were hidden from traditional approaches to instruction and assessment. Reinforced in their efforts by the district, the researchers, and their peers, teachers experimented with practices in actual classroom environments. Discussion of what they did and observed at the monthly meetings legitimized their intuitions and fostered their sense of professionalism.

Linking Assessment and Achievement

Three concurrent developments emerged in teachers' classroom assessment practices: the investigation of instructionally embedded assessment; the selection, design and use of assessment tasks; and the scoring and grading of student work. Both formally and informally, teachers in the collaborative began to discuss their assessment, scoring, and grading approaches and shared their experiences with assessment tasks in MiC and the challenges they faced in using tasks to assess student reasoning and problem solving. At their monthly meetings, teachers deliberated the pros and cons of different approaches and talked about assessment methods used in language arts, science, and social studies. The framework for planning classroom assessment developed by Rebecca Mauston was shared with teachers outside of the mathematics department and was adapted for use by language arts and social studies teachers. During the final month of the collaborative, Grade 6 mathematics teachers received district funding to develop MiC unit assessment plans, a project that included the identification of instructional goals in light of the district's standards and the selection of unit benchmark tasks. Several teachers found that their use of reform assessment practices and the identification of appropriate informal classroom assessment opportunities had influenced their instructional decision making, the selection of curricula, the selection and modification of assessment tasks (and how they were used), and even the instructional pathways they used to guide student discourse.

Certainly teachers found that their change in practice deepened and changed their instruction, but the change in their classroom assessment practices also had a notable impact on student disposition toward learning mathematics. Instead of using assessment primarily to judge whether students had successfully *completed an activity*, teachers had begun to use assessment to evaluate whether students *understood mathematics and science concepts*. When one teacher adjusted her expectations of the ways students should demonstrate their understanding of mathematical skills and concepts, her students improved their work according to the expectations they had *for themselves*. For example, when students were presented with a range of mathematical competencies they needed to demonstrate to earn an "B," students took greater ownership in the learning process and decided which concepts they wanted to learn and demonstrate on the targeted assessments.

As teachers created contexts to support student learning of mathematics and science with understanding, they discovered that middle-grades students could model fairly sophisticated phenomena. Sixth-grade students were able to represent cases of constantly accelerated motion in a variety of ways, and these representations supported discussions about concepts central to kinematics (e.g., speed, time, distance). Overall, students were also able to arrive at multiple conceptual organizations of concepts such as accelerated motion on a ramp (in order to measure changing speed) and could recognize representation issues and could refine their depictions accordingly. Importantly, students were able to demonstrate that they had learned to support what they found through argumentation.

Summary

Overall, the collaborative was effective in supporting and sustaining the effort toward improving instruction and in building school capacity to provide all students with a mathematics program that was comprehensive and meaningful. The context of university collaboration also validated teachers' efforts to administrators and parents, provided incidental technical support, and supported teachers' efforts through provision of necessary material resources such as MiC units, Boxer software, and computers (which would not have been purchased by the district because of limitations in budget allocations for mathematics resources).

After the researchers pulled back, the project was sustained by further exploration of practice by project teachers, dissemination of support structures and classroom practices across the Prairie Creek School District, and "travel" of project findings to researchers and teachers in new projects and new school districts. Field notes gathered by researchers who continued to observe monthly meetings after the conclusion of the collaborative confirmed sustained deliberation of curricular decisions and instructional approaches in a continual effort to improve student learning. The level of teacher discourse remained focused on problems of practice, further suggesting that the norms of professional discourse developed during the collaborative were maintained, at least in the short term.

Clearly, to sustain growth in teachers' classroom practice, teacher collaboration in content areas must continue as a means to support further experimentation and development of teacher knowledge central to teaching and learning mathematics with understanding. Such collaboration could focus on teachers' (a) engagement in new activities that promote teacher learning of mathematics content, (b) recognition and collection of evidence of students' understanding and learning, (c) elicitation and interpretation of that evidence, and (d) exploration of assessment methods appropriate for assessing student understanding. Although such collaboration might include university personnel, over time we expect teachers to take greater ownership in maintaining collaborative structures and continuing and sustaining the discussion of teaching and learning mathematics.

We note in closing that the key issues—district leadership support and the emergence of teacher community—had great effect on the collaborative, which continues informally despite major changes in the district and schools themselves. We also wish to point out the importance of the shift in assessment practice, which—when used not to rank students, but to inform instruction—led to better quality of instruction and substantial increase in student achievement, as measured both by informal classroom assessment and external formal testing.

AUTHOR NOTE

The research reported herein was supported in part by a grant from the U.S. Department of Education, Office of Educational Research and Im-

provement, to the National Center for the Improvement of Student Learning and Achievement in Mathematics and Science (Grant No. R305A60007-98), and by the Wisconsin Center for Education Research, School of Education, University of Wisconsin–Madison. Any opinions, findings, or conclusions are those of the authors and do not necessarily reflect the position, policy, or endorsement of the supporting agencies.

REFERENCES

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (1999). *How people learn: Brain, mind, experience and school*. Washington, DC: National Academy Press.

Brown, A. L., & Ferrara, R. A. (1999). Diagnosing zones of proximal development. In P. Lloyd & C. Fernyhough (Eds.), *Lev Vygotsky: Critical assessments: Vol. 3. The zone of proximal development* (pp. 225–256). New York: Routledge.

Campione, J. C., & Brown, A. L. (1987). Linking dynamic assessment with school achievement. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 82–115). New York: Guilford Press.

Carpenter, T. P., & Lehrer, R. (1999). Teaching and learning mathematics with understanding. In E. Fennema & T. A. Romberg (Eds.), *Classrooms that promote mathematical understanding* (pp. 19–32). Mahwah, NJ: Lawrence Erlbaum Associates.

Dekker, T., Querelle, N., van Reeuwijk, M., Wijers, M., Feijs, E., de Lange, J., Shafer, M. C., Davis, J., Wagner, L., & Webb, D. C. (1997–1998). *Problem-solving assessment system*. Madison, WI: Wisconsin Center for Education Research, University of Wisconsin–Madison.

Fennema, E., & Nelson, B. S. (1997). *Mathematics teachers in transition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Fennema, E., & Romberg, T. A. (Eds.). (1999). *Classrooms that promote mathematical understanding*. Mahwah, NJ: Lawrence Erlbaum Associates.

National Center for Research in Mathematical Sciences Education & Freudenthal Institute. (1996–1998). *Mathematics in context*. Chicago: Encyclopædia Britannica.

Webb, D. C., Ford, M. J., Burrill, J., Romberg, T. A., Reif, J., & Kwako, J. (2001). *Prairie Creek Middle School design collaborative: Year 3 student achievement data technical report*. Madison, WI: Wisconsin Center for Education Research, University of Wisconsin–Madison.