Effects of Peer Coaching on Teachers' Collaborative Interactions and Students' Mathematics Achievement

SARAH MURRAY
Centre College

XIN MA
JOAN MAZUR
University of Kentucky

ABSTRACT. The authors examined peer coaching in the context of the Mentored Implementation Program developed in the Appalachian Mathematics and Science Partnership. The experimental design contained 6 teachers receiving peer coaching with their 202 students and 5 teachers in the control group with their 105 students. Teachers considered peer coaching a positive experience, identifying scheduling and distance as roadblocks. Centering around (a) organization of learning, (b) management of classroom, and (c) mathematical content and pedagogy, collaborative interactions during the post-classroom-observation conferences were brief and showed (a) a lack of analysis, (b) a positive tone and support, (c) a proportional pattern of talk, and (d) a lack of depth in discussion. Peer coaching was not associated with any improvement in mathematics achievement of students.

Keywords: collaborative interactions, mathematics achievement, peer coaching

Because "students learn mathematics through the experiences that teachers provide" (National Council of Teachers of Mathematics [NCTM], 2000, p. 16), teachers' professional development is key for student success in mathematics. Hassel (1999) defined *professional development* as "the process of improving staff skills and competencies needed to produce outstanding educational results for students" (p. 9). Guskey (2000) asserted that "one constant finding in the research literature is that notable improvements in education almost never take place in the absence of professional development" (p. 4). Despite this universal recognition of the importance of professional development, educators and researchers do not agree on which professional development models help teachers the most.

Professional development models, though abundant in the research literature, can be put into two main categories: teacher mentoring and peer coaching (see Conley, Bas-Isaac, & Scull, 1995). According to these researchers, teacher mentoring involves a hierarchical relationship between senior teachers (sometimes educational specialists and university faculties) and junior teachers, but peer coaching involves a mutual consultation between teachers of equal status. The purpose of the present study was to examine the effectiveness of a job-embedded professional development model of peer coaching. Research questions focused on teachers' peer-coaching experiences (i.e., collaborative interactions) and whether peer coaching could improve students' mathematics achievement. Our research effort joins that of other researchers to accumulate empirical evidence (currently insufficient in the research literature) for a critical comparison of effectiveness between the two professional development models.

Research Setting

Supported by the National Science Foundation, the Appalachian Mathematics and Science Partnership (AMSP) was a collaborative effort of the University of Kentucky, University of Tennessee, University of Virginia, Kentucky Science and Technology Corporation, 38 Kentucky school districts, 9 Tennessee school districts, 5 Virginia school districts, and regional universities, colleges, and agencies to develop an overlapping network of partnering and mentoring relationships across kindergarten to Grade 16 (K–16) levels. The goal was to enable students, teachers, school administrators, and higher education faculties to share their expertise and support one another's continuing learning.

One key initiative of the AMSP was the professional development of K–12 personnel, with two integral phases: the mathematics and science summer institutes and the Mentored Implementation Program (MIP). The summer institutes consisted of 1- to 2-week sessions on specific mathematics and science content held regionally at partner higher education institutions and were facilitated by higher education faculties and teacher leaders. Following the summer institutes, teachers participated in the MIP professional development activities. MIP differs from the majority of

Address correspondence to Xin Ma, University of Kentucky, 335 Dickey Hall, Lexington, KY 40506-0017, USA. (E-mail: xin.ma@uky.edu)

professional development models because it provides necessary support to teachers as they implement in their own classrooms the standards-based and inquiry-oriented mathematics and science activities from the summer institutes. Traditional professional development activities often lack this crucial follow-up component.

In the MIP model, peer partners are experienced teachers who have attended one or more of the summer institutes and serve as coaches as they observe and critique one another during the implementation phase. Teachers have the opportunity to choose with whom they are paired. As much as possible, peer partners attend the same institute and teach in the same school or school district. Peer partners are expected to observe and coach one another a minimum of two times throughout the school year (after a summer institute). Coaching, assisting, and supporting are provided not only through a network of peers but also through lead mentors (district specialists, college teachers, and university professors). Lead mentors are assigned a group of peer partners in their region and facilitate participants' having a positive and rewarding experience by (a) helping to schedule and coordinate mentoring activities, (b) communicating with each peer partner regularly, and (c) observing each peer partner a minimum of one time to provide further mentoring and coaching.

Theoretical Foundation

The *zone of proximal development* (ZPD) is a challenging level of development that an individual reaches through social interaction (Vygotsky, 1978). According to Goos, Galbraith, and Renshaw (2002),

> Applied to educational settings, this view of the ZPD suggests there is learning potential in peer groups where [partners] have incomplete but relatively equal expertise—each partner possessing some knowledge and skill but requiring the others' contribution in order to make progress. (p. 195)

Also, "Clarification, elaboration, justification, and critique of one's own or one's partner's reasoning" are identified as characteristics of the "collaborative ZPD" (Goos et al., p. 199). These are critical characteristics of what would be considered a successful peer-partner conference (after a classroom observation) in the MIP, in which teachers ask each other to clarify, elaborate, and reflect on their teaching practices.

Specifically, Goos et al. (2002) identified self-disclosure, feedback request, and mutual monitoring as key types of interactions that can be observed during the peer-mentoring process. *Self-disclosure* includes self-oriented statements and responses that clarify, elaborate, evaluate, or justify an individual teacher's own teaching. *Feedback request* includes self-oriented questioning that encourages a peer to critique or reflect on his or her partner's teaching. *Mutual monitoring* includes statements, questions, and responses that represent an effort of peer partners trying to understand each other's thinking. When teachers challenge or question each other's thinking, the collaborative problem-solving process becomes more successful because "challenge is a defining feature of the zone of proximal development" (Goos et al., p. 218).

Romagnano (1994) emphasized the importance of collaboration to facilitating change in behavior. The key is to place teachers in what Romagnano called "collaborative insider" (p. 174) positions. The MIP adopted this concept to encourage peer-partnered teachers to become agents to implement and support change. Lyons and Pinnell (2001) asserted that teachers become more analytic about their own practice through reflective collaboration with peers and that the support teachers receive from peers is critical to applying new knowledge, strategies, and techniques in the classroom.

Peer Coaching

Existing coaching and mentoring approaches consist of apprentice models in which one teacher is the mentor or coach and the other, less experienced teacher is the protégé. West and Staub (2003) stated that

> content-focused coaching is related to apprenticeship, in which an apprentice is observed while carrying out a task and the master craftsman offers hints, provides support, gives feedback, models, gives reminders, and poses new tasks aimed at bettering performance. (p. 2)

This coaching approach serves as the program foundation for the MIP with a twist. In the MIP, one teacher is not viewed as more of an expert than the other. Instead, they work in partnership, equally supporting and learning from each other to implement teaching activities and strategies from the summer institute for the improvement of classroom practices. Therefore, the MIP is clearly a peer-coaching professional development model (see previous definition).

In the present study, we considered it appropriate and meaningful for the MIP to adopt peer coaching because the goal was to give teachers the opportunity to learn from one another in what Wenger (1998) referred to as a "community of practice," defined as "shared histories of learning" with a focus on learning as social participation (p. 86). The MIP encourages teachers to reflect on and rethink their teaching practices in line with Wenger's emphasis on the "urgent need for reflection and rethinking" (p. 9). He described three dimensions that characterize the community of practice. First, mutual engagement involves teachers' working together and negotiating meanings with one another. In the MIP, this process takes place during the postobservation conference, through e-mail discussion, and even in the hallway for those teachers in the same school as peer partners discover, develop, define, and establish evolving forms of mutual engagement.

Second, joint enterprise involves teachers' negotiating the meanings and vocabulary tied to a given task or activity. Teachers in the MIP have the joint task of

implementing materials from the summer institutes and holding one another accountable at the postobservation conference by discussing, reflecting on, and critiquing a lesson taught. Third, shared repertoire involves a wide variety of things from routines to tools. In the MIP, teachers use their experiences from the summer institutes to renegotiate meanings that define what the classroom should look like when they implement materials from the institutes. According to Wenger (1998), constant renegotiation exists within the community of practice. The MIP offers teachers the opportunity to participate in such a community of practice in which the meaning of effective mathematics teaching is constantly renegotiated.

Again, the MIP fits into the definition of peer coaching as a process in which teachers work collaboratively to solve problems and answer questions pertaining to the implementation of innovations (Joyce & Showers, 2002). According to Ackland (1991), peer coaching is appropriate for programs that are (a) nonevaluative, (b) based on classroom observation with feedback, and (c) intended to improve instructional strategies and techniques. The MIP creates such a process for teachers to collaborate and support one another as they implement materials from the summer institutes. Teachers were asked from the onset of the summer institutes to partner with a colleague, and that request was reinforced by follow-up activities in the MIP. Teachers were also provided with training on coaching and mentoring their peer partner. This training, with the supervision of the lead mentors, was intended to prevent peer coaching from becoming a "shared ignorance" (see Guskey, 1999, p. 11).

Status of Literature

After a review of theory and practice of peer coaching, Joyce and Showers (2002) stated that peer coaching can be applied with good promise to professional development in which the objective is substantial improvement in (a) knowledge, (b) skill, and (c) transfer of training into the classroom. These researchers emphasized that teacher mentoring (coaching by trainers) may produce similar improvement but is not practical or possible in most educational settings. Joyce and Showers estimated how successful the peer coaching can be in transferring training into the classroom, and their conclusion was 95% of participants are likely to transfer their training in a satisfactory way to their classrooms, compared with at most 5% of participants without peer coaching.

In perhaps the largest survey research of coaching so far, Greene (2004) found that (a) teachers are positive about the support provided through coaching, especially when they plan together with coaches; (b) the amount of joint planning depends on personalities of coaches; (c) teachers view coaches who make frequent informal contacts as more supportive; (d) teachers notice inconsistencies between theory and practice among coaches; (e) although coaches' demonstration (modeling) in the classroom does not go beyond teachers' instructional expertise, it does help teachers to rethink and modify their instructional practice; and (f) factors hindering the effectiveness of coaching are lack of time, teacher resistance, and scheduling conflict.

Bowman and McCormick (2000) compared peer coaching with traditional supervision in an experimental design. Focusing on the development of "clarity skills," "pedagogical reasoning," and "levels of satisfaction" (pp. 256–257) among preservice teachers, they concluded that

> collaboration fosters expert instruction. Evidence presented indicates that peer coaching is a feasible vehicle for instituting collaborative efforts; therefore, peer coaching warrants consideration as a potentially serviceable solution for strengthening field-based training of prospective teachers. (p. 261)

Showers, Joyce, and Bennett (1987) asserted that "combinations of four components (theory, demonstration, practice, and feedback) appear necessary to develop the levels of cognitive and interactive skills that permit practice in the classroom" (p. 86). Teachers who are provided with continued technical assistance by peer experts are more likely to achieve greater classroom implementation of their training (Joyce & Showers, 2002).

According to Kohler, Crilley, Shearer, and Good (1997), researchers have typically focused on three outcomes to evaluate the effects of peer coaching: (a) procedural change in teachers' pedagogical strategies, methods, or techniques; (b) teachers' interactions with peer coaches and their ongoing satisfaction with the process; and (c) student outcomes such as academic skills and competencies. Kohler et al.'s own research showed that although more procedural change occurs during peer coaching than when teachers work independently, teachers feel uncertain whether the instructional innovation that they are implementing under the guidance of peer coaches is effective. Positively, peer-coaching activities do promote a high level of student engagement in learning (Kohler et al.).

Unfortunately, most claims for the effectiveness of peer coaching are anecdotal in nature. Although success stories continue to emerge, there were few empirical data to support any major claim, particularly in regard to the relation between peer coaching and student learning. For example, Sparks and Bruder (1987) reported that although 70% of teachers claimed that peer coaching improves student achievement, they offered little supportive data. Russo (2004) insisted that little data evinced that peer coaching leads to improved student achievement. It is understandable that research on the effectiveness of peer coaching is so thin because peer coaching, though showing promise, is a relatively new form of professional development. The severe lack of empirical data underscores the significance of any empirical research on peer coaching. We performed the present study to fill this gap in the research literature: Thus, we went beyond classroom observations and teacher interviews to collect empirical data on collaborative interactions of peer coaches during postobservation conferences

and investigated the improvement in mathematics content knowledge of students of peer partners by using an experimental design with pretests and posttests.

Method

To determine the effects of peer coaching on teachers' collaborative interactions and students' academic achievement, we used an experimental design with experimental and control groups. Qualitative methods were used to examine collaborative interactions and perceptions of the MIP among mathematics teachers. Quantitative methods were used to analyze pre- and posttest scores of students in the experimental and control groups and explore the relation between peer coaching and student achievement.

Participants

Participants were 14 teachers in six schools (one K–8 school, three middle schools, and two high schools) from four school districts. Of the 14 teachers, 9 were in AMSP professional development (the summer institute and the MIP), forming the experimental group;[1] and 5 were not, forming the control group. Teaching experience for the experimental group ranged from 2 to 16 years, and that for the control group ranged from 1 to 32 years. Teachers in the experimental group had 202 students (44 in Grade 7, 121 in Grade 8, and 37 in Grade 9); teachers in the control group had 105 students (30 in Grade 7, 52 in Grade 8, and 23 in Grade 9).

All AMSP school districts had an equal opportunity to send mathematics teachers to the summer institute, Transition to Algebra, held in Kentucky in 2005. Although teachers received a stipend of $1,000 for participating in the institute, participation was voluntary. To account for this fact, we recruited teachers for the control group from similar school districts with comparable student populations. We matched school districts on a variety of background characteristics including (a) daily attendance, (b) revenue per student, (c) expenditure per student, (d) number of classified staff, (e) number of certified personnel, (f) percentage of students eligible for free or reduced-price lunch, and (g) overall student academic performance in reading and mathematics for the previous 2 school years as reported in the state testing program. We realized that lead mentors would work with peer partners in the MIP. To control for the lead mentor effects, we made certain that recruited peer partners came from the same region so as to have the same lead mentor.

Measures

For the qualitative component that examined mathematics teachers' perceptions of the MIP, a short perception survey instrument was designed and administered to teachers in the experimental group to assess the qualitative dimensions of mathematics teachers' perceptions of the MIP. Teachers remarked on their experience with the MIP, in particular their perceptions of the benefits of and barriers to the MIP.

For the quantitative component that examined the effects of peer coaching on student achievement, we faced the challenge that participating school districts adopted different textbook series and we lacked the resources to develop a curriculum-based mathematics achievement test. A common methodological solution in research situations like ours is to develop a skill-based test that is independent of any curriculum. Adopting this strategy, we sought existing skill-based mathematics achievement tests. The *Programme for International Student Assessment* (PISA; Organization for Economic Cooperation and Development, 2000, 2003) uses a skill-based mathematics achievement test independent of mathematics curriculum in any participating country. Instead of assessing students' mastery of a particular mathematics curriculum, PISA focuses on students' ability to use their mathematical knowledge and skills to solve real-life problems. This was the main reason why we adopted items from PISA.

We selected 19 items from 11 PISA 2000 and 38 PISA 2003 sample items. The criterion for inclusion was how closely an item represented an area in the Kentucky Core Content for Mathematics (Kentucky Department of Education, 2005). The test was piloted among 19 students in one middle school (12 seventh-grade and 7 eighth-grade students). Some psychometric analyses resulted in the deletion of 3 items. A reliability of .65 was obtained for our mathematics achievement test. This reliability was not as high as we expected, and we believed that the lower reliability may be related to the fact that PISA targets eighth-grade students but our study included seventh- to ninth-grade students. Nevertheless, our test was still considered a fairly good representation of educational expectations for students because adopted items were well aligned with the state mathematics curriculum. Overall, we believed that this relatively low reliability was not a serious threat to our study.

Analysis

For the qualitative component that examined mathematics teachers' collaborative interactions in the MIP, we audiotaped postobservation conferences of the 6 teachers from the experimental group. Audiotapes were transcribed, and the data were coded, examined, and evaluated for evidence of teacher collaboration and reflection. Qualitative data analysis focused on the identification of the number of times each peer partner spoke without interruptions; the various types of comments that peer partners made, including questions, statements, and compliments; and the content of those questions, statements, and compliments.

A multiple-regression approach to analysis of covariance (ANCOVA) was used to determine whether the difference in the posttest scores between the experimental and control groups was statistically significant.[2] The dependent

variable was the posttest scores, the covariate was the pretest scores, and the independent variable was the treatment, coded as a dummy variable to identify whether a student was in the experimental or control group. We first tested the interaction effect between the treatment and the pretest scores. If this interaction is statistically significant, then the treatment effects depend on the pretest scores. If this interaction is not statistically significant, then a second model can be run without the interaction term to determine whether the treatment has any effect on the posttest scores (with control of the pretest scores).

Results

Mathematics Teachers' Perceptions of Peer Coaching

All teachers in the MIP responded to an open-ended survey on their perceptions of the MIP.[3] Table 1 presents a breakdown in terms of the benefits and barriers of the MIP that teachers perceived, organized from responses given most often to those given least often. The primary beneficial aspect of MIP participation indicated by teachers was the opportunity to share ideas, techniques, and strategies with their peer partners. Peer partners also identified communication with and supporting one another as key benefits of the MIP. Each of these benefits provided evidence of collaborative interactions between teachers. However, teachers identified scheduling as the major barrier or roadblock to achieving intended purposes of the MIP. Peer partners also reported distance as a key barrier. Each of these barriers could interfere with peer partners' commitment and opportunity for collaborative interactions.

TABLE 1. Perceived Benefits and Barriers of the Mentored Implementation Program Among Peer Partners

Category	Frequency (%)
Benefits	
Sharing ideas, techniques, and strategies	47
Getting feedback and the perspective of another teacher	25
Observing another teacher	23
Discussion or communication with another teacher	19
Supporting one another	4
Barriers	
Scheduling	39
Distance	14
Timing	8
Getting a substitute	3

Mathematics Teachers' Collaborative Interactions During Peer Coaching

In addition to peer partners' responses to the perception survey of the MIP, we examined the collaborative interactions of peer partners in their postobservation conferences. In the experimental group, 9 teachers were involved in four postobservation conferences (from four different groups) that we audiotaped and transcribed. Table 2 presents a summary of the analyses. The average conference time was 13 min, and the average number of topics discussed was 12. The topics focused on organization of learning, management of classroom, and mathematics content and pedagogy. The types of interactions during the postobservation conferences included questions, statements, and compliments. The percentage of questions during a postobservation conference ranged from 6% to 13%. The percentage of statements was the greatest, ranging from 83% to 91%. The amount of compliments given during a postobservation conference ranged from 0% to 9%.

Analysis of the transcripts from the four postobservation conferences in Table 2 further revealed four substantial themes and patterns of collaborative interactions among peer partners. First, teachers were more descriptive of (as opposed to reflective or analytical about) the actual classroom observation (lack of analysis). Peer partners provided minimal critique to the classroom teachers whom they observed, often describing what they saw rather than providing any analysis. Neither did the observers ask any question that would effectively motivate reflection or analysis. Similarly, the observees were more descriptive than reflective. The following examples illustrate this pattern.

Group 1 Observer: "[For] student involvement, I thought most of the students were involved and interested. On task and that is just one of [your] talents to get an Algebra 1 class always attentive and interested."
Group 1 Observee: "One student. But I can't complain."
Group 1 Observer: "The others were very well behaved I thought."

Group 2 Observer: "Is this something that you think you would want to use again?"
Group 2 Observee: "Yes. It would be really good in 8th grade because I know in 8th grade they have questions that deal with blocks and what their views would look like." [referring to questions on state assessments]

Group 3 Observer: "Was there anything else that you wanted to add?"
Group 3 Observee: "I guess not. Actually this group had done this earlier in the year. They were a little familiar with it but I felt that they were forgetting the why. They got use [sic] to using the rules but were confusing the rules so we went back to the manipulatives."

Group 4 Observer: "The one thing that I liked was the contract that you showed me."
Group 4 Observee: "What we do is the kids sign a contract and their parents sign saying that if the calculator gets damaged by the student that they will pay the replacement cost which is $140."

TABLE 2. Summary of Postobservation Conferences of Four Groups of Peer Partners

Group	Time (%)					
	Length (min)	Interchanges	Topics	Organization and management	Content and pedagogy	Other topics
1[a]	19	76	18	44	50	6
2[b]	13	39	13	38	62	0
3[c]	7	20	7	14	86	0
4[d]	14	125	10	50	10	40

[a]Group 1 had the observee teaching Grade 9 and the observer teaching Grades 10–12 at the same high school. [b]Group 2 had the observee teaching Grade 7 and the observers teaching Grade 8 at different middle schools in the same school district. For the other groups, teachers conferenced over two classroom observations, with each teacher being both observer and observee. [c]Group 3 had teachers teaching Grade 8 at the same middle school. [d]Group 4 had one teacher teaching Grade 8 at a middle school and the other teacher teaching Grades 10–12 at a high school in the same school district.

As can be seen from the conversation excerpts, although some observers did attempt to encourage reflections from the observees, their questions were simply not focused, leading to many possible descriptive responses.

Second, teachers were very positive and supportive of one another (positive tone and support). They supported one another's efforts through encouraging comments. Overall, 100% of the conversations were positive in tone without a single negative comment made. Most comments were based on a description of what was observed during the classroom observation (e.g., homework assignment, educational resources, classroom activities). The following examples illustrate this pattern:

Group 1 Observer: "I can definitely tell in your classes that you have such a good rapport with your students. You really do."

Group 2 Observer: "I thought it was neat how you opened with a real-world connection. Got everyone together and focused. You gave them time to play with their manipulatives before. I thought that was really good so they weren't distracted."

Group 3 Observer: "I thought that was an excellent way to do that with those blocks."

Group 4 Observer: "I like the story."

Third, instead of the domination of one teacher over the discussion, there was relatively equal sharing of thoughts and ideas (proportional pattern of talk). This pattern was most evident in two groups. In Group 1, the least number of times that an individual spoke without interruption was 24, whereas the most was 28. In Group 3, one partner spoke 11 times without interruption, and the other partner spoke 10 times. Last, although peer partners discussed a wide variety of topics, their discussion appeared to be superficial (lack of depth). We noticed that Group 1 discussed 18 topics in 19 min, Group 2 discussed 13 topics in 13 min, Group 3 discussed 7 topics in 7 min, and Group 4 discussed 10 topics in 14 min during the postobservation conference. The average was 1 topic per minute. Such little time on a certain topic could only allow a superficial discussion.

We also noticed that in two of the four postobservation conferences the roles of observee and observer were not clearly defined because each peer partner was trying to discuss what he or she could recall from a previous classroom observation. For all conferences, discussions between peer partners consisted of explanations of what occurred during the classroom observations rather than meaningful analyses of how classroom instructions could be improved. This situation led teachers to become equal partners in a narrative conversation, blurring the roles of observee and observer. One example illustrates this phenomenon:

Group 2 Observee: "When they use manipulatives, they think that they are playing but they also learn."

Group 2 Observer: "They look like a bunch of architects in there."

Effects of Peer Coaching on Students' Mathematics Achievement

Multiple regression analysis of pre- and posttest mathematics achievement scores and the (MIP) treatment effects revealed no statistically significant interaction effect between the pretest scores and the treatment on the posttest scores. Therefore, the relation between the MIP treatment and the posttest scores was independent of the pretest scores. This lack of interaction allowed us to perform further multiple regression analysis without the interaction term to examine the main effects associated with the treatment (MIP) on the posttest scores (see Table 3). The result indicated that peer coaching as implemented through the MIP had no statistically significant effect on students' mathematics achievement. In other words, the MIP did not statistically significantly improve mathematics achievement of students whose teachers took part in peer coaching. Last, our model accounted for 31% of the

TABLE 3. Results of Multiple Regression Analysis Examining the Effects of Peer Coaching on Students' Mathematics Achievement

Variable	Effect	*SE*
Pretest scores	0.57*	0.05
Treatment (MIP vs. non-MIP)	0.15	0.33
Intercept	2.38*	0.37
R^2	0.31	—

Note. MIP = Mentored Implementation Program.
*$p < .05$.

total variance in the posttest scores. Given that this figure represents almost one third of the total variance, we were satisfied with the performance of our model.[4]

Discussion

Principal Findings

Teachers considered the MIP (peer coaching) a positive experience in their professional development. Peer partners enjoyed sharing ideas, techniques, and strategies; getting feedback from and communicating with one another; and supporting one another. Meanwhile, peer partners identified scheduling and distance as the major barriers or roadblocks to the MIP. Transcripts of the four postobservation conferences provided evidence that teachers discussed organization of learning, management of the classroom, and mathematical content and pedagogy. Collaborative interactions among peer partners during the postobservation conferences were relatively brief and could be characterized as having (a) a lack of analysis, (b) a positive tone and support, (c) a proportional pattern of talk, and (d) a lack of depth in discussion. Overall, peer partners did not challenge or question one another's classroom practices. As one of the consequences, the roles of observer and observee became blurred during the postobservation conferences. Comparison of the experimental and control groups using multiple regression analysis showed that peer coaching, as implemented through the MIP, was not associated with any improvement in mathematics achievement of students whose teachers participated in the MIP.

Connection to the Literature

Although peer partners in the MIP did have professional interactions, future researchers should ask this question: Were they truly collaborative in the sense that Goos et al. (2002) have outlined? The MIP peer partners were definitely more descriptive than they were either reflective or analytical with respect to each other's classroom observations. Questioning in the postobservation conferences was not in the form of feedback request in which teachers are encouraged to critique each other's thinking. Also, statements were not at the levels of self-disclosure and mutual monitoring as defined by Goos et al. Wenger (1998) emphasized the importance of reflection and rethinking in the community of practice. However, peer partners in the MIP lacked reflection and rethinking of each other's classroom instructions. Overall, in the present study, we concluded that even though the postobservation conferences involved social interactions among peer partners, there was neither evidence of true collaboration nor evidence of renegotiating as outlined by Wenger. The present findings serve as a good warning to researchers and educators that without substantially structuring or standardizing peer coaching, many may mistake it as a panacea for professional development.

Sparks and Bruder (1987) reported that "70 percent of the teachers said that their students were 'very likely' learning more as a result of the skills they had gained from the Peer Coaching Project" (p. 56). However, those researchers admitted that few teachers had specific examples to support their claims about greater student success. Unlike their work that relied on teachers' perception of student success, the present study specifically examined student mathematics achievement by using a pretest and posttest design, and we found no evidence of improvement in the mathematics achievement of students whose teachers were peer partners in the MIP. The present study showed that for peer coaching to improve student learning as Sparks and Bruder indicated, more than 1 year is necessary for such improvement to occur.

Loucks-Horsley, Hewson, Love, and Stiles (1998) identified curriculum replacement units, institutes, coaching and mentoring, professional networks, and courses for training of professional developers as effective formats of professional learning for teachers. Each of these formats was provided through the MIP. Transcripts from the postobservation conferences indicated that the MIP promoted professional learning. Teachers frequently referred to strategies and techniques learned from the summer institutes, and they tried to implement those by working together with their peer partners. In addition, peer partners reported that they learned new ideas through their work with fellow teachers. Therefore, the MIP can be considered an experiment to see how various possible formats of professional learning as described in Loucks-Horsley et al. can be integrated to promote professional development of teachers. Considering the present study, we suggest that even though the collaborative interactions were neither analytical nor reflective, the MIP showed some potential to become a comprehensive professional learning format for teachers, likely an effective one with certain measures or reforms.

The literature describes three types of peer coaching: technical, collegial, and challenge types (see Garmston, 1987). The MIP emphasized the first two in that it transferred what was learned in the summer institute to the classroom (technical coaching) and increased communication

about the teaching profession (collegial coaching). We emphasize that the two kinds of peer coaching through the MIP promoted collegiality and professional dialogue. Peer partners were positive about their experience with fellow teachers and particularly appreciated the sharing of ideas, techniques, and strategies; getting feedback from and communicating with one another; and supporting one another. Although peer partners dealt with the transfer of training into the classroom (technical coaching) in a superficial way, it procedurally worked together with the communication about the teaching profession (collegial coaching). In this sense, we suggest that it is not impossible for the two types of peer coaching to become a coordinated effort of professional development.

The America's Choice Model (Greene, 2004) indicated that factors such as lack of time, teacher resistance, and scheduling difficulty restrict the effectiveness of peer coaching. Time and scheduling occured as major barriers to the MIP. Peer partners often found it difficult to schedule a convenient time not only to observe but also to conference. A postobservation conference did not always directly follow a classroom observation, with some conferences occurring days or even weeks after actual classroom observations. Because our study had a rural focus, we found that distance was another major barrier to peer coaching in a rural region. Therefore, we add to the literature that rural teachers have even more difficulties to overcome as far as peer coaching is concerned.

Last, our study supports the literature on the role of ambiguity in peer coaching (Poglinco et al., 2003). In the MIP, the roles of observer and observee were often vague. Without a leader (i.e., a coach), the focus of the postobservation discussion was often unclear. The lack of clearly defined roles created particularly difficult situations for those peer partners who tried to conference over two classroom observations. In these cases, it was often unclear from the transcripts which teacher was serving as the peer coach. Therefore, our study provides more evidence to the literature on the need for clarity in defining the roles of peer partners.

Policy Implications

The main piece that appears to be missing from the peer-coaching model as implemented through the MIP is the process of helping teachers to think more deeply about their work. This aspect of the MIP could be improved through a mandated training program that would focus on conferencing strategies to encourage collaborative reflection and analysis of classroom practices. We emphasize that the need for training by experts for teachers in peer-partner relationships is also related to the issue of role ambiguity in peer coaching. Training that involves more role playing can help to clarify the role of each peer partner.

Furthermore, the critique and closer examination of practice require time, which was the one barrier to in-depth professional development consistently reported in the literature and reinforced in the findings of the present MIP study. Peer coaching that includes the key dimension of challenge as noted by Garmston (1987) would certainly emphasize collegiality as a prerequisite (as evidenced in the MIP). However, a mandate to ensure the crucial time needed seems equally essential.

From the present findings, we suggest the following requirements for the MIP specifically and for peer coaching in general: (a) teachers' attendance at the content-specific institute with a peer partner from the same school or district, (b) written agreement between teachers and principals that guarantees the accomplishment of peer-coaching requirements and the time needed to fulfill those commitments, (c) coaching and mentoring training by experts that illustrates and emphasizes collaborative interactions as well as models dialogic critique, and (d) educators' requiring peer partners to videotape the observed lessons, particularly if a postobservation conference cannot occur soon after a classroom observation.

Specifically, we emphasize that in general, video clips of classroom observations and postobservation conferences can enhance the discussion between peer partners and help them to understand their roles as observer and observee. By fast-forwarding or rewinding when necessary, peer partners can view video clips and discuss specific aspects of the lesson in real time rather than struggling to recall it from memory. Such access to video would have been particularly helpful for Groups 3 and 4 in our study because they were unable to conference soon after their classroom observations. Overall, reflection and analysis become easier with the help of video clips, leading to meaningful collaborative interactions that promote professional growth.

On the basis of the results of our study, we believe that the school administration plays a critical role in the success of peer coaching such as the MIP, particularly in rural regions where teachers face more challenges to adopt peer coaching as a way of professional development. Principals can help to effectively solve two of the three major roadblocks to the MIP identified in our study—scheduling and time—by providing teachers with more flexibility for daily routine and more time for observation and conference. The use of technology such as video conferencing and perhaps even *blogging* (a technique that provides an online text box structure for easy entry of content including video and audio files for quick publishing) holds promise to overcome the challenge of distance, another major roadblock to the MIP.

Limitations

Our study has a couple of limitations. Primarily, participation in our study was voluntary, and the recruitment was difficult. Teachers had to be willing to collect student assent and parental consent forms and administer pretests and posttests to their students, in addition to classroom observations and postobservation conferences. Therefore,

participation largely depended on personal connections. Although recruitment through personal connections is an acceptable strategy in qualitative research, it can introduce bias to quantitative research on mathematics achievement of students for those volunteering teachers. To address this issue in the present study, we could only pick teachers for the control group from schools of a similar background. Such a strategy may alleviate but cannot eliminate the selection bias in the quantitative component. Time was an additional limitation. Pretest and posttest scores were collected from a single semester (about 6 months), and our entire study spanned 1 year of implementation of the MIP. Although the first 6 months appeared to be a time for peer partners to settle into the MIP, the overall duration of 1 year is rather short, particularly in terms of improvement in student mathematics achievement.

Recommendations for Further Research

As we mentioned, a longer longitudinal study of the impact of the MIP on teachers and students is warranted for a better understanding of peer coaching. Furthermore, in the MIP, teachers were asked to visit their peer partners a minimum of two times during the school year, giving peer partners a total of four times to meet and conference with each other. We question whether this is enough time for teachers to form a partnership in which they feel comfortable critiquing each other's classroom practices. We suggest further researchers determine what would be considered an appropriate number of observations and conferences that build and strengthen the peer-coaching relationship.

How can teachers serving as peer coaches challenge one another to be critical, reflective, and analytical of their own classroom practices? This is an important question for future research. We have no doubt that the training of peer partners by experts is imperative for the success of peer coaching. Peer partners need to learn appropriate conferencing skills and strategies that enable them to become effective coaches and mentors for their peer partners during the postobservation conference. But future researchers need to design and test specific training programs for this purpose.

Last, teacher mentoring and peer coaching continue to be competitive concepts of professional development. To determine the value of peer–coach relationships, we recommend a comparative study among three separate groups: (a) teachers working in the classroom without outside support, (b) teachers in teacher–mentor relationships, and (c) teachers in peer–coach relationships. We believe that comparing these three specific situations will help to clarify the unique contribution of peer coaching to teachers' professional development.

NOTES

1. All 9 teachers in the experimental group participated in the qualitative component of our study. For the quantitative component that focused on students of these teachers, some of the 9 teachers did not contribute students. In all, 2 teachers had senior high school students who were not suitable for the mathematics achievement test that we adopted. Also, 1 teacher was a collaborative special education teacher who did not have a stable cohort of students. Therefore, 6 (out of 9) teachers in the experimental group contributed students for the quantitative component of our study.

2. The multiple regression and correlation (MRC) approach to ANCOVA has meaning about the analytical hierarchy: MRC is a more general statistical framework than ANCOVA (see Cohen & Cohen, 1983). In other words, as a special case of MRC, all analytical functions of ANCOVA can be achieved in MRC. The reason why we chose MRC was that it more readily offers insight into how well our model fit the data (i.e., the R^2 statistic). A model-data-fit evaluation is important when making knowledge claims.

3. Instead of using responses from the 9 teachers in the experimental group only, we used responses from all teachers participating in the MIP to achieve a much larger sample. Consequently, the results were much more representative of teachers' perceptions of the MIP in rural regions.

4. Gaur and Gaur (2006) stated that "while in natural science research it is not uncommon to get *R* square values as high as 0.99, a much lower value (0.10–0.20) of *R* square is acceptable in social science research" (p. 109). Therefore, we were very satisfied with the performance of our model.

REFERENCES

Ackland, R. (1991). A review of the peer coaching literature. *Journal of Staff Development, 12*(1), 22–27.

Bowman, C. L., & McCormick, S. (2000). Comparison of peer coaching versus traditional supervision effects. *The Journal of Educational Research, 93,* 256–261.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correction analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Conley, S., Bas-Isaac, E., & Scull, R. (1995). Teacher mentoring and peer coaching: A micropolitical interpretation. *Journal of Personnel Evaluation in Education, 9,* 7–19.

Garmston, R. (1987). How administrators support peer coaching. *Educational Leadership, 44*(5), 18–26.

Gaur, A. S., & Gaur, S. S. (2006). *Statistical methods for practice and research: A guide to data analysis using SPSS*. Thousand Oaks, CA: Sage.

Goos, M., Galbraith, P., & Renshaw, P. (2002). Socially mediated metacognition: Creating collaborative zones of proximal development in small group problem solving. *Educational Studies in Mathematics, 49,* 193–223.

Greene, T. (2004). *Literature review for school-based staff developers and coaches*. Oxford, OH: National Staff Development Council.

Guskey, T. (1999). Apply time with wisdom. *Journal of Staff Development, 20*(2), 10–15.

Guskey, T. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin.

Hassel, E. (1999). *Professional development: Learning from the best*. Oak Brook, IL: North Central Regional Educational Laboratory.

Joyce, B. R., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Kentucky Department of Education. (2005). *Core content for mathematics assessment*. Frankfort, KY: Author.

Kohler, F. W., Crilley, K. M., Shearer, D. D., & Good, G. (1997). Effects for peer coaching on teacher and student outcomes. *The Journal of Educational Research, 90,* 240–250.

Loucks-Horsley, S., Hewson, P. W., Love, N., & Stiles, K. E. (1998). *Designing professional development for teachers of science and mathematics*. Thousand Oaks, CA: Corwin.

Lyons, C. A., & Pinnell, G. S. (2001). *Systems for change in literacy education*. Portsmouth, NH: Heinemann.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Organization for Economic Cooperation and Development. (2000). *Measuring student knowledge and skills: The PISA 2000 assessment of reading, mathematical and scientific literacy*. Paris: Author.

Organization for Economic Cooperation and Development. (2003). *The PISA 2003 assessment framework: Mathematics, reading, science and problem solving*. Paris: Author.

Poglinco, S. M., Bach, A. J., Hovde, K., Rosenblum, S., Saunders, M., & Supovitz, J. (2003). *The heart of the matter: The coaching model in America's Choice schools*. Philadelphia: Graduate School of Education, University of Pennsylvania, Philadelphia.
Romagnano, L. (1994). *Wrestling with change*. Portsmouth, NH: Heinemann.
Russo, A. (2004). School-based coaching: A revolution in professional development or just a fad? *Harvard Education Letter, 20*(4), 1–3.
Showers, B., Joyce, B. R., & Bennett, B. (1987). Synthesis of research on staff development: A framework for future and state-of-the-art analysis. *Educational Leadership, 45*(3), 77–87.
Sparks, G. M., & Bruder, S. (1987). Before and after peer coaching. *Educational Leadership, 42*(3), 54–57.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
Wenger, E. (1998). *Communities of practice: Learning, meaning and identity*. New York: Cambridge University Press.
West, L., & Staub, F. C. (2003). *Content-focused coaching: Transforming mathematics lessons*. Portsmouth, NH: Heinemann.