Analyzing CLASSROOM PRACTICE

Fusing Lesson Study and Authentic Achievement: A Model for Teacher Collaboration

After several years of working with school districts on various systemic reform efforts, Mr. Stewart and Mr. Brendefur conclude that the best way to bring about positive change at the classroom level is to adopt a model wherein small groups of teachers work in collaborative learning communities focused on improving day-to-day instruction.

BY ROGER A. STEWART AND JONATHAN L. BRENDEFUR

In his February 2004 Kappan article, Mike Schmoker called into question the effectiveness of many of the major reform efforts of the past two decades. In particular, he made the case that strategic planning efforts and whole-school reform efforts have been mostly ineffective because planning cycles are too long and implementation is too complex and cumbersome, leading to “fragmentation and overload.” To solve these problems and to make schools work better for students, he recommended that small groups of teachers work collaboratively on relatively short-term goals. He wrote, “We must replace complex, long-term plans with simple plans that focus on actual teaching lessons and units created in true ‘learning communities’ that promote team-based, short-term thought and action.”

As we read Schmoker’s article, we nodded our heads in agreement; we had experienced exactly what he described and had arrived at similar conclusions. Over the course of several years, we had worked with a group of districts, helping them through a series of improvement cycles. They were supported in their efforts by the J. A. and Kathryn Albertson Foundation, a private, Idaho-based foundation for which we consulted. The districts started with strategic planning and implementation. We then worked with them through curriculum alignment, based on Fenwick English’s work; data-driven decision making, based on Mike Schmoker’s work; and assessment literacy, based on Richard Stiggins’ work.2 When we detected that only some administrators and teachers in the districts were “getting on board,” we decided that the problem was at the administrative level and that we needed to introduce the administrators to the bigger picture of school reform, so we brought in Michael Fullan’s work.3 When we realized that teachers were not consistently and systematically communicating and collaborating, we saw the need to develop professional learning communities, so we brought in Richard DuFour and Robert Eaker’s work.4

But as the years went by, we saw little change in test scores and actual classroom practice, and as the data stream continued to show little improvement in the
schools, we became increasingly concerned that the reform efforts were having “only the most negligible impact” on the central work of the schools: instruction. Although the districts had successfully expanded their capacity, we realized that we were still witnessing a lack of clarity and coherence. The systems we were working with were fragmented and overloaded, and the time lines were too long. Some teachers and administrators felt that there was too much going on. Others wanted to opt out because they could see little direct connection between the reform efforts and their day-to-day working lives.

We set about to remedy these problems, and our remedy was exactly what Schmoker recommended. We decided we needed a model that would get to the classroom level as quickly as possible and start deep conversations about curriculum, instruction, and student learning. What we had going for us that previous reformers had lacked was the rich background outlined above. We could draw from all of the work of the likes of Schmoker, Fullan, DuFour, Eaker, English, and Stiggins and synthesize a model that had the highest probability of creating change at the classroom level.

What we came up with was a fusion of lesson study and Fred Newmann’s authentic achievement. This combination synthesized much of the work we had drawn on before, but this time we put all of the ideas and directives into a manageable package. Teachers would come together and work collaboratively on that which interested and engaged them most, namely, their day-to-day instruction.

LESSON STUDY

The districts we were working with had engaged in important conversations about data, about curriculum alignment, and about student-centered assessment. But, despite the value of these conversations, there was little evidence that they were having an impact on student achievement, because teachers’ day-to-day classroom practices were not changing — at least not enough. All the reforms these districts had been exposed to were failing to get down to the classroom level to any great degree. What was needed was a vehicle to foster close examination of the teaching process by those intimately involved in that process, the teachers.

We turned to the lesson study model, wherein a collaborative group of teachers follows a recursive cycle of articulating a clear lesson plan, observing the teaching of this lesson, and discussing and revising the plan and instruction. Data from high-stakes tests can, of course, inform and perhaps direct this cycle, but the focus remains at the classroom instructional level — what teachers and students do to achieve specific instructional objectives.

The process of developing a detailed lesson plan and then discussing it with a group of colleagues proved to be a powerful tool for instructional improvement. Once the lesson plan had been critiqued, the teacher taught the lesson, which was videotaped in progress. Working with another teacher during a planning period or getting released time to do the videotaping served as a good way for teachers to start observing one another’s practice and to become involved in something that could contribute directly to their own effectiveness. When the groups met to view the videotapes and discuss the actual delivery of the lesson, rich conversations about what had gone well and what had not turned out as planned were common. Revisions to the lesson plan became a natural consequence of these discussions. Teachers who did not have the luxury of reteaching the lesson in its revised form because they had to keep moving through their curricula often said that they were looking forward to teaching the lesson the following school year to see how much better it would be. For those team members who could go back and reteach the revised lesson and see the fruits of their collaborative labors, the results were often quite striking. The lesson improved dramatically.

But lesson study was only one part of the model. From our prior experience working with groups of teachers on a variety of reform initiatives, we knew that structuring their conversations was critically important. But we also knew that too much structure would be detrimental. The teachers needed coherent but subtle guidance that would allow them to truly explore the depth and breadth of their practice. The guidance had to be intimately related to the day-to-day work of the classroom, namely, instruction and assessment. Our focal point, therefore, became Newmann’s model of authentic achievement.

NEWMANN’S AUTHENTIC INSTRUCTION AND ASSESSMENT

Newmann and his colleagues have worked with a variety of schools and districts around the country, helping them understand and incorporate specific criteria for authentic academic achievement into their curriculum, instruction, and assessment. The results have been
quite convincing. Student achievement improves when teachers focus on three criteria: “construction of knowledge,” “disciplined inquiry,” and “value beyond school.”

These criteria are derived from an analysis of significant adult accomplishments. Newmann and Gary Wghan use the example of designing a bridge. In order to design a bridge, an adult employs all three criteria. Construction of knowledge occurs when information is organized, synthesized, interpreted, and evaluated — all higher-order thinking activities. Disciplined inquiry is the process that leads to “in-depth understanding.” As the bridge designer works, he or she draws deeply on an established base of knowledge and theory to solve significant and meaningful problems. Elaborated communication is the product of this intellectual work. The designer creates detailed plans and may write extensively or communicate in other ways about the bridge. All of these products, including the bridge itself, “have an impact on others beyond simply demonstrating that [the designer is] competent.” In the context of the classroom, this broader purpose is what is meant by “value beyond school.”

Newmann and his colleagues’ classroom model includes authentic instruction, authentic assessment, and authentic student performance. The three criteria underpin all of these components in the same way. Whether a teacher is developing a lesson plan or task, teaching a lesson, designing an assessment, or evaluating student performance, the goal is to incorporate construction of knowledge, disciplined inquiry, and value beyond school. Where this has been accomplished, student achievement improves.

How do we as teachers know when these criteria have been met? For each component of the model, Newmann and his colleagues have developed standards by which to gauge our efforts. The standards for authentic instruction, for example, include higher-order thinking, deep knowledge, substantive conversation, and connection to the world beyond the classroom. Higher-order thinking is a subcriterion of construction of knowledge. Deep knowledge and substantive conversation are subcriteria of disciplined inquiry. And connection to the world beyond the classroom is a direct corollary to value beyond school.

Some educators who have adopted Newmann’s model have reduced all of this information into a series of rubrics that can be used during the stages of lesson planning, lesson delivery, and assessment of student work to ascertain what level of expectations a lesson plan manifests, what level the lesson delivery (i.e., instruction) articulates, and what level students achieve as evidenced by their work. The partial rubric presented in Figure 1 illustrates how the model is operationalized in the world of the classroom. This particular rubric was created to evaluate the quality of the mathematics tasks a teacher has developed, but there are rubrics for other content areas, for evaluating instructional delivery, and for evaluating student performance and work. The full version of the rubric shown in Figure 1 is available at the website http://csi.boisestate.edu/ilt/rubrics.htm, which also includes a variety of other rubrics. Quest High School in Humble, Texas, a magnet school of choice that worked closely with Newmann and with us, was kind enough to allow us to use the rubrics its staff developed. All of these rubrics include variations on the three criteria and their subcriteria. For example, in its com-

| FIGURE 1. Partial Rubric for Scoring Mathematics Tasks | | 
|---|---|---|
| Construction of Knowledge | Elaborated Written Mathematical Communication |
| 4 | N/A | Analysis/Persuasion/Theory. Explicit call for generalization AND support. The task requires the student to show his/her solution path, AND to explain the solution path with evidence such as models or examples. |
| 3 | The task’s dominant expectation is for students to interpret, analyze, synthesize, or evaluate information, rather than merely to reproduce information. | Report/Summary. Call for generalization OR support. The task asks students, using narrative or expository writing, either to draw conclusions or make generalizations or arguments, OR to offer examples, summaries, illustrations, details, or reasons, but not both. |
| 2 | There is some expectation for students to interpret, analyze, synthesize, or evaluate information, rather than merely to reproduce information. | Short-answer exercises. The task or its parts can be answered with only one or two sentences, clauses, or phrasal fragments that complete a thought. Students may be asked to show some work or give some examples, but this is not emphasized, and not much detail is requested. |
| 1 | There is very little or no expectation for students to interpret, analyze, synthesize, or evaluate information. The task’s dominant expectation is for students to retrieve or reproduce fragments of knowledge or to repeatedly apply previously learned algorithms and procedures. | Fill-in-the-blank or multiple-choice exercises. The task requires no extended writing, only giving mathematical answers or definitions. |
complete form, the rubric that is partially shown in Figure 1 has four column headings: Construction of Knowledge, Elaborated Written Mathematical Communication, Disciplinary Content, and Connection to Students’ Lives. All of these headings represent either one of the criteria themselves or one of their subcriteria (e.g., Elaborated Written Mathematical Communication is an element of Disciplined Inquiry). Most of the rubrics are supplemented by “tips for scoring.” For the first column of the rubric in Figure 1, for example, there are “Tips for Scoring Construction of Knowledge,” one of which states, “Possible indicators of mathematical interpretation are tasks that ask students to consider alternative solutions or strategies, to create their own mathematical problems, to create a mathematical generalization or abstraction, or to invent their own solution methods.” We found the tips to be quite helpful to teachers as they deliberated on which level of the rubric best fit what they were evaluating.

JAPANESE LESSON STUDY MEETS NEUMANN’S PRINCIPLES OF INTELLECTUAL QUALITY

We combined the models of lesson study and Neumann’s authentic achievement to create a new model for instructional learning teams. Teachers would come together around the lesson study process and use Neumann’s criteria and the rubrics developed from them to structure and focus their conversations. We knew that the model would have to be quite flexible, since we would be forming numerous teams in the districts we were working with. Some teams would be made up of teachers from the same grade level in the same school, but others would cut across grade levels and schools. Secondary teams would have members from multiple content areas and perhaps even grade levels.

If the teachers on a team were from one grade level in the same school and were all focused on the same lesson, then they could follow the “ideal” lesson study model quite closely. They could work on a single lesson, teach the lesson, evaluate how it went, revise it accordingly, and then teach it again to see if improvements happened. When a team cut across grade levels or content areas, an individual teacher could bring a lesson to the group for critique. He or she could then teach the lesson and videotape it, and then the group could watch the tape and provide feedback. When student work came in based on the lesson, the group could also critique it, using the rubrics. No matter what the makeup of the group, the rubrics would always be the centerpiece of the subsequent conversations.

WHAT WE LEARNED ABOUT SUCCESSFUL GROUPS

Over 50 teams of teachers were formed in the 13 districts we worked with around our state. Teams met every week or every other week, depending on schedules and time of year. Sometimes two weeks would go by between meetings if report cards were due or parent/teacher conferences were coming up. We monitored the teams’ progress through questionnaires administered before the project began and at its completion, and through portfolios that the teachers submitted at the conclusion of the school year. The portfolios included reflections on the quality and impact of the experience, examples of lesson plans as they evolved, and examples of student work. In addition to monitoring the teams using these methods, we decided to follow five teams more closely. We chose ones that were close to our homes so that we could regularly observe the groups and collect artifacts like lesson plans, videotapes, and student work.

Our model was no panacea for the challenges facing public schools today, but it worked well in many instances. For example, one middle school team was so successful and its members felt so strongly about the experience that they, along with their district instructional coach, convinced the other feeder middle school and the high school to adopt the model for their professional development the following year. The administrators of the three buildings approached the superintendent for permission to be released from the regular school inservice days. That way, the teachers could form interdisciplinary teams and meet on their own time. For example, teams met on Saturdays for breakfast and before and after school to work on lesson study using authentic achievement criteria. In return, they were given the regular inservice days off. Because this particular district was not scoring well on high-stakes accountability tests in the area of mathematics, the teams used the model to focus on numeracy across the curriculum.

Some salient features of successful teamwork began to emerge. A focus on the rubrics was a common characteristic of successful groups. Groups had a tendency at times to wander in their conversations and either talk about specific students’ progress or talk in generalities that did not lead to focused ideas for improving a lesson. But those groups that were successful either had a leader who kept them on task or had a natural tendency to defer to the rubrics for guidance. As the conversation would wander, a member of the group...
might say, “Hey, let’s take a look at the rubric and see what it shows us about the videotape.” Teachers mentioned the importance of the rubrics in their reflections:

We learned to stick to the language of the rubric and offer ways to increase the quality of each person’s instruction. We had a powerful, thoughtful, and helpful team.

*   *   *

The rubrics were useful in scoring tasks as well as student work. They were beneficial in eliminating bias and promoting more objective scoring.

*   *   *

By the time we put our lessons together, we were beginning to think with the rubrics in mind.

The successful groups we observed were able to stay focused on the rubrics, but not so rigidly that conversation was stifled or narrowed too much. Instead, the rubrics served to open up discussion. Often there was debate over just where something fell on the scale. Successful teams used the scoring tips mentioned above, along with their personal opinions and experience, to help them adjudicate disagreements. All of this was done in a collegial spirit, knowing that determining right and wrong answers was not the goal. Instead it was the thinking, reflection, and dialogue about practice that became important. As teams gained experience working with the rubrics, their discourse became increasingly sophisticated, using the language and constructs of the model.

Another important factor was adherence to the model. Lessons could not always be retaught because many of the teachers were at secondary levels and could not go back and reteach content — they had to keep moving forward in their curricula. Elementary teachers also felt the pressure to move forward. But it was still important for the groups to complete the cycle as outlined in our model. Successful groups started out by critiquing a lesson plan. Once the plan had been taught in the classroom, the group scheduled a viewing of the videotape at its next meeting to critique the instruction. A similar schedule was followed when student work came in that directly reflected on the lesson plan and instruction. At the conclusion of this cycle, the lesson plan and instruction had been thoroughly critiqued from multiple angles, and revisions had been made for the next time the lesson would be taught. It was the accrual of insights from all three stages of the cycle that added up to enough “bang for the buck” to get groups hooked on the process. The experience of one group we observed illustrates the need to carry out the entire cycle.

This group was made up of middle school teachers from various disciplines who all taught eighth grade. A highly accomplished social studies teacher came to a meeting with a lesson on the Renaissance to be critiqued. It scored quite high on the criteria. Some adjustments were suggested, but nothing extensive. At the following meeting, the group watched the videotape of the teacher teaching the lesson. The instruction also scored quite high on the rubric. Everything was looking pretty good — strong lesson followed by strong instruction based on the lesson. A week after the group viewed the videotape, the student work that was a direct outgrowth of this lesson was available for critique.

As the group sat in its circle, each member perused a half dozen or so pieces of student work. The assignment was an essay asking the students to describe the Renaissance in Italy to a friend or relative from another part of the world. The students could use dialogue and be creative in their responses, but as we read we discovered that most of the student work was quite drab. There was little insight into how people lived during the Renaissance or what they believed and felt. We were all quite surprised, but no one more so than the teacher who had developed and taught the lesson. She immediately started thinking about what she needed to do the next year to make this lesson more effective. Her colleagues chimed in, and in 25 minutes or so of discussion, the lesson was significantly revised with additional steps, more in-depth content, richer discussions, and much more teacher modeling of the type of writing she wanted from the students. As in this example, student work often served as the litmus test for the overall quality of the lesson and its delivery.

The teachers we observed were not accustomed to writing detailed lesson plans. The small block in a plan book and teachers’ memories appeared to be the dominant repositories for the day’s lessons. Thus when teachers became part of learning teams, they had to adjust their practice for more public consumption. Often the first lesson a teacher brought to the group was so cryptic that the other members had to question the presenter extensively to get enough details about the plan to start making suggestions. This was not all that bad, because the questions opened up all sorts of ideas, and the probing pointed out to the presenter that the lesson plan lacked focus and clarity and thus important details about procedures and outcomes. Often the teacher would return to the next meeting with a more elabo-
rately written plan that the group could really sink its teeth into, but in a collaborative and supportive way. All of the teachers we observed had the ability to produce such plans without any coaching or additional professional development.

In their reflections, teachers talked about how their planning had been improved through the process:

There is power in collaborative planning. There is value in observing colleagues teach. My thought processes were stimulated, and it helped to organize my thoughts about teaching a lesson. My focus on instruction has been brought to a higher level. This process is less teacher directed and more student centered in lesson planning. There is more focus on your desired result.

This brings us to perhaps the most important characteristic of successful groups, namely, a willingness among members to take risks. Videotaping was a surprisingly tall hurdle for some teachers and groups to overcome. We saw this as an indication of the extreme lack of collaboration and sharing that was a part of their schools’ cultures. Schools have been characterized as egg crates, where teachers in their individual rooms never come into contact with their neighbors for fear of having their shells broken.

We witnessed this phenomenon firsthand. Sonal Chokshi and Clea Fernandez have listed challenges to importing lesson study into American classrooms. One of them is that “American teachers will be too nervous and self-conscious to open their classrooms to their peers.” 2 Successful groups either overcame this fear or had little of it to begin with. Most teachers had never been videotaped, and most had never been observed by their colleagues. Thus the lesson study model was quite a departure from common practice. But many took on the challenge and in the end enjoyed being observed and observing. Some, however, could not get past their fears. We witnessed one group of high school English teachers founder at just this juncture of the model. They brought lessons to share, and they taught the lessons and provided animated commentary about their lesson delivery to the group, but none would ever volunteer to be videotaped or observed. Members of this group were also reluctant to share student work with one another. In the end the group dropped the lesson study model and disbanded.

We have no formula for success in this area. Strong group leadership that was perceived as caring and kind appeared to be important. But a professional focus on the task at hand, namely, working collaboratively to improve instruction, also appeared to help groups overcome their shyness at having their practice exposed to colleagues. As one teacher from a successful group reflected, “We bring lessons to be scrutinized by our colleagues without the fear of being ridiculed. I felt I was being mentored in the truest sense of the word. As we spent time together, we learned to value each other’s opinions.”

Another important factor in teams’ success was early insight into and discussion about how much of teaching can or should be at the higher levels of the rubrics. Most groups started the process under the naive but understandable impression that all lessons should score high on the criteria. But it did not take them long to realize that teaching was more complex than what any rubric can capture. Some lessons would score quite low on the scales, and that was just fine. These lessons were building the base from which to launch into higher levels of Newmann’s criteria. The rubrics and the discussion around them caused the teachers to realize that they needed to take their instruction systematically from lower to higher. Nothing earthshaking or new here, but this was an important insight for teachers as they learned to work together to craft their instruction in more sophisticated and strategic ways.

REFLECTIONS

We feel fortunate to have had the opportunity to be exposed to all of the insightful thinkers whose ideas infused this work. Without their cumulative knowledge in the fields of school reform, curriculum, and assessment, we would not have been able to arrive at a model that so closely parallels Schmoker’s recommendation: small groups of teachers working in collaborative learning communities focused on day-to-day instruction that leads to short-term goals of student improvement. But we are the first to admit that not all of our learning teams functioned well. As we said above, our model is no panacea, but it does provide an additional tool in the toolbox as educators continue the struggle to find ways to improve schools and raise student achievement. As one administrator said, “Here we have a master teacher talking about how much she has learned through this process. I think, if she can learn this much, what can a new teacher learn? The process is a powerful tool that an administrator can use to facilitate collaboration and instructional improvement.”

What we found is that the fusion of lesson study and Newmann’s criteria created a flexible environment in
which teachers could examine their practice. Although we encountered some of the challenges that Chokshi and Fernandez detailed, such as lack of time for lesson study or self-consciousness about openly sharing one’s practice with colleagues, overall the model functioned quite well. Successful teams realized that the process was what was powerful, not necessarily the creation of a perfect lesson plan. They also realized that it was not the number of lesson plans that were covered by the group but the depth and quality of the coverage that influenced their teaching the most. Teachers who were in teams that cut across grade levels or content areas talked about how they gained general insights into how to improve their teaching, even though the discussion could be about content that they did not teach or grade levels other than their own. One teacher said, “I learned one thing after another that helped me in my own lesson planning. The rubrics made me aware of things I could improve across all areas of my teaching, not just the lesson plan I focused on or the plans of my team members.” Another teacher said:

When you plan to this level of detail and go through this process, it changes your way of thinking. It allows you to ask, “What are they going to know? What questions are they going to ask? How will I get them to focus on strategy instead of just the answer?” It changes the way you teach — you can anticipate. What you do here spills into other topics and makes them better too.

The teachers in these learning groups have been developing what Deanna Burney has referred to as educators’ “craft knowledge.” Burney shares our conviction that if this kind of professional growth is to occur, we must end teachers’ isolation. “Transforming thinking and practice requires people to take risks,” she notes. “They can develop their expertise only if they are willing to experiment, make mistakes, and analyze those mistakes — with everyone else and in front of everyone else. There is no other way for new knowledge to infuse the system and create stronger instructional practice.”

A student teacher who was on one of our teams confirmed the benefits of de-privatized practice. “As a preservice teacher, I think I’m lucky to be involved in this process. The team provides a great process for improving your lessons. I often think that it would be nice to teach every lesson twice. In a way this is what we’re doing. We’re teaching as a team and watching others teach. I learn a lot from this.”

Thus we have seen that teachers across the career spectrum can benefit from this approach. We close with the comments of an experienced teacher, who captures the essence of the model and its potential impact on schooling:

Perhaps the most important lesson I learned was that this process is what I expect my students to do. Being a project-based school, we ask students to develop rubrics, conduct peer critiques, analyze their work, and make revisions until they produce a quality product. We ask them to reflect on their work processes and products, always looking for strengths and areas for improvement. We require them to showcase some of their work to authentic audiences outside the school community. Being involved in the learning team asked me to do the same tasks. As their teacher, I could share my own feelings about showing my work to my peers, taking a hard look at myself, and finding the courage to present this to a real audience. I am living what I am teaching; nothing could be more powerful.

8. Ibid., p. 89.
9. Newmann, Secada, and Wehlage, op. cit.
10. See http://ghs.humblek12.tx.us for more information about the school.