

Moving the Learning of Teaching Closer to Practice: Teacher Education Implications of School-Based Inquiry Teams

Ronald Gallimore

University of California, Los Angeles

Bradley A. Ermeling
William M. Saunders

Learning Teams Group Pearson

Claude Goldenberg

Stanford University

Abstract

A 5-year prospective, quasi-experimental investigation demonstrated that grade-level teams in 9 Title 1 schools using an inquiry-focused protocol to solve instructional problems significantly increased achievement. Teachers applying the inquiry protocol shifted attribution of improved student performance to their teaching rather than external causes. This shift was achieved by focusing on an academic problem long enough to develop an instructional solution. Seeing causal connections fosters acquisition of key teaching skills and knowledge, such as identifying student needs, formulating instructional plans, and using evidence to refine instruction. These outcomes are more likely when teams are teaching similar content, led by a trained peer-facilitator, using an inquiry-focused protocol, and have stable settings in which to engage in continuous improvement.

Claims that teacher education programs have improved teaching very little and rest on a shallow knowledge base (e.g., Morris & Hiebert, 2009, in this issue) have sparked controversy and diverse perspectives (e.g., Darling-Hammond & Bransford, 2006; Sikula, 1996). Acknowledging diversity of opinion, we take the view that much remains to be done to make teacher education and its knowledge base as robust as they need to be.

In particular, we argue that teacher education is likely to benefit from closer links with school-based efforts to improve teaching (see also Lampert & Graziani, 2009, in this issue). We agree with Grossman and McDonald (2008), who recently suggested that a "stronger connection to research on teaching could inform the content of teacher education, while a stronger rela-

tionship to research on organizations and policy implementation could focus attention on the organizational contexts in which the work takes shape" (p. 184). One knowledge source for improving teaching and teacher education might be lessons learned from school-based, teacher-inquiry programs. There are two aspects of this idea, which we will deal with separately.

First, how do we construct, sustain, and support school-based settings for continuous teacher inquiry that increase student achievement? Second, what is learned about improving teaching and what are the implications for teacher educators? We were invited to address these questions by drawing on several decades of investigating teacher inquiry in school-based learning teams.

We begin with a limited review of school-based professional development research. Next, we summarize two investigations of teacher teams using a theory-based framework for recursive inquiry—analogue to the method of continuous improvement highlighted in the introduction to this issue. Last, we present and illustrate four critical operational features that sustained productive continuous improvement in our investigations of teacher teams and have since been deployed in subsequent scaling efforts in more than 180 schools in six states. We hypothesize that these four features focus teacher teams on shared problems long enough to develop solutions that improve student outcomes. Once teachers begin to attribute student gains to their own efforts, inquiry and continuous improvement are recognized as useful, satisfying, and worth the changes in school settings and routines that they require of busy educators in challenging circumstances.

The school-based inquiry process we investigated corresponds to a proposal for reorienting teacher education using a contemporary version of Dewey's (1965) idea of creating "intensive, focused opportunities to experiment with aspects of practice and then learn from that experience"

(Grossman & McDonald, 2008, pp. 189–190), Megan Franke (personal communication, July 16, 2008) characterizes this proposal as "moving the learning of teaching closer to practice," from which the title of this article is taken and gratefully acknowledged.

School-Based Teacher Learning: History and Research

"Nobody would disagree . . . that schools are primarily for the education of children. [But the] assumption that teachers can create and maintain conditions . . . stimulating for children, without those same conditions existing for teachers, has no warrant" (Sarasohn, 1972, pp. 123–124). Early empirical support for this thesis was provided by Little's (1982) case study of six schools. Teachers in successful schools more often jointly planned, designed, and evaluated instructional materials, and taught each other the practice of teaching. Interest in school-based teacher learning accelerated in the 1980s as criticism of conventional methods increased (Bird & Little, 1986).

Subsequent ideas about school-based teacher education and development, such as professional learning communities (PLCs) and teacher learning teams (LTs), can be traced to many sources (e.g., Cochran-Smith & Lytle, 1999a; Fullan, 1991; Hord, 1997; Kruse, Louis, & Bryk, 1994; Lieberman, 1995; Little & McLaughlin, 1993; Peterson, McCarthy, & Elmore, 1996; Rosenholtz, 1989; Rowan, 1990). The long-term trend is apparent in evolving standards for professional development that include school-based opportunities for teachers to jointly learn, plan, and problem solve (National Staff Development Council, 2001).

School-based teacher development has intuitive appeal that must be measured against a sobering reality. Although teachers prefer such approaches and consider them more valuable than conventional approaches (Garet, Porter, Desimone, Birman, & Yoon, 2001), there is limited evidence of impact on teaching and/or achievement

(Vescio, Ross, & Adams, 2008; Whitehurst, 2002). Vescio et al. (2008) identified 11 empirical studies that examined teaching and/or achievement changes—eight of which reported limited effects on achievement. All had design or measurement limitations, such as posttest only, no comparison groups, and so on.

In addition to the varying PLC definitions included in Vescio et al.'s (2008) review, there are other promising forms of school-based teacher learning and inquiry, including U.S. adaptations of Japanese lesson study (Lewis, Perry, & Murata, 2006), cognitively guided instruction or CGI (Kazemi & Franke, 2004), and teacher-as-researcher (Cochran-Smith & Lytle, 1999b), among others. As promising as these approaches are, similar to PLCs, their application has either been limited or the evidence base consists of case studies and demonstration projects and too few larger-scale or replication studies. This is an old problem in clinical and teaching research: to test for important effects, the researchers must provide sufficient training and sustain an experimental program long enough to achieve a sufficient implementation "dosage" (Gallimore & Santagata, 2006). Despite these limitations, PLCs, learning teams, lesson study, CGI, teacher-as-researcher, and similar approaches remain promising and worthy of serious scale-up trials.

Constructing School-Based Learning Opportunities for Teachers

Before the Beginning

Many studies of school-based teacher learning found programs before "they were 'up and running'" (Fullan, 2000, p. 4). What is missing are studies that follow schools implementing an explicit framework and document the resulting changes in student achievement—from before the beginning rather than retrospectively (Sarason, 1972).

Our "before the beginning" investigations were carried out by several practitio-

ner and researcher groups (Ermeling, 2005; Goldenberg, 2004; McDougall, Saunders, & Goldenberg, 2007; Saunders & Goldenberg, 2005; Saunders, Goldenberg, & Gallimore, in press; Tharp & Gallimore, 1988). Like others, we struggled to implement and sustain teaching changes in challenged schools. We experimented with different approaches to professional development (e.g., Goldenberg, 1992/93; Saunders & Goldenberg, 1999). However, teaching changes were short-lived, limited to a few teachers, or difficult to scale.

One barrier was the lack of a context in the schools where these studies were conducted in which teachers could collaboratively and persistently work on improving instruction. Like most schools that have department and/or grade-level meetings, teaching and learning were seldom on the agenda, and continuous improvement was rarely practiced. Creating alternatives to conventional settings was difficult because of limited resources, and were not durable because they were temporary "projects" and not rooted in school or district culture. Would student achievement increase if familiar settings such as grade-level meetings were regularly devoted to "focused opportunities to experiment with aspects of practice" (Grossman & McDonald, 2008, pp. 189–190)? How might they be reoriented to sustain such inquiry into teaching? What effects might this have on student achievement?

Framework for Transforming School Settings into Teacher Learning Opportunities

Overview. The aforementioned line of questioning led us to focus on the before-the-beginning problem (Sarason, 1972). Where to begin to reorient or transform grade-level and staff meetings so they provide continuing learning opportunities for teachers is a central question, one that was the focus of a 6-year, prospective case study of a struggling Title 1 elementary school in

a district serving students from immigrant Latino families (Goldenberg, 2004; Sullivan, 1994). Using the framework described next, achievement scores at the case study school rose from worst to best in the district. A detailed accounting of the achievement scores from before the beginning of the case study through the beginning, middle, and end phases is presented elsewhere (Goldenberg, 2004).

Framework description. The framework that guided the 6-year prospective case study originally comprised four elements: (1) *goals* that are set and shared, (2) meaningful *indicators* that measure progress, (3) *assistance* by capable others from within and outside the school, and (4) distributed *leadership* that supports and pressures goal attainment. These four elements are staples of the school improvement and "effective schools" literature of the 1970s and 1980s (e.g., Bliss, Firestone, & Richards, 1991). These elements also are found in the outside-of-education knowledge-building systems described in the introductory essay (Morris & Hiebert, 2009, in this issue). A fifth element we added to the framework was *setting*. By setting we mean, very generally, "any instance in which two or more people come together in new relationships over a sustained period of time in order to achieve certain goals" (Sarason, 1972, p. 1). Settings provide the context in which the other four elements of the framework can be enacted. Later in this article, we return to consider some more general theoretical aspects of the settings construct.

The framework predicted that acting in concert, these "change elements" could affect teachers' attitudes and cognitions and lead to instructional changes that positively affect student learning and achievement (Goldenberg, 2004). One expected change was an increase in teacher attribution of better student outcomes to the improved instruction that their team had developed. We return to this idea when we consider implications for teacher education.

Settings for teacher learning and in-

quiry. The focal settings for continuous teacher inquiry were grade-level teams, an instructional leadership team, and faculty meetings. The instructional leadership team (ILT) was composed of grade-level representatives, a reading coach, the principal, and a researcher. The ILT provided a school-wide setting for sustaining and supporting grade-level teams.

A portion of grade-level meeting time was to be routinely devoted to inquiry and trying out instruction to address common student academic needs. The intent was to get teachers to put aside for 2–4 hours per month their individual work and concerns to work as a team and solve an academic problem they agreed their students shared. Teams were trained to use a recursive cycle of collaborative inquiry—set and share an explicit goal for student learning, jointly plan instruction to address it, implement the plan, use common assessments to track student work to monitor progress, and either move on to a new shared goal or cycle back if the current one had not been reached. Many of these inquiry cycle phases overlap with practices in lesson study, CGI, teacher-as-researcher, and other reflective teaching approaches cited earlier as well as continuous improvement processes used in other fields (Morris & Hiebert, 2009, in this issue).

Grade-level teams struggling with the inquiry/improvement cycle received assistance from a researcher, the principal, or teachers who sat on the ILT. The principal refocused portions of faculty meeting time previously devoted to school operations to improving teaching and learning, such as reviewing and analyzing periodic and end-of-year assessments, and sharing instructional plans and materials.

Scale-Up Study

Transforming grade-level and faculty meetings to support inquiry and instructional improvement seemed to have had an effect in the case study school. But would

such an approach scale? Would it scale without the external assistance the researchers had provided the case study school? Could principals use knowledge developed in the case study to set up teams deploying the recursive inquiry cycle? To our knowledge, no study had tested, prospectively and experimentally, whether common school contexts can be transformed into settings for this kind of teacher inquiry focused on improving student achievement. This was the goal of the scale-up study.

Fifteen Title 1 schools serving 14,000 mostly low achieving, limited English proficient students participated in the scale-up study. All 15 schools were required by their district to adopt a school-improvement model: Nine voluntarily chose our framework, and six comparison schools chose from a district-approved list of national and local models. At the beginning of the study, there were no statistically significant differences between the two groups on the Stanford 9 and demographic factors (McDougall et al., 2007; Saunders et al., in press).

During the first 2 project years, researchers (authors W. Saunders and C. Goldenberg) provided 2-hour monthly training sessions for principals in the nine scale-up schools; principals were tasked by their superintendent to use this information to implement the case study framework, including setting up grade-level teams working with our recursive inquiry process.

After 2 years, there were no improvements in achievement in the nine schools implementing the framework relative to either comparison schools or the district. This was not surprising given the lack of implementation achieved. Researcher observations and principals' reports indicated framework-specified ILTs and grade-level settings were never established, met infrequently, or were not sufficiently focused to initiate and/or complete inquiry cycles. This principals-only approach turned out

to be phase 1 of the scale-up study because it did not work and a phase 2 intervention was developed.

Would the framework that guided the case study be effective with an increase in direct assistance to principals and teachers? For the final 3 years (phase 2), the design was significantly augmented to include summer (2.5 days) and winter (1 day) institutes and external assistance to leadership teams (ILTs) from each school (the principal and teacher facilitators from each grade level). Based on accumulated experience, the inquiry process used in the case study and phase 1 was articulated as a formal protocol, published in a manual that included protocols for ILTs as well, and distributed to all participants. The published protocol for grade-level teams set out a several-step process for recursively identifying shared student academic problems, developing and planning instruction, and analyzing student work (protocol available in McDougall et al., 2007). The protocol structured but did not prescribe the focus or direction of grade-level team inquiry, and instructional improvement was applicable to any area of the curriculum and could be repeated as each student need was successfully addressed. Once schools opened, research staff attended monthly ILT meetings at each school and met monthly with each principal individually and monthly with all principals as a group. As time permitted, researchers occasionally met with grade-level teams.

Effects of the augmented implementation became apparent during the first year of phase 2 (the project's third year). Grade-level and ILT meetings were both more consistently held and more focused on inquiry and improving instruction and learning. In the final year of phase 2 (fifth year of the scale-up study), an external evaluator assessed framework implementation at a subset of randomly selected scale-up ($n = 4$) and comparison schools ($n = 3$). Several findings distinguished meetings at scale-up and comparison schools (McDougall et al.,

2007). At the three comparison schools, meetings were more loosely structured, more frequently canceled or rescheduled, and more often addressed operations and related business matters. At the four scale-up schools, however, there was more focus in grade-level and ILT meetings on student academics, systematic and joint planning, purposeful use of assessment data (of all kinds), and efforts to implement and evaluate jointly developed instruction. Apparently, fidelity of framework implementation required external support that was midway between the rich level of external support provided in the case study and the limited amount provided in the principals-only phase 1 of the scale-up study.

What about student achievement? Over the 3 years of phase 2, scale-up schools significantly outperformed comparison schools on the Stanford 9. Starting out well below the district average before and during phase 1, scale-up schools surpassed the six comparison schools and even the district average by the end of the third and final year of phase 2. Conservatively estimated effect sizes were greater than 0.8. Comparison schools made essentially no gains relative to the district during the 5-year project (data and analysis details in Saunders et al., in press).

Due to resource limitations, the scale-up study did not attempt to assess and establish process mediators of improved achievement. Were teachers planning better lessons? Were they more aware of student needs? Did they learn and use better instructional practices in the classroom? Focus groups and interviews suggest such changes did occur, but we cannot satisfactorily specify the contribution of several plausible alternative mediators (McDougall et al., 2007; Saunders & Goldenberg, 2005; Saunders et al., in press).

More to the point of this special issue, what did we learn of possible interest to teacher educators? From informal and formal evaluation efforts, we accumulated enough evidence to offer one hypothesis

and some lessons learned (readers can see videos of teacher meetings, classroom instruction, and teacher reflections on the process used in the scale-up study at http://www.stanford.edu/~clauddeg/CD1/getting_results/index.html).

Moving the Learning of Teaching Closer to Practice: Implications of Learning Teams Investigations for Teacher Education

We hypothesize that a few critical operational features, developed and refined during the framework-guided studies of learning teams (see framework description above), allowed a form of instructional inquiry that drew teachers' attention to and helped them discover causal connections between their teaching and student performance. The importance of this hypothesis lies in our claim that seeing such cause-effect connections places teachers on a path of continuing teaching improvement. In the following sections, we first describe attribution changes observed in the scale-up study that suggest teachers began to discover causal connections. We next discuss consequences and examples of teachers making such connections. Finally, we present and discuss the four critical operational features that contributed to making causal connections.

Attribution Changes

Focus groups and interviews revealed that teacher attributions shifted once inquiry settings stabilized and teachers started (1) focusing on concrete learning goals, (2) tracking progress indicators, and, most critically, (3) getting tangible results in student learning. In contrast to comparison teachers, scale-up teachers began to attribute student gains to their own teaching to the extent that they experienced visible improvements in academic achievement associated with results-producing inquiry processes (McDougall et al., 2007). In the following excerpt, a teacher in a

scale-up school attributed improved results to the assistance teachers were providing each other to accomplish the instructional goals they had set for their students.

[A teacher reports that] . . . part of these grade-level meetings started becoming demonstration time doing little mini-lessons. . . . We shared different lessons . . . someone would demonstrate a lesson—whatever's working in the classroom. So, I think it was also very helpful for us . . . and then, when you find out what's working in one class, then you can try it in the other classroom, and so eventually everybody's doing the same thing because it works. . . . [I]t goes back to something that I said earlier. Um, yeah, with whatever pressure and whatever downside [to this process], *we are showing results*. (Field notes, scale-up school team meeting, emphasis added)

In contrast, in comparison schools, where teachers did not work on common and agreed-upon learning goals within a framework focused on student outcomes and collegial assistance, teachers were more likely to attribute achievement gains to external factors or student traits, such as socioeconomic conditions, inexperience with the English language, academic inability, or lack of parental involvement (McDougall et al., 2007). For example:

Coach distributes a one-page document with standardized test results from last year. Teachers note that third-grade scores from last year were high and relate that this year's [grade 4 students] should do well on the test this year. Male T states reason why the scores were good and his students are doing well in math is because most of the students [including his] are male. He also states that's why the reading scores are low; girls do better with reading, males do better with math. Another teacher comments [that the students] are not ready "developmentally" for writing tasks . . . [and] references their (lack of) English reading/English language proficiency. (Field notes, comparison school grade-level team meeting)

Consequences of Teachers Making Cause-Effect Connections

We believe the shift in attributions was mediated by continuous, recursive, and collective work on academic student problems until teachers had solved them—an outcome dependent on assistance from peers (and others) with support and accountability pressures from building leadership that helped maintain and focus stable settings. One principal made the case that this represented an important developmental step for teachers:

Principal: [In the learning teams work, we are] trying to break the mold or trying to break the train of thought of, "I taught the best that I [could]—I did the best job, it's these kids. They didn't learn it." [We're] trying to break that cycle and get the teachers to [change from], "I tried it this way and the kids didn't learn it" [to] maybe, "I need to modify and try it this way. And, if that doesn't work, try it this way and try it this way." But, it's too easy for them to say they just didn't get it. . . . but now what are you going to do to ensure that they do get it? . . . a lot of them don't go to that step. They just give up, [and say the students] didn't get it. Whereas . . . the good teachers will say, "OK. They didn't get it. I have to try it this way and this way." . . . that's been the hardest thing to deal with, with some teachers overall. It's . . . a crutch. It's a crutch. [They say it's] not their fault because they taught the lesson. They did a beautiful lesson. It's the kids who didn't get it. So, that's been very frustrating . . . to work with . . . some teachers. But we're getting . . . through to some of these people. . . . [and] *the grade-level meetings help*. And just the whole involvement with [the researchers] has helped also. It's just that kind of a climate at this school that we need—to just try whatever. *We need to do whatever we need to do to help these kids succeed*. (Transcript, scale-up principal interview, emphasis added)

According to this principal, sustaining inquiry led some teachers to question an assumption many initially shared: "I planned and taught the lesson, but they didn't get

it." This assumption reflects an "activity-driven" approach to instructional planning (Shavelson & Stern, 1981). Teachers in teams using the inquiry protocol began to assume "you haven't taught until they've learned," to borrow an aphorism often used by John Wooden—the greatest coach of the twentieth century—who always credited his coaching successes partly to what he learned teaching high school English (Nater & Gallimore, 2005).

We hypothesize that critical learning opportunities arise when teachers focus on a specific student need over a period of time and shift to an emphasis on *figuring out* an instructional solution that produces a detectable improvement in learning, not just *trying out* a variety of instructional activities or strategies. There are parallels between this process and the continuous-improvement approach that Morris and Hiebert (2009, in this issue) propose as one approach for building a knowledge base for teacher education and development. One teacher at the case study school who was experienced in recursive inquiry makes that point in a comment on the value of teachers persisting until problems are solved: ". . . if there's not a lot of follow-up and support [for conventional professional development], the teachers eventually are not going to be doing [what] you've taught them. Because . . . they just won't follow through. [But at our school] because there's the weekly commitment to show up to a [grade-level meeting] and work on a specific skill, . . . whatever the teachers decided to focus on for that year. There's a constant weekly support group to keep that person focused on that new skill. *And to brainstorm how to make it work if it's not working, and to push them on to the next step*" (teacher interview, case study school, emphasis added).

We claim it is not how long a team works on a problem that determines if they see a cause-effect connection, but whether they persist until it is solved. Novice groups might take months to finish their

first inquiry cycle, and complete perhaps only two during their first year. Proficient teams might learn to identify more vexing problems that thread through the curriculum and school year, work through multiple inquiry cycles, and develop instruction that builds student mastery of a complex concept that has broad performance implications.

A Caveat

These outcomes are unlikely in the absence of building leadership that supports and holds teacher teams accountable for sustaining the inquiry process until they see tangible results. Asked in a focus group what advice teachers would offer to a principal who wants to implement this process in their school, one teacher responded, "Try to build trust between the members of the group . . . and [with] the administration that this is not a way of being critical, it's a way of working for the good of the children and everyone involved" (scale-up school focus group). Another teacher provided a different, but not conflicting, perspective:

I think [it's important that the administrators stay] firm because a lot of us, we're going in different directions saying, "Well I don't know if I can do this. I don't know if my children can do this" (other teachers voice agreement). [In the beginning] everyone felt like because of some situation . . . personal to themselves, it may not work. And because administration was really firm with us and said, "Look this is what you need to do. So like it or not, do it" (various agreements among the other teachers). I was in the beginning giving my own [assessments] and the [assessments] that they were doing and then suddenly I went, "Wait a minute." As I saw the results, I was like, "Wait, okay" (Another teacher says, "Exactly."). (Scale-up school focus group)

With a balance of administrative support and pressure, teacher groups are more likely to persist with addressing problems long enough to make causal connections

between instructional decisions and achievement gains. This form of learning ramifies beyond the particular problem because it shifts teachers' focus away from what they can't control, to what they can. We hypothesize that this focused and experimental approach to a few specific problems each year fosters a "cause-effect" mindset that diffuses through daily practice and decisions, just as a drop of food coloring gradually permeates water.

Examples of Cause-Effect Connections

Installations of the learning team framework in many schools after the scale-up study provide examples of causal connections teachers make when they employ the inquiry protocol. Examples are given below.

Grade 1: Writing. A first-grade team selected learning goals focused on writing. Drawing from recently adopted state standards, the team chose to work on one first-grade standard in particular: students will write multisentence narratives about a single event. Initially, there was considerable skepticism among the group about the capacity of first graders to write compositions involving eight or more sentences. However, with the support and leadership of the teacher-facilitator in the group—who consistently encouraged her colleagues to "take a leap of faith"—the team developed several lessons and practices that over the course of the year produced positive student results. Indeed, based on the assessments developed by the teams (collaboratively scored student writings), by March the vast majority of the students were producing compositions of more than 10 sentences that developed a coherent narrative about a single event. Teachers attributed their results to daily modeling and think-aloud practices and also teacher-led sharing and feedback sessions. Both practices were familiar to the team of teachers, but over the course of the year, as they worked with each other during grade-level meet-

ings, brought in students' writing, and shared successful practices, they refined how to use modeling and sharing/feedback specifically to help students write longer but coherent narratives.

Grade 3: Mathematics. During the first quarter of the school year, a third-grade team chose to work on helping students understand multiplication as repeated addition (with arrays). Reviewing their newly adopted math instructional guide—authored by the district with the intent of pushing teachers to use rich problems to explore mathematical concepts—teachers were clearly challenged by the practice of introducing a rich problem, allowing time for groups of students to attempt solutions, and then delivering the intended directed lesson(s) about the related math concept. Indeed, teachers felt it was somehow unfair to students to "withhold" the directed lesson(s) until after students had grappled with the problem. With the assistance of the mathematics coach (a former teacher at the school), teachers developed procedures for introducing the problem, asking supportive questions while groups of students worked on the problem, leading discussions about the solutions groups developed, and explaining to students how their solutions connected to the concept of multiplication as repeated addition. For each quarter that school year, the team developed plans to use the district-developed "rich" problem, introducing the problem first and delivering the directed lessons second. End-of-year discussions with some of the teachers on the third-grade team revealed that they had altered their conceptualization of the initial introduction of the problem. Initially, they thought they were being "unfair" to students by giving them a problem before they had received related lessons. By the end of the year, they thought they were providing students with the opportunity to apply and show what they already knew about mathematics. According to some in the group, the teachers arrived at a better understanding of how to

observe students as they tried to work through the problem, and how to use those observations to deliver a more responsive subsequent lesson.

Grade 4/5: Reading comprehension. A team of fourth- and fifth-grade teachers chose to work on reading comprehension. At the time, the state had just begun assessing comprehension through open-ended responses, responses that required students to summarize grade-appropriate text and also explain the theme or main idea of the text. Three of the five teachers in the group were extremely skeptical about their students' capacity to demonstrate comprehension through written responses. They questioned whether students had sufficient writing as well as reading skills. The teacher-facilitator, however, helped the team approach the work as an investigation: "Well let's work on it and just see what happens." Through his calm and patient facilitation and with the assistance of a reading/writing consultant (working with various teams at the school), the team worked on this project over the entire year and produced significant results. While most students were producing ones and twos on their responses (scale of 1–4), in November, by the end of the year, most students were producing twos and threes. Almost every student had improved by one point. Over the course of the year, the team had tried several instructional approaches, including teacher modeling (of how to write a summary), conferencing (group and one-on-one), and paired writing (pairs writing responses together). However, they concluded that discussion about example student responses was the most effective instructional approach they had tried—that is, leading discussions in which students compared a summary scored as a 1 to a summary scored as a 2 (or a 2 compared to a 3). One of the realizations the group articulated in their final meeting of the year was that they learned how to take an objective that seemed hard for the students and break it down into manageable steps

(e.g., teaching a few things that make a level 2 summary better than a level 1 summary).

Middle school: ESL. A middle school ESL team chose to address the persistent problem of helping students use proper punctuation and capitalization. Over a period of several months (from September to January) the group worked through the protocol to collectively incorporate more journal writing activities focused on punctuation and capitalization as well as strategic reminders to reinforce punctuation and capitalization rules on daily writing assignments from the ESL text. By January, the team discovered that over 90% of the students were using capitals and end punctuation correctly, which the teachers directly attributed to the daily practice and strategic reminders. The group articulated these insights in a reflective conversation with the assistant principal:

T1: If we teach [capitalization and punctuation] right from the beginning, it is so much easier. I don't have to go around and "Oh, you forgot the capital letter, and you forgot the capital letter." And I was able to move to writing paragraphs. In the intro level, it is a big achievement because now they are writing paragraphs and I just see that they have a period and a capital letter. I don't get these long paragraphs with the "and ...and" and never ending sentence. Now I have short sentences and I see a lot of periods and capital letters . . .

AP: I'm getting the feeling that you had the "aha," like a sense of small victory, or that something went right.

T2: I think that's a good way to put it.

AP: How does that make you feel in the rest of your classes, that something went well? And I remember when Candace told me that you got 90%, how excited we all got, not only us, even down in the district, that you went through the first cycle [using the protocol] and it worked. Does that carry over into your other classes and to your other teaching? Do you feel more a sense of accomplishment?

T3: Absolutely.

T2: And more inspired to move on and realize that you can have success in other areas.

High school: Chemistry. A Chemistry team chose the challenge of helping students analyze data as they wrote conclusions for lab reports. The teachers felt that students either had never been appropriately taught or were not taking the time to complete this correctly. Therefore, the teachers built a series of small activities to help the students analyze data and write a synopsis of what the data were telling them. The teachers gave small groups different data sets, allotted time for them to study the data, and had the small groups write a brief synopsis of what the data were indicating. The students were then asked to join a larger team and combine the data sets from both groups, revise their initial synopsis as needed, and present their data to the whole class. The facilitator described the group's cause-effect findings in the following observations by a teacher:

We found when given time and structure students had no problem correctly analyzing data—it was beautiful? This is what we ended up calling “the pause.” . . . The “pause” is especially relevant in complex/multipart labs. In our second lesson of the year, we dealt with a complicated lab that had two parts. In the past, the students would just work through both parts without stopping. We rebuilt the lab completely to have the students analyze the first set of data/concept like we did previously. They studied the data, wrote a synopsis, answered some deeper questions in small groups and then joined another group to compare. After they were finished with this section, they moved onto the next part (did the same thing). Teachers reported that the student understanding was higher than past years. Breaking the lab into parts with different spots of analysis helped them with both their analysis and application questions—they were using critical thinking now!

Each of these accounts reflects the impact of teachers discovering “cause-effect” relationships. In each case, it was not only the identification of a promising strategy or the analysis of student work that produced an effect on practice, but it was sticking with a problem over time and jointly committing to specific instructional plans that produced an observable and important (to these teachers) improvement in student learning.

For the most part the professional knowledge teachers gained represented incremental improvements in how to teach. While these improvements were regarded by teachers as worthwhile for teaching something better, and a change they intended to continue, they seldom matched teaching practices aspired to by some critics of traditional instructional practices. These small and valued gains were detailed, concrete, specific, and linked to practice—qualities some argue are necessary building blocks of a professional teaching knowledge base (e.g., Hiebert, Gallimore, & Stigler, 2002).

Providing teacher teams with substantially more support and resources than the scale-up framework can achieve more dramatic changes in teaching practices, perhaps closer to those sought by critics of conventional classroom instruction such as teacher-dominated discussions. In parallel with the scale-up study, small-scale experiments conducted by Saunders and Goldenberg (1999) secured more dramatic changes, including reading-comprehension lessons with conversational features such as fewer known-answer teacher questions, and longer, more connected student speaking turns. Although these experiments produced significant gains in reading comprehension (What Works Clearinghouse, 2006), the cost of the enhanced support needed to achieve these more dramatic instructional changes greatly exceeded what was provided to scale-up schools for teacher teams and required special funding that not all schools have. This suggests

teacher teams can achieve more dramatic improvements, perhaps meaning that incremental gains need not be the only aspiration for this approach.

Features Sustaining School-Based Teacher Inquiry and Continuous Improvement

The following four operational features we hypothesize are critical to teachers sustaining and benefiting from instructional inquiry. No doubt there are others, but investigations of our learning-teams framework suggest that these four are critical because they enable teachers to work long enough on academic problems to construct causal links between classroom instruction and student learning. Through these operational features, the learning-team framework elements are enacted (goals, indicators, assistance, and leadership), leading to school-wide settings for the continuous improvement of teaching and learning (Goldenberg, 2004).

Job-alike teams. In the framework we investigated, a "learning team" or "teacher workgroup" is typically composed of three to seven individuals teaching the same grade level, course, or subject area. Absent a common task immediately relevant to each teacher's own classroom, it is difficult to create and sustain the kind of inquiry cycle observed in the scale-up schools and others in which we now work. In elementary programs, grade-level teams fulfill this function. At the secondary level, we have been most successful when teachers are organized into course-level (or subject area) teams, such as seventh-grade pre-algebra, or ninth-grade English. To be successful, teams need to set and share goals to work on that are immediately applicable to their classrooms. Without such goals, teams will drift toward superficial discussions and truncated efforts to test alternative instruction.

Trained peer-facilitators. Selecting a team facilitator is critical to sustaining in-

quiry long enough for cause-effect connections to be made by a teacher team. Selection can take many forms, depending on the context. But even the most motivated teams need a "point person"—at least one member identified and trained to guide their colleagues through the process over time. The role can be shared, and members can rotate in and out from year to year as capacity grows. Teams are more effective with peers leading rather than administrators or content experts in the facilitator role for several reasons. Peer-facilitators are uniquely positioned to model "a leap of faith," frame the work as an investigation, help the group "stick with it," and guide protocol use as a full participant in the inquiry process. Teacher-facilitators are trying out in their classrooms the same lessons as everyone else in the group. In addition, the use of teacher-facilitators frees up coaches and content experts to play a knowledgeable resource role rather than team-leader role; this significantly lessens the chances the setting is converted from inquiry-focused to a more conventional professional development (PD) "presentation" structure that puts teachers in a passive rather than active role. Distributed leadership also permits administrators to circulate and provide appropriate support and accountability for multiple teams and facilitators. The inquiry process and its requisite settings are more reliably sustained by administrator buy-in, support, and willingness to hold everyone accountable (including themselves).

Inquiry-focused protocols. The scale-up protocol we employed is similar to a number of approaches that feature recursive inquiry and continuous improvement, but might be less prescriptive in terms of the kind of instruction that is the end goal (e.g., CGI), or the number of criteria lesson planning must meet (e.g., lesson study). Our investigations suggest it is critical to define and publish a protocol that articulates specific inquiry functions: jointly and recursively identifying appropriate and worth-

while goals for student learning; finding or developing appropriate means to assess student progress toward those goals; bringing to the table the expertise of colleagues and others who can assist in accomplishing these goals; planning, preparing, and delivering lessons; using evidence from the classroom to evaluate instruction; and, finally, reflecting on the process to determine next steps.

Such knowledge and skills are often taught in teacher-preparation courses. Training peer-facilitators to use and support a protocol such as we deployed brings learning and application of these closer to classroom practice. Effective teams hone new knowledge and skills while using them in their classrooms as well as in collaborative contexts that hold everyone accountable for collecting feedback on the effects of their teaching. By using our protocol or a similar process, we hypothesize that teachers might gain several benefits:

1. In the best circumstances, the detailed recursive planning and analysis of individual lessons—related to specific goals—yields specific cause-effect insights (internal attributions) about addressing a specific area of student need.
2. By working through the cyclical process until meaningful student results are achieved, teachers develop findings and insights about teaching that can transfer beyond the specific work of the team.
3. Finally, we hypothesized that slowing down and making planning and analysis visible in a collective and intentional way affects general patterns of cognition, including:
 - a) cause-effect analyses become part of daily planning;
 - b) more attention is more consistently paid, with greater clarity and care, to particular areas of student need;
 - c) greater attention is paid to classroom interactions and artifacts as opportunities to get feedback on the effects of teachers' efforts;

- d) more regular and productive questioning of existing instructional practices;
- e) greater interest in gaining more knowledge about practice and alternative approaches; and
- f) greater reliance on evidence to drive planning and decisions.

Stable settings. For the teams to stick with the protocol long enough to see and attribute improved student learning to their teaching, there must be a stable, protected setting in which the work of inquiry can get done. Establishing stable settings for this purpose is a challenge, even when support and accountability are provided by administrators. But this is not surprising. The reorientation of settings attempted in our investigations is greatly constrained by school and district ecology. Conserving as few as 20 hours a year for teacher inquiry was nearly always a struggle for various reasons. For example, even the most effective teams are able to devote only about 75% of the allotted time to work on improving instruction; less effective ones utilize about 50% (Powell, Goldenberg, & Cano, 1995). With multiple, uncoordinated reform initiatives hitting schools, time for teacher inquiry is often sacrificed for competing demands, such as mandated PD or the responsibilities for parent and IEP conferences. The immediacy and urgency of day-to-day operations gobble up time and put everyone's commitment to the test. In candid moments, teachers battling overload and fatigue report that there are times they feel like just going home, or completing other tasks rather than attending a grade-level meeting to engage in their team's chosen inquiry.

There are cultural as well as ecological constraints. Teacher-driven inquiry conflicts with default cultural schema about how to improve instruction. For example, in one post-scale-up installation of learning teams, a skeptical district expert was invited to observe a secondary school inquiry

team. She was shocked at what she considered the group's low level of mathematical knowledge and discourse. She told teachers they were mangling the instructional program, and took over the meeting to deliver an ad hoc PD presentation, assigned the teachers a different student need to work on, modeled a lesson to address it, and assigned them to teach it. The spirit of inquiry the protocol was designed to nurture was dampened that day. Later, our project advisor was unable to persuade the expert that she was missing a diagnostic window into the teachers' actual level of functioning, perhaps an indication of the limited effect of the conventional PD that she and her unit had previously delivered. She did not recognize that the teachers needed both conventional PD presentations to deepen their pedagogical content knowledge as well as a stable inquiry and learning setting in which to convert that knowledge into better lessons and practices.

Like all institutions with a long history, public school settings evolved over time into contexts that supported practices participants came to take for granted (Elmore, 2000; Lortie, 1975). School, home, and workplace settings are observable manifestations of ecology and culture in everyday life; they are the hard-won compromises between the possible (ecological resources and constraints) and the desirable (cultural beliefs and personal values) (Gallimore, Goldenberg, & Weisner, 1993). Settings resist change because that is one of their functions; they are sources of predictability, coherence, and personal agency when social, economic, and other changes threaten adaptation and survival in both a narrow and broad sense. If change involves renegotiating compromises that make a setting stable, resistance will arise because people prefer their existing, if not perfect, daily settings and their embedded practices. Humans are satisficers, not maximizers (Simon, 1957); most prefer an adaptation that is working just well enough to an innovation that might risk reawakening stresses and con-

flicts that required the original compromises reflected in existing settings. The most common stimulus to change is a significant social or economic perturbation; a common response is making only those changes that restore as much as possible pre-perturbation settings and routines (Edgerton, 1992). This reflects a common human preference that the more things change, the more we want them to stay the same.

If existing settings, such as grade-level teams or faculty meetings, are an evolutionary product of adaptation to multiple ecological and cultural resources and constraints, reorienting them takes advantage of their durability to sustain new practices such as inquiry and continuous improvement. Not surprising, students of educational reform argue that continuously improving schools "adjust and adapt the routines of the workplace . . . with the primary purpose of creating settings where teachers, administrators, and outside experts can interact around common problems of practice" (Elmore, 2000, p. 30).

Setting-focused interventions involve culture change. Reorienting existing settings to support teacher inquiry means changing an adaptation that has evolved over time—including taken for granted assumptions about the purposes of the commonplace, such as grade-level or department meetings. It means enlarging conceptions of teacher education and development to include both individual- and setting-focused programs. Individual-focused courses and PD are needed to augment content and pedagogical content knowledge, and reorienting school settings is needed so new knowledge and skills coalesce through teacher inquiry into better classroom practices. Otherwise, as Crandall et al. (1982) documented, what's learned in individual-focused programs may be so thoroughly assimilated into existing practices that their impact is diluted.

Even in schools and districts committed to continuous improvement through inquiry, including those that have seen

achievement gains, maintaining that focus in grade-level/departmental settings remains a constant challenge. Teacher inquiry might appeal as a means of continuous professional learning, but it and its essential settings are not rooted in the ecology and culture of many U.S. districts—a development that might partly depend on parallel changes in teacher education programs.

Conclusion

The method of continuous improvement, highlighted in the introduction to this issue (Morris & Hiebert, 2009, in this issue), is unlikely to work as expected if teachers assume, "I planned and taught the lesson, but they didn't get it." Teachers shift from this stance by working on learning problems long enough to solve them and discover causal connections between their instruction and student outcomes. Seeing causal connections is afforded by stable settings and peer facilitators that support job-alike teams' use of articulated inquiry protocols that support continuous improvement of teaching. Stable settings, job-alike teams, peer facilitators, and protocols create "intensive, focused opportunities to experiment with aspects of practice" (Grossman & McDonald, 2008, pp. 189–190) that move the learning of teaching closer to practice. They nurture a more productive assumption of professional learning and continuous improvement: "You haven't taught until they've learned."

Note

We acknowledge the support of the Spencer Foundation, Office of Educational Research and Improvement, Center for Culture and Health (UCLA), and LessonLab Research Institute. Subsequent to the research summarized in this article, a program based on the results is now offered by Pearson Education, which employs three of the authors (Gallimore, Ermeling, and Saunders). The views expressed are the authors' and do not necessarily reflect those of funding agencies or Pearson Education.

References

- Bird, T., & Little, J. W. (1986). How schools organize the teaching profession. *Elementary School Journal*, 86(4), 493–512.
- Bliss, J., Firestone, W., & Richards, C. (Eds.). (1991). *Rethinking effective schools: Research and practice*. Englewood Cliffs, NJ: Prentice-Hall.
- Cochran-Smith, M., & Lytle, S. L. (1999a). Relationships of knowledge and practice: Teacher learning in communities. *Review of Research in Education*, 24(1), 249–305.
- Cochran-Smith, M., & Lytle, S. L. (1999b). The teacher researcher movement: A decade later. *Educational Researcher*, 28(7), 15–25.
- Crandall, D. P., Loucks-Horsley, S., Bauchner, J. E., Schmidt, W. B., Eiseman, J. W., Cox, P. L., et al. (1982). *People, policies, and practices: Examining the chain of school improvement*. Andover, MA: NETWORK, Inc.
- Darling-Hammond, L., & Bransford, J. (2006). *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco: Jossey-Bass.
- Dewey, J. (1965). The relation of theory to practice in education. In M. Borrowman (Ed.), *Teacher education in America: A documentary history* (pp. 140–171). New York: Teachers College Press.
- Edgerton, R. B. (1992). *Sick societies: Challenging the myth of primitive harmony*. New York: Free Press.
- Elmore, R. F. (2000). *Building a new structure for school leadership*. Washington, DC: Albert Shanker Institute.
- Ermeling, B. A. (2005). *Transforming professional development for an American high school: A lesson study inspired, technology powered system for teacher learning*. Unpublished doctoral dissertation, University of California–Los Angeles.
- Fullan, M. (1991). *The new meaning of educational change*. New York: Teachers College Press.
- Fullan, M. (2000). The three stories of educational reform. *Phi Delta Kappan*, 81(8), 581–584.
- Gallimore, R., Goldenberg, C., & Weisner, T. (1993). The social construction and subjective reality of activity settings: Implications for community psychology. *American Journal of Community Psychology*, 21(4), 537–559.
- Gallimore, R., & Santagata, R. (2006). Researching teaching: The problem of studying a system resistant to change. In R. R. Bootzin & P. E. McKnight (Eds.), *Strengthening research methodology: Psychological measurement and*

- evaluation (pp. 11–28). Washington, DC: APA Books.
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915–945.
- Goldenberg, C. N. (1992/93). Instructional conversations: Promoting comprehension through discussion. *Reading Teacher*, 46(4), 316–326.
- Goldenberg, C. N. (2004). *Successful school change: Creating settings to improve teaching and learning*. New York: Teachers College Press.
- Grossman, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal*, 45(1), 184–205.
- Hiebert, J., Gallimore, R., & Stigler, J. W. (2002). A knowledge base for the teaching profession: What would it look like and how can we get one? *Educational Researcher*, 31(3), 3–15.
- Hord, S. M. (1997). *Professional learning communities: Communities of continuous inquiry and improvement*. Austin, TX: Southwest Educational Development Laboratory.
- Kazemi, E., & Franke, M. L. (2004). Teacher learning in mathematics: Using student work to promote collective inquiry. *Journal of Mathematics Teacher Education*, 7(3), 203–235.
- Kruse, S., Louis, K., & Bryk, A. (1994). *Building professional community in schools*. Madison, WI: Center on Organization and Restructuring Schools.
- Lampert, M., & Graziani, F. (2009). Instructional activities as a tool for teachers' and teacher educators' learning. *Elementary School Journal*, 109(5), 000–000.
- Lewis, C., Perry, R., & Murata, A. (2006). How should research contribute to instructional improvement: The case of lesson study. *Educational Researcher*, 35(3), 3–14.
- Lieberman, A. (Ed.). (1995). *The work of restructuring schools: Building from the ground up*. New York: Teachers College Press.
- Little, J. (1982). Norms for collegiality and experimentation: Workplace conditions of school success. *American Educational Research Journal*, 19(3), 325–340.
- Little, J., & McLaughlin, M. (Eds.). (1993). *Teachers' work: Individuals, colleagues, and contexts*. New York: Teachers College Press.
- Lortie, D. C. (1975). *Schoolteacher*. Chicago: University of Chicago Press.
- McDougall, D., Saunders, W., & Goldenberg, C. (2007). Inside the black box of school reform: Explaining the how and why of change at Getting Results schools. *International Journal of Disability, Development and Education*, 54(1), 51–89.
- Morris, A. K., & Hiebert, J. (2009). Introduction: Building knowledge bases and improving systems of practice. *Elementary School Journal*, 109(5), 000–000.
- Nater, S., & Gallimore, R. (2005). *You haven't taught until they have learned: John Wooden's teaching principles and practices*. Morgantown, WV: Fitness International Technology.
- National Staff Development Council. (2001). *NSDC Standards for Staff Development*. Oxford, OH: National Staff Development Council. Retrieved January 28, 2006, from <http://www.nsdc.org/standards/index.cfm>
- Peterson, P., McCarthey, S., & Elmore, R. (1996). Learning from school restructuring. *American Educational Research Journal*, 33(1), 119–153.
- Powell, A., Goldenberg, C. N., & Cano, L. (1995, April). *Assisting change: Some settings for teacher development work better than others*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Rosenholtz, S. J. (1989). Workplace conditions that affect teacher quality and commitment: Implications for teacher induction programs. *Elementary School Journal*, 89(4), 421–439.
- Rowan, B. (1990). Commitment and control: Alternative strategies for the organizational design of schools. In C. Cazden (Ed.), *Review of Research in Education*, 16(1), 353–389.
- Sarason, S. (1972). *The creation of settings and the future societies*. San Francisco: Jossey-Bass.
- Saunders, W., & Goldenberg, C. (1999). The effects of instructional conversations and literature logs on limited- and fluent-English proficient students' story comprehension and thematic understanding. *Elementary School Journal*, 99(4), 277–301.
- Saunders, W., & Goldenberg, C. (2005). The contribution of settings to school improvement and school change: A case study. In C. O'Donnell & L. Yamauchi (Eds.), *Culture and context in human behavior change: Theory, research, and applications* (pp. 127–150). New York: Peter Lang.
- Saunders, W., Goldenberg, C., & Gallimore, R. (in press). Increasing achievement by fo-

- cusing grade level teams on improving classroom learning: A prospective, quasi-experimental study of Title 1 schools. *American Educational Research Journal*.
- Shavelson, R. J., & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. *Review of Educational Research*, 51(4), 455-498.
- Sikula, J. (Ed.). (1996). *Handbook of research on teacher education*. New York: Simon & Schuster.
- Simon, H. (1957). *Administrative behavior* (2nd ed.). New York: Macmillan.
- Sullivan, J. (1994). *Producing change in an urban, second-language, elementary school*. Unpublished doctoral dissertation, Graduate School of Education, University of California-Los Angeles.
- Tharp, R. G., & Gallimore, R. (1988). *Rousing minds to life*. Cambridge: Cambridge University Press.
- Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education*, 24(1), 80-91.
- What Works Clearinghouse. (2006). *Intervention: Instructional conversations and literature logs*. Institute of Education Sciences. Retrieved July 21, 2008, from http://ies.ed.gov/ncee/wwc/reports/english_lang/icll/
- Whitehurst, G. J. (2002, March). *Research on teacher preparation and professional development*. Paper presented at the White House Conference on Preparing Tomorrow's Teachers, Washington, DC.