8

HOW TO REFORM *REFORM*

Mary M. Kennedy

MICHIGAN STATE UNIVERSITY

The U.S. education system today is awash in reforms of all sorts, and especially in reforms that aim to improve the quality of teaching practice. Teacher education programs seek to revise their approaches to preparing teachers, and K-12 schools seek to revise their methods of evaluating and improving teachers' classroom practices. Such efforts are part of our education culture, for both school districts and teacher education programs have repeatedly sought to reform themselves in the past. In fact, reforms aimed at improving teaching practice have been so plentiful and so persistent over the years that historians have built careers examining the history of reform itself (Cuban, 1984, 1990; Tyack, 1987; Tyack & Cuban, 1995; Tyack & Tobin, 1994). They have offered several hypotheses for why reforms have failed to alter teaching practice – the culture of schools, the variety of purposes they are designed to serve, and constraints inherent in school life. They have used the phrase "grammar of schooling" to refer to structural constraints that limit the ability of schools to change. The grammar of schooling includes such things as the way students are grouped into batches, the way the daily schedules are set, and other things which constrain reformers who seek to change teaching practice. Years ago, I followed up on those hypotheses with a qualitative study of teaching practices (Kennedy, 2005) and found a lot of evidence to suggest that poor practices were indeed a function of that grammar. The problems teachers encountered during their lessons were more often due to organizational issues than to failures in their own knowledge or commitments. Teachers made mistakes when they had too little planning time, were distracted by non-teaching duties, had difficulty coordinating with the librarians, tech support, special education teachers and so forth. The number and variety of management and coordination tasks required of teachers distracted them from their central responsibilities, sent them into

their classrooms unprepared to teach, or disrupted their lessons with extraneous announcements and visitors.

My aim in this chapter is to look more closely at the reforms themselves, and specifically at how reformers expect their messages to influence teachers' practices. I contrast two different models of influence. In one model, reformers provide teachers with verified knowledge that is universally applicable. It can be summarized in textbooks, described in lectures, or demonstrated in a variety of ways. When reformers use this model to influence teachers, they rely on direct instruction. They often give lectures on the body of knowledge they want teachers to learn, or illustrate the kind of practices they believe teachers should use, or ask teachers to examine videorecorded lessons or other classroom artifacts that can be used to illustrate particular points. To find out whether teachers have learned their content, they sometimes give teachers tests and sometimes use observation instruments to measure changes in specific teaching practices.

In the second model, knowledge is more context-dependent. To share it with teachers, reformers have to find a way to embed it in teaching situations. They often rely on video recordings of lessons or samples of student work and they spend much of their time interrogating teachers about these artifacts so that teachers learn to see the new ideas within the situations presented. To find out whether teachers have learned the new knowledge, they look at how teachers design their lessons or how they interpret and respond to specific episodes of teaching.

In this chapter, I label the first model *knowledge transmission* and the second *knowledge development*. The knowledge-transmission model has a longer and deeper tradition within U.S. culture and especially within U.S. school culture. Curriculum knowledge is partitioned into school subjects, each with its own textbooks that can be transported from room to room. Teachers schedule lessons to cover each discrete topic and testing programs measure how much knowledge students have acquired. But the knowledge-development model has always had some adherents as well. There have always been a few schools or programs that take a more developmental approach to influencing students, always some curricula that seek to develop knowledge rather than transmit it.

This same pattern also appears in the world of reformers. The majority of reformers rely on a model of knowledge transmission but a few prefer knowledge development. Members of the first group have identified specific bodies of knowledge that they want teachers to acquire or specific changes in practice that they want to see teachers implement. Those in the second group want teachers to interpret classroom events differently and generate different kinds of lesson plans. An important difference between the two approaches to reform is that the knowledge-transmission model assumes the reformers already know the teaching practices that teachers should use whereas the knowledge developers view teaching practices as contingent on circumstances.

My plan in this chapter is to estimate the merits of these two models by examining the effectiveness of professional development (PD) programs that are

based on each model. But before looking at the programs themselves, I first examine the kind of intellectual work involved in teaching itself, in the hope that a better understanding of how teachers develop their practices in the first place might help us better understand how each of these models would likely influence their work. After examining teaching practice itself, I compare PD programs that are based on these two models. Finally, in a third section, I discuss prospects for educational reform more broadly.

What Do Teachers Actually Do?

One problem faced by both teacher educators and school reformers is how to characterize the details of teaching practice in order to talk about it. When we watch teachers, we see a seamless stream of activity, but when we discuss it, we need language that can partition that stream into coherent parts that can be analyzed. As a field, teacher education has been somewhat successful in sorting out clarified, articulated, and categories of *knowledge* that aspiring teachers might need to learn – cultural foundations, learning theory, or classroom management – and we can identify a few discrete practices, such as cooperative groups or core teaching practices. But we lack a larger map that sorts the whole of classroom practice into meaningful segments.

One problem has been finding the right "size" of teaching activity. Teachers sometimes make very small gestures that can be significant, such as patting a student on the back, and they also engage in very large and complex activities, like designing a six-week-long unit on a the solar system. In perhaps the first effort ever to partition teaching practice, Charters and Waples (1929) identified over 1,000 discrete things teachers did. Some were trivial, others extremely complex. But the list did not lend itself easily to the development of teacher education curricula.

This leads us to a second problem, which is ensuring that the pieces of practice that we sort out are indeed meaningfully relevant to the overall phenomenon we call "classroom teaching." I don't mean that they are *statistically* related to a specified outcome, but rather that they are *functionally* relevant to that outcome, that we can understand their role in the overall endeavor. Here, I identify four such functions and argue that they can be used to characterize (or to analyze or to evaluate) virtually all observable teaching events. Further, they are all requirements of the job. If teachers are unable to do any of these four things, most of us would agree that they are unable to teach in classrooms.

Representing Curriculum Content

The first function of teaching is to represent curriculum content in real time and in a way that makes it comprehensible to naïve thinkers. If students could (and would) learn content simply by reading textbooks or other available documents, there would be no need for teachers. But they do not. So teachers must find a

way to reveal content in some way, and they must do this within a specific amount of time with a specific set of materials for a specific group of children. Thus, we see teachers provide demonstrations, pictures, movies, hypothetical problems, walked-through examples. We see them posing questions and answering questions from students, and we see them asking students to engage in a variety of learning activities on their own – reading, solving problems, thought experiments, writing, gathering data, and so forth.

Notice that the decisions teachers make regarding how they will represent their content will depend heavily on the *specific set of students*, the *specific time frame*, and the *particular materials that are available*. Thus, we are likely to see variations in how teachers address this function. One teacher might use a physical example because she happens to have one available, while another might engage students in a thought experiment. One might put a diagram on the board, and ask students to label its parts or speculate about how different parts work, while yet another might ask students to engage in a group activity and still another might show a video or simply write on the board. Even teachers who teach the same curriculum from year to year are likely to change their representations over time, responding to the needs of different students, different time constraints, or to their own need for variety. But these representations will also reflect the extent to which teachers feel capable of using particular representations, whether they have sufficient planning time to develop particular representations, and whether they subscribe to the first or the second conception of knowledge that I outlined above.

Enlist Student Participation

The second function of classroom teaching is to enlist student participation. This is especially difficult because, although education is mandatory, *learning is not.* Thus, teachers face a captive audience, and sometimes a resistant audience. Cohen (1988) suggested that teachers belong in a class of "human improvement" professions, like psychotherapists or fitness trainers, in which one's success depends entirely on the clients' willingness to improve themselves. If clients do not wish to learn, lose weight, improve their golf swing, or save for retirement, then professional help will likely be ineffective. Similarly, teachers cannot succeed unless their clients, students, choose to learn.

Further, even students who are willing to participate may resist thinking about difficult or complicated topics. School learning requires what Kahneman (2011) calls "slow thinking;" the kind of thinking that requires concentration and effort. In contrast, most of life outside the classroom calls for "fast thinking," thinking that is reflexive and that allows us to jump to conclusions, rely on rules of thumb, or rely on habits formed in the past. Fast thinking is easier and is the default method used by most human beings for most activities, thus producing a persistent problem for teachers who must persuade their students to invest the time and concentration needed to engage in intentional thinking.

Notice, too, that the need to enlist student participation can compete with the need to represent content. This can happen when, for instance, the most accurate representations might be less engaging, or when the most engaging learning activities might also be less informative. Further, teachers often become resigned to the presence of one or two students who are simply unwilling to participate at all, so that their goals for these students are often just to minimize the chances that these students will interfere with the lesson. In this case, their method of enlisting participation often reflects a negotiated compromise between students, content and learning activities.

Exposing Student Thinking

The third function of classroom teaching is probably the least evident to external observers: Teachers need to continually elicit feedback from students so that they can estimate how well students are grasping new content. Each day's lesson follows from the lessons taught the day before, and students need to grasp essential points from each lesson if they are to make sense of the next lessons. Thus, we see teachers asking students to read aloud, solve problems, share their findings, respond to one another's ideas, show their work on the board, or turn in assigned projects for review. Sometimes we see teachers orchestrate more complex projects such as debates or group experiments that allow them to watch students as they develop their own nascent ideas. These interrogations are especially visible in elementary classrooms, where teachers use a constant question-and-answer format both to maintain students' attention and to assess their understanding.

This need to expose students' thinking is especially important in light of the sheer variation in how students make sense of new ideas. Generally speaking, we all make sense of new ideas by connecting them to things we already know. Students who enter with different prior experiences are likely to interpret the content in different ways as they each connect it to their own prior knowledge and experiences. A roomful of highly engaged students is actually a roomful of different ideas, conceptions, confusions, questions, and insights, so that the teacher's task calls to mind the popular analogy, *herding cats*. One reason that experienced teachers tend to be more effective than novices is that they know their audiences. They have learned what kind of confusions and misconceptions to expect from their students.

Containing Student Behavior

The fourth function of classroom teaching is to contain student behavior. This is necessary in part as a matter of public safety but also to ensure that students are not distracting each other, or distracting the teacher, from the lesson. Classrooms contain from 20–40 students in a relatively confined space. Further, these students are young, energetic, restless, and less able to control their own behavior than are adults.

Like the other functions of classroom teaching, this one also turns out to be complicated. One of the earliest studies of teaching practice (Kounin, 1970) sought to better understand what the authors called "desists," meaning teacher actions intended to stop misbehaviors. As they observed their teachers, they discovered that most classroom management consisted of *prevention* of misbehaviors rather than stopping misbehaviors once they had occurred. For instance, teachers continuously demonstrated to students that they were aware of what everyone was doing. A teacher might say, "So you can see that the area of this triangle, Christina, is ..." The teacher is speaking to the entire class, but simply by mentioning one student's name, she warns the named student to settle down and she reminds all the other students that she is alert to their individual actions as well.

Accommodating their own Personal Needs

In addition to these four functions, teachers must also satisfy a requirement that is not addressed in any policy manuals, nor is it addressed in teacher education courses, yet it is essential for teachers themselves: Teachers must find a way to address the four central functions of teaching in a way that is consistent with their own personalities and personal needs. Some teachers may need an orderly and predictable classroom, while others may enjoy more spontaneity. Some may need to reduce the overall volume of noise while others are energized by the class's energy. These are matters not only of personal taste, but often of personal need as well. If a teacher cannot find a way to create an atmosphere that she or he is comfortable living in, she or he is not likely to remain teaching for very long.

Furthermore, these required functions present different kinds of problems with each group of students and with each lesson topic, so that teachers are continuously adjusting and revising and re-thinking their strategies over time.

A complete definition of what teachers try to do, and spend their time thinking about, therefore might look like this:

- They strive to represent curriculum content in a way that meets constraints of time and space and that renders the content comprehensible to specific naïve minds;
- For students who differ greatly in their interest in learning;
- And whose grasp of the content is not readily visible to the teacher;
- And who are restless and easily distracted;
- In a way that satisfies their own personal needs.

This portrait of classroom teaching introduces two new hypotheses for why past reforms might have failed. One is that reforms might focus on only one of these functions, without regard for how the new idea might affect the others. For example, a program that emphasizes containing student behavior might inadvertently reduce student motivation to participate, or a program that aims to fully

engage students might generate so much enthusiasm that their behavior is harder to contain. When this happens, the reform will likely be abandoned because it creates as many problems as it solves.

The second hypothesis for the failure of reform has to do with the models of influence that reformers use to influence teaching practice. Teachers face different configurations of students from class to class, and different content from hour to hour, so that each lesson plan represents a unique solution to these four central functions. Today's geography lesson might be adapted to accommodate the surly boy on the left, the fact that a teacher down the hall has borrowed the globe, or the fact that half of the board is unavailable for use because it is covered with fire drill instructions. Viewing teaching in this way, we might hypothesize that reforms based on a knowledge-transmission model might be less effective than those based on a knowledge-development model, for the knowledge-transmission model does not recognize any need for situational adjustments.

With all of this in mind, then, we can now examine the methods reformers use to try to change teachers' practices.

How Do Reformers Reform Teaching?

At some point in the reform cycle, reformers must themselves become teachers. When this happens, they face the same requirements that all teachers face. They must find a way to represent their ideas so that teachers can understand them, they must enlist teacher participation in their program, and they must monitor whether or how well teachers are understanding and adopting their ideas. Further, they face the additional requirement of trying to persuade teachers to *abandon their former practices* in order to implement the reformer's new ideas.

This is where professional development, or PD, becomes a part of reform efforts, and why an examination of PD can be a useful approach for gaining insights into when and how reforms work. For, like teachers themselves, reformers have ideas about how to represent their ideas to teachers, how to enlist teacher participation, etc., and ultimately to influence teaching practice. The design of their PD programs reflects these assumptions.

I recently reviewed 27 best-evidence studies of PD (Kennedy, 2016) to learn more about the relative merits of alternative approaches to PD. By "best evidence," I mean that these studies (1) used randomization or an equally justifiable approach for forming comparison groups, rather than simply comparing participating teachers to peers with similar teaching assignments,[1] (2) followed teachers for at least a full school year, and (3) included measures of student learning in their list of outcomes.

In the past, reviews of PD studies have tended to characterize programs by observable characteristics such as their duration, intensity, or their use of coaches or technology. I focus instead on their apparent underlying assumptions about how they expect to influence teaching practice. I sorted them into two groups

which reflect the two models of influence described above. One approach, *knowledge transmission*, is based on the traditional conception of learning as a process of acquiring authoritative and unambiguous knowledge or rules. The other approach, *knowledge development*, is based on the idea that each teacher needs to independently formulate the new content and think through how this content can help them better address the four functions of their jobs. The first model of influence has historically dominated professional development programs. Programs based on the second model tend to be more recent, are often smaller, and often have less clearly defined curricula.

Among the studies I reviewed, knowledge transmission programs looked just like their name implies: PD providers gave teachers direct instruction on specific school subjects or gave them explicit guidance about specific practices. They tended to present their ideas in a lecture format but they also provided a lot of demonstrations and illustrations and included question-and-answer sessions or small-group discussions to help teachers digest the new knowledge. Many also provided coaches who would observe teachers and give them constructive feedback about how they could alter their practices to more closely conform to the stipulated practices. Generally speaking, these programs presented what they believed to be the best methods for teaching specific content, or grade levels, and then tried to make sure all teachers learned to implement those methods as accurately as possible.

In the knowledge development programs, PD providers spent more time helping teachers examine their own teaching. Sometimes they organized conversations about video recordings of classroom lessons. Sometimes they asked teachers to try using the same learning activities that their students would use later so that they could experience the learning activities first hand. Sometimes they asked teachers to jointly develop lesson plans. In all cases, the PD leader introduced new knowledge and new concepts into these discussions and encouraged teachers to use those concepts in their analyses and interpretation of classroom events.

In a sense, the differences between these two sets of programs are analogous to the distinction made by Stigler, Hiebert and Givven (this volume) between changes that are transmitted into schools from the outside, versus those that are developed within the system itself as part of ongoing self-improvement processes. In this case, the PD programs aiming for knowledge transmission brought their ideas into the schools and expected teachers to adopt them. But the analogy is not entirely accurate because, in this case, the developmental PD providers also typically came from outside. Still, their aim was to improve local capacity for self-improvement rather than offer predefined best practices for local teachers to implement.

To shed light on the merits of these different models of influence, I present several matched pairs of programs. Programs within each pair had similar strategic goals, but they differed in their underlying models for how to influence teaching

practice. The outcomes of these pairs of programs are displayed graphically in Figure 8.1, where each icon represents a finding from an independent study of a particular approach to PD, relative to its own comparison program. The vertical guide on the left measures program effects in standard deviation units and in each case effectiveness is measured relative to an alternative program that lacked the salient feature of the focal PD. Outcome measures varied from study to study, of course, but in Figure 8.1, I show results either from a commercial standardized test or from the test that most closely approximates that.

I used these icons to represent program effects because the icons provide a lot of other information about the studies themselves. The size of each icon reflects the number of teachers participating in the PD (bigger icons mean more teachers participated in the PD) and the shading of each icon reflects the number of hours teachers spent in the program (a darker shade means more total contact hours). Thus, a large, dark circle tells us that a program included a lot of teachers and spent a lot of time with them.

Each icon, or icon cluster if needed, is also labeled with the name of its PD program, and these names are all followed by either an asterisk (*) indicating that the program followed a knowledge-development model, or with a small square (■) indicating that the program followed a knowledge-transmission model.

The first thing that Figure 8.1 tells us is that *all* studies of PD yielded relatively small effects on student learning. Virtually all program effects are below .3 of a standard deviation. This is consistent with the nature of PD: there is a lot

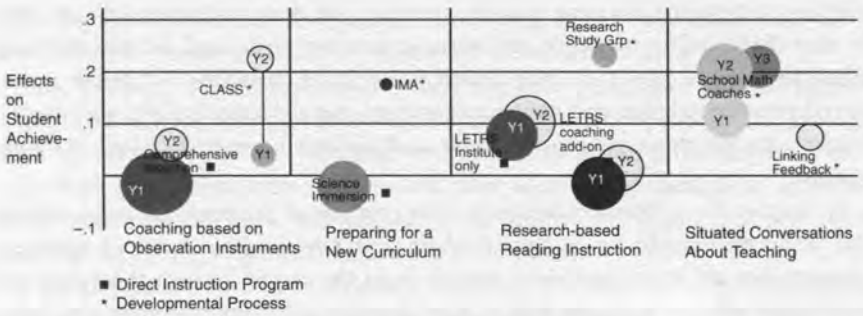

FIGURE 8.1 Effectiveness of PD based on knowledge transmission vs knowledge development

Legend: Each cluster of icons represents an independent study of a particular PD program relative to its own comparison program. Programs based on a knowledge-development model have an asterisk after their name and those based on a transmission model have a box after their name. Larger icons represent larger programs and darker icons represent more time spent with teachers. Comparison programs were whatever the districts already provided or a program explicitly designed to be comparable. Outcome measures were either commercial standardized tests or whatever most closely approximated those. The Science Immersion program outcome consisted of curriculum-embedded unit tests and the IMA outcome was a test of computational skills.

of slippage as new ideas move from one level to another, from the PD designer to the PD provider to teachers and finally to students. Nevertheless, there are meaningful differences among these programs, and in particular there are meaningful differences within each pair. So let's now look at these pairs.

Learning from Observation Instruments

The two icons in the left-most section of Figure 8.1 represent outcomes from two PD programs that used observation instruments as tools to convey their reform ideas to teachers. Both programs interacted with teachers for a full school year; both assessed student achievement with a standardized test at the end of that year; and both assessed student achievement again at the end of the following year, even though their programs did not continue into the second year. Thus, each program is represented with two icons: a shaded icon that reflects the intensity of treatment activity during the first year and a white icon that represent the lack of programmatic activity during the second year.

The larger icons represent outcomes from a program called *Comprehensive Induction* (Glazerman et al., 2008; Glazerman et al., 2010; Isenberg et al., 2009). I put a little square after its name to show that it relied on knowledge transmission to influence practice. The program was organized around an observation instrument that measured a specific set of recommended practices. Coaches used the observation rubric to define best practices for novices, to demonstrate these practices to novices, to evaluate novices' use of these practices and to give novices feedback. The observation rubrics enabled coaches to be very explicit about what good practice should *look like*, and what teachers should *do*, but did not help coaches explain how these behaviors worked, why they worked, or how they contributed to any particular teaching goal. These are the questions that knowledge development programs focus on.

In the second program (Allen, Pianta, Gregory, Mikami, & Lun, 2011), the observation instrument was used to stimulate teachers' thinking about what was happening during their lessons. Pianta (this volume) provides more information about this observation instrument and about the PD program itself. In this PD, coaches weren't even inside teachers' buildings but instead provided telephone consultations based on concepts embedded in the *Classroom Assessment Scoring System* (CLASS). Teachers videorecorded sample lessons approximately every two weeks and sent their video recordings to an online "teaching partner." Then the two would talk about the lesson. The teaching partners were far less directive than the Comprehensive Induction coaches. Instead of correcting teachers' behaviors, prescribing good practice, or evaluating what they saw on the video, they used "prompts" to help teachers examine and think about specific events that had occurred. For instance, a "nice work" prompt might say, "You do a nice job letting the students talk. It seems like they are really feeling involved. Why do you think this worked?" And a "consider this" prompt might look like this:

> One aspect of "Teacher Sensitivity" is when you consistently monitor students for cues and when you notice they need extra support or assistance. In this clip, what does the boy in the front row do that shows you that he needs your support at this moment? What criteria did you use to gauge when to move on?

Notice that the teaching partner was not directly recommending any specific procedures or rules for teachers to follow. Instead, the partner used the observation protocol to pose questions that might help teachers think harder about their classroom experiences, about the relationship between their own behaviors and the behaviors of their students, and about the enacted meaning of the concepts underlying the CLASS observation instrument. Rather than directly instructing teachers about what they *should have done*, coaches sought to help teachers think more strategically about their own teaching problems and how their own behavior influenced their students.

Figure 8.1 shows that both program outcomes were slightly above the zero baseline at the end of their treatment year, and that both were further above the baseline at the end of the second year, a finding that suggests that teachers do need time to internalize new ideas from PD and that, over time, they continue to find ways to incorporate those new ideas into their ongoing systems of practice. Neither of the year-1 effects was statistically significant, and only the Teaching Partners program was significant in year 2. Apparently, the questions raised by the teaching partners encourage teachers to gain a deeper understanding of the concepts embodied in the observation instrument, to use these concepts to examine about their own situations, goals, and strategies more closely, and to develop their own solutions to their ongoing teaching problems. This second-year contrast is especially important because neither program was actively working with teachers during the second year, so the differences tell us the extent to which teachers themselves, independently, were able to use the program's lessons to improve their own practices. And apparently, the knowledge-development PD fostered more self-generated improvements than did the knowledge-transmission PD.

Learning a New Curriculum

The second comparison in Figure 8.1 contrasts two PD programs that were designed to help teachers learn a new curriculum. One program introduced teachers in Los Angeles to a new science curriculum called *Science Immersion* (Borman, Gamoran, & Bowdon, 2008). The other introduced teachers to a new mathematics curriculum called *Integrated Mathematics Assessment* (Saxe, Gearhart, & Nasir, 2001).

The new science curriculum was intended to "immerse" students in the scientific process by asking them to generate hypotheses and conduct experiments. Although the curriculum appeared to reflect a knowledge-development model of influence on children's learning, with its emphasis on exploration and hypothesis testing, the PD

for teachers was based on a model of knowledge transmission. To help teachers learn this new curriculum, the PD gave them detailed manuals and associated guidelines regarding how to implement each unit of the immersion curriculum. Comparison teachers were given the new curriculum but not the PD.

To get a sense of the prescriptive detail in the Science Immersion PD, take a look at this passage from the teachers' manual describing a single fourth-grade unit (*Rot it right: The cycling of matter and the replication of energy. 4th grade science immersion unit*, 2006, p. 21):

1. To set the tone for this investigation as an exploration, generate a class discussion and class list about what plants need for growth and development.
2. Use the Think Aloud technique to model how to refine a wondering into a good scientific investigation. From the students' list about what plants need, form the question – What effect does sunlight have on radish plant growth and development?
3. Continue the Think Aloud to model assembling the Terraqua Columns using proper experimental procedures, and designing an experiment that has only one factor that is varied.
4. Have students record and explain their predictions for each set of columns for later reference.

Each lesson description begins with an enumeration of the steps that need to occur. The manual for this unit alone is 206 pages long.

The second PD program used a knowledge-development model to introduce teachers to the new mathematics curriculum. Teachers met in a small study group throughout the year as they implemented the new curriculum. Within each curricular unit, the PD meetings addressed four issues: the mathematical concepts themselves, how students reasoned about those concepts, what motivates children to participate in mathematics, and how to assess students' understanding in the context of class discussions, through their written work, or in other ways. Notice how closely these four topics reflect the four central functions of teaching.

The second section of Figure 8.1 shows the outcomes of these two programs and shows that the second PD, based on a knowledge-development model, was far more effective than the PD based on knowledge transmission. But there is another important finding from the second study that is not visually apparent. The knowledge-development PD included two different comparison groups, one of which was given the same new curriculum with no guidance (a comparison analogous to the comparison used in the Science Immersion study); the other relied on professional learning communities – but without the aid of an external guide – to examine the new curriculum and think about how to implement it. The icon shown in Figure 8.1 reflects the average of these two comparisons. It is important to note that the focal PD was more effective than the professional

learning community. This tells us that effective knowledge development consists of more than simply prompting teachers to think about what they are doing. In this case, teacher learning communities alone were less productive than was the community with a discussion moderator, who presumably introduced new concepts and raised provocative questions to push teachers' thinking.

Notice, too, that this approach to knowledge development was quite different from the approach in the Teaching Partner program, even though I consider both to be based on a "developmental" process. Though their pedagogies are quite different, they reflect the same underlying assumptions. Both are oriented toward developing teachers' capability to make their own decisions in the classroom and both also introduced new *concepts* (rather than prescriptions) that they wanted teachers to use as they designed their own lessons.

Readers may have also noticed that studies of the two knowledge-development programs described so far were very small relative to their contrasting studies of knowledge-transmission programs. Small studies are more commonplace among reformers using the knowledge-development model, and in fact, this study of the mathematics PD was the smallest study in the entire population reviewed, with only nine teachers in the treatment group and six in the comparison group. I suspect one reason for these smaller research projects may be that knowledge-development processes are more heavily reliant on the expertise of the PD providers themselves. Programs based on knowledge transmission often directly train their PD providers using the same logic that they use to train teachers themselves, so larger-scale projects are more feasible within that model of change.

Learning to Use Research-based Literacy Practices

The third contrast shown in Figure 8.1 compares two studies of PD that aimed to help teachers learn research-based strategies for teaching language arts. Both programs focused on practices identified several years ago when a National Reading Panel (2000) report provided an exhaustive summary of teaching practices whose effectiveness had been documented. So these programs were teaching essentially the same content, and this content could be considered to be very authoritative, warranted by research, and universal, thus inviting a knowledge-transmission approach to PD. Further, both programs wanted teachers to learn the specific teaching practices that had been found to be effective in promoting specific kinds of student learning. Yet the programs used quite different methods for teaching this content to teachers.

The first program, *LETRS, or Language Essentials for Teachers of Reading and Spelling* (Garet et al., 2008), relied on knowledge transmission. Teachers attended a series of eight full-day seminars interspersed throughout the school year, each covering one of the practices identified as "essential" by the National Reading Panel. Each seminar was accompanied by a complete textbook on that topic. A subset of participating teachers also received a parallel coaching program in which

coaches provided additional support by demonstrating these practices, observing teachers, and giving them feedback. Figure 8.1 shows two icons for the LETRS programs, one for the institutes alone and another for the combination of institute and coaching.

Even though the LETRS icons are displayed as slightly above the zero baseline, neither of these effects was statistically significant. However, the program did have a significant effect on teachers' *knowledge* and on their *implementation* of recommended procedures. In other words, teachers acquired the knowledge they were expected to acquire and they implemented the procedures they were expected to implement. Yet neither of these changes increased student achievement. For reformers, this is a puzzling outcome. Teachers gained both knowledge and technique, but were not able to use their new knowledge or techniques effectively to raise student achievement. But the finding should not be surprising at all, for knowledge transmission tells teachers only what their practices should *look like* rather than what their functions are within each lesson. The findings tell us that knowledge transmission is effective in the sense that teachers can learn to do the specific things they are told to do, but it is not effective in helping teachers learn to *teach*. I will return to this important point later.

The second program in this pair was called *Teacher Study Group* (Gersten, Dimino, Jayanthi, Kim, & Santoro, 2010). In this program, teachers met in study groups to examine the same National Reading Panel findings and then consider how they might design their forthcoming lessons in light of these findings. Meetings followed a four-step process: First, teachers would report what happened when they implemented their previously planned lessons. Then they would discuss their latest readings about research-based reading instruction. In this phase, the group facilitator focused their attention on the central concepts to make sure everyone understood them. In the third phase, they would review the publisher's recommended lesson and discuss its strengths and weaknesses. Finally, they would work together to design a lesson of their own that incorporated the research principle they had just read about.

The difference between the LETRS program and the Teacher Study Group program is especially useful because of their common content, of course, but also because that content is quite specific about what constitutes effective teaching. The content would seem to invite knowledge transmission. It is clearly specified and has the kind of universal authority that can often be granted by research. Yet the two programs differ in their underlying assumptions about the nature of teaching and of teacher learning. In one program, teachers were told everything they should do, and were observed to ensure that they did it all as specified. In the other, they evaluated the merits, rationale, and purposes of various practices, and decided for themselves when and whether to use them. Remember, too, that knowledge transmission was successful at increasing teachers' knowledge and their implementation of the recommended practices, so the problem is not that knowledge transmission per se failed to convey the content.

I can think of two hypotheses for why this approach to knowledge transmission failed to improve teaching practice. One has to do with the role these practices serve in normal classrooms. Most of the practices that reformers advocate today have been identified through correlational studies of the sort described by Gitomer (this volume). These studies look for relationships between (1) specific observable teaching behaviors and (2) overall student learning. These studies tell us that, when teachers in the past have enacted a particular practice, their students learned more. Missing from this message, though, is *why* the original teachers did those things in the first place. For the original teachers, the observed behaviors were part of a broader strategy for achieving their instructional ends – perhaps enlisting student engagement or monitoring students' understandings. Further, teachers may have made further adjustments to their lessons based on what happened after they did the originally observed behavior. The researchers do not know what these purposes were, so when new teachers were taught to use these procedures, the procedures were separated from their original instructional purpose. Knowledge transmission describes the form, but not the function of the recommended moves. It reminds us of a problem first identified over a century ago when Gore (1903) pointed out the problem of teaching "methods" separately from content in teacher education. He summarized the problem by saying "Content without form is blind, and form without content is empty" (p. 642). In this case, teachers learned to duplicate specific forms identified by reading researchers but the forms were emptied of their meaning when they were presented as ends in themselves.

The second hypothesis is that, by its very nature, knowledge transmission has the potential to *interfere with practice*, both by ignoring the role of context in making decisions about what to do next, and by distracting teachers' attention from their broader instructional purposes. By asking teachers to copy a set of movements that are independent of context and independent of any instructional purpose, knowledge transmission diverts teachers' attention away from the very thing they should be thinking about: context and purpose. Teachers begin designing their lessons to duplicate the recommended practices, but not to achieve any particular instructional goals. They focus on getting the *protocol* right rather than getting the *lesson* right.

Learning from Situated Examinations of Practice

The fourth section of Figure 8.1 differs from the others in that both programs in this comparison appear to be aiming for knowledge development. Both programs encourage teachers to work together and to share ideas and experiences, presumably to become better at reasoning about their own practice and developing better strategies for meeting their goals. But they differ substantially in their effectiveness, inviting us to speculate about why.

The larger and more effective of these two program (Campbell & Malkus, 2011) relied on coaches who were prepared for their job by taking several credit hours of

mathematics and mathematics education. Coaches did not rely on any standardized protocols to define effective practice but instead used their own judgment, presumably informed by their improved understanding of mathematics content, to guide their work. Coaches worked with all teachers in a school, engaging in co-teaching and co-planning, and they also attended school-level team meetings. Notice that their work with teachers involved *co-planning*, which is very different from coaching programs that *observe and evaluate teaching practice*. Co-planning focuses on strategic thinking, whereas evaluation focuses on errors in past practice. Discussions about lesson *plans* invite discussions about goals, about the relation between means and ends, about possible conceptual confusions in today's content, and about the unique needs of specific students. Notice, too, that the study itself was larger than has been the case for other knowledge-development programs, thus suggesting that it is possible to manage larger-scale PD within the knowledge-development model.

The second program (Supovitz, 2012) sought to foster data-driven decisions among teachers. In this case, the districts had already formed local professional learning communities which met at regular intervals to review and discuss each new wave of achievement test data. For their study, the authors augmented the unit test data with observation data in the hope that the combination of observation data and achievement data might stimulate greater thought among teachers and hence greater changes in their practice. The program had a modest impact, but not as great as its comparison PD.

I present this last pair-wise comparison to remind readers that the knowledge-development process is a relatively new approach to PD and that it is not as clearly defined as knowledge transmission is. The general approach is to focus on fostering strategic thinking about means and ends, and to provide teachers with new concepts that can guide their strategic thinking. But beyond this general outline, these programs have been quite various. Perhaps future studies might be modeled after Heller and others (2012), who compared alternative PD approaches *within the developmental paradigm*. An accumulation of studies like this will help us gain greater understanding of how this paradigm works. In the meantime, how can we use the findings above to further clarify the practice of knowledge-development PD?

Implications for Knowledge Development

The series of pair-wise comparisons shown above suggests that efforts to foster knowledge development are generally more effective than efforts to foster knowledge transmission. Based on those studies, I now offer some hypotheses about the salient features of effective PD.

Effective PD focuses on underlying goals rather than visible behaviors. My first hypothesis is that PD programs are more effective when they help teachers plan their lessons rather than when they critique those lessons. This is especially apparent in the two programs introducing effective literacy practices.

Teachers were more able to adopt research-based language-arts practices when they spent their time thinking about how to incorporate these practices into their lessons rather than when they were told to use these techniques and then were given performance feedback on how well they did. Similarly, in the comparison of curriculum-oriented PD, teachers were better able to implement their new curriculum when they spent time thinking about the content in relation to their own students rather than following step-by-step guidelines about how to implement the new curriculum.

Teaching is a purposeful activity and I suspect that many failures within the knowledge-transmission model are due to its lack of attention to purpose. The difference between a teaching behavior and its purpose was revealed in a small study by Ghousseini (2015), who was attempting to teach "core practices" to a novice teacher called Linda. One recommended practice was "managing discussions," which consisted of a set of discrete moves such as asking students what other students had just said, or asking students how many agreed or disagreed with something another student said. But as Ghousseini watched Linda use these moves, she found that Linda used them to achieve different *purposes* than Ghousseini had intended. For instance, when a student offered an incorrect idea, Linda was more likely to ask the class how many agreed or disagreed with the students' idea. Conversely, when a student offered a good idea, Linda was more likely to ask another student to repeat what the first student had said. This bias in her use of these moves suggests that, although Linda had learned to enact the moves, she didn't have a complete understanding of their purposes. For Linda, the goal of the discussion was to get students to say correct sentences, whereas for Ghousseini, the goal was to foster and monitor student thinking. Without clarity regarding purpose, Linda had learned to use the moves for her own purposes. This sort of shift may be common in programs that focus on knowledge transmission.

Effective PD content is conceptual rather than factual or procedural. The comparisons above also suggest that conceptual content is more valuable than procedural or factual content. Concepts are more abstract, allowing teachers to particularize them in different ways to accommodate the particulars of their own situations. This difference is especially apparent in the two PD programs that were based on observation instruments. In the Comprehensive Induction program, the observation instrument was used to define discrete visible practices that teachers should use. In the Teaching Partner program, the observation instrument was used to introduce concepts such as "teacher sensitivity," recognizing that there are many ways in which teacher sensitivity might be enacted within a given lesson.

Change takes time. The third thing we learn from these studies is that teachers need time to incorporate new knowledge into their ongoing systems of practices. Regardless of the nature of the original program, when researchers followed teachers beyond the conclusion of the PD itself, the teachers continued to

become more effective over that time period. This should not be a surprise. Teaching is a complex practice in which teachers are necessarily striving to do many things at once – manage specific behavior problems, motivate students, represent complex substantive ideas, conform to a rigid schedule with a limited set of materials, and so forth. In their review of teachers' thought processes, Clark and Peterson (1986) noted that teachers make decisions as often as every two minutes, and those decisions are necessarily based on their interpretation of events at that specific moment. It makes sense that teachers would need time to work with new ideas and it makes sense that, over time, they would get better at incorporating those new ideas into their ongoing systems of practice.

This important aspect of teacher learning is often ignored in the flurry of reforms. As reforms proliferate, teachers are more likely to be learning something new every year, and thus no reform is fully digested or incorporated into their practice before another reform is introduced.

Effective PD pedagogy encourages teacher-to-teacher sharing. The notion of teacher-to-teacher sharing has recently become popular, and for good reason: There is increasing evidence that student achievement improves when teachers are given shared planning time (Ladd, 2009), when they engage in collaborative activities (Ronfeldt, Farmer, Mcqueen, & Grissom, 2015), or when their school provides a supportive professional environment (Kraft & Papay, 2014). In one study (Papay, Taylor, Tyler, & Laski, 2016), teachers who had fostered larger student achievement gains were paired with teachers whose students achieved less, and they were able to help their colleagues improve their practices. These findings reinforce sociocultural learning theories that were introduced in the 1990s (Brown, Collins, & Duguid, 1989) and entered educational discourse at the turn of the century (see, e.g., Putnam & Borko, 2000).

However, as a result of findings such as these, some reformers are rushing to increase teacher *agency* (Calvert, 2016) or teacher *autonomy* (e.g., Education Evolving, 2014), without considering whether teachers actually have very many useful concepts to share with one another. This leads to my last point:

Effective PD also introduces new ideas. Not only did the programs reviewed here rely on outside change agents to introduce new ideas, but some also used autonomous professional learning communities as their *comparison group*. In these cases, the main contribution of the focal program was that it provided an outsider whose job it was to introduce new ideas into local conversations. That person might be called a coach, a teaching partner, or a group facilitator, but his or her job was not merely to engage teachers in analytic discussions about teaching, but also to introduce new concepts into those discussions. And in each case where researchers compared their PD learning communities to autonomous learning communities, the PD had a greater impact on student learning than local learning communities alone. So while collegial discussion and self-exploration are important learning *processes*, real improvements come when these conversations include someone who can introduce new ideas into conversations.

Prospects for Improving Teaching Practice

So far, I have mostly focused on the kind of professional development that reformers provide. Specifically, I have shown that programs that aim for knowledge development are more effective than those that aim for knowledge transmission. This is an important finding, particularly given the dominance of the knowledge-transmission model among education reformers.

But I want to add a second point, having more to do with the entire reform enterprise: The number and variety of reforms we impose on teachers each year has itself become dysfunctional. Our education system is awash in reforms of all shapes and sizes, to the point where teachers could not possibly accommodate them all, let alone understand them well enough to incorporate them meaningfully into their systems of instruction. Talk of reform dominates our schools, our school districts, and our national education institutions. Districts adopt one observation protocol after another, conveying to teachers the idea that their task is to implement pre-specified sets of activities rather than to help students think about and learn particular school content.

Further, most of these reforms are conveyed to teachers through a knowledge-transmission model, an approach that can actually distract teachers from thinking about the four central functions of teaching or about their lesson-specific problems and goals. So the number and variety of reforms introduced each year, combined with the fact that they are introduced through a knowledge-transmission process, is likely to have the net effect of reducing, rather than enhancing, the quality of teachers' practice. Every new policy, from zero tolerance to team teaching, from curricular reform to new student assessments to new observation systems, asks teachers to *stop* thinking about how they will achieve the four central functions of teaching and to think instead about how to enact the latest reform.

One could argue, of course, that these distractions are temporary and are justified by their long-term value, but because the urge to change practices in our schools has become so ubiquitous, teachers are learning something new virtually every year, and often more than one new thing each year. They do not have time to fully incorporate any of these reforms into a coherent and strategic approach to teaching before the next innovation is introduced. We are stuck in a *reform habit* that continuously distracts teachers from thoughtfully engaging with their lessons and leads them to believe that teaching consists of a hastily put-together collection of ad hoc practices rather than a strategically designed sequence of events.

How, then, can we shift our education system toward a more developmental stance? Put another way, how can we institutionalize knowledge-development habits? Stigler, Hiebert, and Givvin (this volume) view this problem as partly an insider/outsider problem. People outside of schools design reforms and then try to push them into schools. In a parallel universe, they suggest, we might imagine institutions that are self-correcting and self-improving, so that they don't need external interventions. The self-improvement process that these authors favor is

called a PDSA cycle (Plan, Do, Study, and Act), something like the lesson study format used in Japan. If such practices became the norm, rather than ending up as just another temporary fad, they believe, schools could evolve into self-improving institutions, places where members collaborate to examine their own systems and find ways to fine-tune them.

One wonders whether such institutions could ever develop in this country. Our educational governance system is built on the assumption that good ideas must be generated by researchers and empirically verified outside the local context. Further, our school culture, our "grammar of schooling," gives teachers far too many non-teaching duties and far too little time to engage in thoughtful examinations of their own practices. I once proposed lesson study to an elementary teacher who was widely recognized for her expertise, but she immediately dismissed the idea as not feasible within her work schedule.

Even apart from a specific approach such as lesson study, we can still wonder whether schools could ever sustain the kind of policy stability needed to promote ongoing self-improvement. Cobb and Jackson (2011) tried to figure out what local school systems would need to maintain a self-sustaining support system for teachers. They concluded that such a system would not be possible in the U.S. school system, for it would require too much institutional stability: explicit goals, an assessment system that was aligned with those goals, a detailed vision of high-quality instruction, a set of materials and instructional tools compatible with that vision, district-wide professional development for specific practices, and school-based professional learning communities to provide ongoing support, and additional supports for struggling students. Such a system would also depend on willing and able local administrators and the political stability of the school board and the state and federal educational agencies.

At this stage in U.S. education, it is hard to imagine any self-sustaining, systemic, or continuous self-development process ever becoming institutionalized, in part because our widespread belief that those of us outside the system have better ideas than teachers appear to have, because of our widespread desire for long-distance control of teaching practice, and because of our widespread faith in knowledge transmission as the appropriate mechanism for change. Every year, a multitude of reformers across our land seek to import new reforms into schools and very few seem willing to abandon their goals in favor of a self-improvement system. Virtually every politician feels obligated to include some sort of education reform in his or her platform, and virtually every school superintendent feels s/he must introduce something new that puts his own personal stamp on his or her district. Reform has become a professional obligation for school administrators as well as politicians. And in a nation of not-really-very-United States, each year brings in a new list of urgent reform proposals, each of which seeks to be duplicated throughout the education landscape. We have trapped ourselves in a system of continuous change that gives teachers no stability and no time to develop coherent, purposeful practices.

Note

1. In most studies of PD, the most salient competing hypothesis for explaining PD outcomes is teachers' motivation to learn. That is, if teachers taking the PD *chose* to enroll in the program, and comparison teachers did not, then the two groups are not comparable in their motivation to learn. All studies included here controlled for teachers' motivation to learn in some way. In some cases, all teachers in a district were randomly assigned (thus, perhaps none were motivated to learn, but the two groups were equal on that). In other cases, teachers volunteered and then were randomly assigned (here they are equal in that they knew they would be randomly assigned when they enlisted). In still other cases, teachers were offered two programs and could select one of them, which still ensured comparable motivation to learn.

References

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science, 333*, 1034–1037.

Borman, G. D., Gamoran, A., & Bowdon, J. (2008). A randomized trial of knowledge development in elementary science: First-year achievement effects. *Journal of Research on Educational Effectiveness, 1*, 237–264.

Brown, J. S., Collins, A. M., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42.

Calvert, L. (2016). *Moving from compliance to agency: What teachers need to make professional learning work*. Oxford, OH: Learning forward and NCTAF.

Campbell, P. F., & Malkus, N. N. (2011). The impact of elementary mathematics coaches on student achievement. *Elementary School Journal, 111*(3), 430–454.

Charters, W. W., & Waples, D. (1929). *Commonwealth teacher training study*. Chicago, IL: University of Chicago Press.

Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 255–296). New York, NY: Macmillan.

Cobb, P., & Jackson, K. (2011). Towards an empirically grounded theory of action for improving the quality of mathematics teaching at scale. *Mathematics Teacher Education and Development, 13*(1), 6–33.

Cohen, D. K. (1988). Teaching practice: Plus ça change… In P. W. Jackson (Ed.), *Contributing to educational change: Perspectives on research and practice* (pp. 27–84). Berkeley, CA: McCutchan.

Cuban, L. (1984). *How teachers taught: Constancy and change in American classrooms, 1890–1980*. White Plains, NY: Longman.

Cuban, L. (1990). Reforming again, again, and again. *Educational Researcher, 19* (1), 3–13.

Education Evolving. (2014). *Teacher powered schools: Generating lasting impact through common sense innovation*. Education Evolving. Retrieved from www.teacherpowered.org/

Garet, M. S., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., … Sztejnberg, L. (2008). *The impact of two professional development interventions on early reading instruction and achievement*. Washington, DC: National Center for Educational Evaluation and Regional Assistance, Institute of Education Sciences. Retrieved from https://ies.ed.gov/ncee/pdf/20084030.pdf

Gersten, R., Dimino, J., Jayanthi, M., Kim, J. S., & Santoro, L. E. (2010). Teacher study group: Impact of the professional development model on reading instruction and

student outcomes in first grade classrooms. *American Educational Research Journal, 47*(3), 694–739.

Ghousseini, H. (2015). Core practices and problems of practice in learning to lead class discussions. *Elementary School Journal, 115*(3), 334–357.

Gitomer, D. H. Promises and pitfalls for teacher evaluation. This volume.

Glazerman, S., Dolfin, S., Bleeker, M., Johnson, A., Isenberg, E., Lugo-Gil, J., … Britton, E. (2008). *Impacts of comprehensive teacher induction: Results from the first year of a randomized controlled study*. (NCEE 2009–4035). Washington, DC: Institute for Education Sciences. Retrieved from http://ies.ed.gov/ncee/pdf/20094035.pdf

Glazerman, S., Isenberg, E., Dolfin, S., Bleeker, M., Johnson, A., Grider, M., & Jacobus, M. (2010). *Impacts of comprehensive teacher induction: Final results from a randomized controlled study*. Washington, DC: National Center for Education Evaluation. Retrieved from http://ies.ed.gov/ncee/pdf/20094034.pdf

Gore, W. C. (1903). Academic and professional training of teachers. II. *Elementary School Teacher, 3*(9), 632–637.

Heller, J. I., Dahler, K. R., Wong, N., Shinohara, M., & Miratrix, L. W. (2012). Differential effects of three professional development models on teacher knowledge and student achievement in elementary science. *Journal of Research in Science Teaching, 49*(3), 333–362.

Isenberg, E., Glazerman, S., Bleeker, M., Johnson, A., Lugo-Gil, J., Grider, M., & Dolfin, S. (2009). *Impacts of comprehensive teacher induction: Results from the second year of a randomized controlled study*. Washington, DC: Institute for Education Sciences. Retrieved from http://ies.ed.gov/ncee/pubs/20094072/pdf/20094072.pdf

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrer, Straus and Giroux.

Kennedy, M. M. (2005). *Inside teaching: How classroom life undermines reform*. Cambridge, MA: Harvard University Press.

Kennedy, M. M. (2016). Parsing the practice of teaching. *Journal of Teacher Education, 67*(1), 3–17.

Kounin, J. (1970). *Discipline and group management in classrooms*. New York, NY: Holt, Rhinehart and Winston.

Kraft, M. A., & Papay, J. P. (2014). Can professional environments in schools promote knowledge development? Explaining heterogeneity in returns to teaching experience. *Educational Evaluation and Policy Analysis, 36* (4), 476–500.

Ladd, H. (2009). Teachers' perception of their working conditions: How predictive of planned and actual teacher movement? *Educational Evaluation and Policy Analysis, 33*, 235–261.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. Washington, DC: National Institute of Child Health and Human Development. Retrieved from www.nichd.nih.gov/publications/pubs/nrp/Documents/report.pdf

Papay, J. P., Taylor, E. S., Tyler, J. H., & Laski, M. (2016). *Learning job skills from colleagues at work: Evidence from a field experiment using teacher performance data*. Cambridge, MA: National Bureau of Economic Research.

Pianta, R. C. Teacher-student interactions: Measurement, impacts, improvement, and policy. This volume.

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher, 29*(1), 4–15.

Ronfeldt, M., Farmer, S. O., Mcqueen, K., & Grissom, J. A. (2015). Teacher collaboration in instructional teams and student achievement. *American Educational Research Journal, 52*(3), 475–514.

Rot it right: The cycling of matter and the replication of energy. 4th grade science immersion unit. (2006). Los Angeles, CA: SCALE: System-Wide Change for All Learners and Educators. Retrieved from notebook.lausd.net/pls/ptl/url/ITEM/2BD0A2671713C038E0430A000210C038

Saxe, G. B., Gearhart, M., & Nasir, N. S. (2001). Enhancing students' understanding of mathematics: A study of three contrasting approaches to professional support. *Journal of Mathematics Teacher Education, 4*, 55–79.

Stigler, J. W., Hiebert, J., & Givvin, K. B. Does VAM + MET = Improved teaching? This volume.

Supovitz, J. (2012). *The linking study—First-year results: A report of the first-year effects of an experimental study of the impact of feedback to teachers on teaching and learning*: University of Pennsylvania Center for Policy Research in Education. Retrieved from www.cpre.org/linking-study-first-year-results

Tyack, D. (1987). An American tradition: The changing role of schooling and teaching. *Harvard Educational Review, 57*(2), 171–174.

Tyack, D., & Cuban, L. (1995). *Tinkering toward Utopia: A century of public school reform.* Cambridge, MA: Harvard University Press.

Tyack, D., & Tobin, W. (1994). The "grammar" of schooling: Why has it been so hard to change? *American Educational Research Journal, 31*(3), 453–479.