Chapter 9

Content *Then* Process: Teacher Learning Communities in the Service of Formative Assessment

Dylan Wiliam

Raising student achievement is important, but not for the reasons many educators think. Forget No Child Left Behind and adequate yearly progress. Forget district and state reports that rank schools by proportion of proficient students. Raising achievement is important because it matters for individuals and society. If you achieve at a higher level, you live longer, are healthier, and earn more money. For those with only a high school diploma, the standard of living in the United States is lower today than it was in 1975; for those with degrees, it is 25 to 50% higher.

In addition, people who earn more money pay more taxes, are less likely to depend on Medicaid or welfare, and are less likely to be in prison. It has been calculated that if a student who drops out of high school would stay to graduate, the benefit to society would be $209,000 (Levin, Belfield, Muennig, & Rouse, 2007). This sum is

made up of $139,000 in extra tax revenue, $40,500 savings in public health costs, $26,600 savings in law-enforcement and prison costs, and $3,000 in welfare savings. Eric Hanushek (2004), a leading economist of education in the United States, has calculated that if we could raise each student's achievement by one standard deviation (equivalent to raising a student from the 50th to the 84th percentile), over 30 years, the economy would grow by an additional 10%, and just the *additional* taxes being paid by everyone would more than pay for the whole of K–12 education.

For these reasons, it is clear that we need to raise student achievement in our schools, but how do we do this? We too often rely on quick fixes, which rarely produce success. To successfully raise student achievement, we must improve the quality of the teachers working in our schools—specifically, we must work to improve the teachers we already have. We must look carefully at both the costs and the benefits of possible reforms. An analysis of the research reveals that helping teachers develop minute-by-minute and day-by-day formative assessment practices is more cost-effective than any other strategy. However, changing what teachers do, day in and day out, cannot be done effectively through traditional methods, such as the summer workshop. This chapter argues that to raise student achievement, educators must form building-based learning communities in which teachers use a format of five strategies for formative assessment, hold each other accountable, and provide support for one another.

The Search for Solutions

Given the gravity of the benefits of increased student achievement, both for individuals and society, it is not surprising that we have been looking for solutions to improve the effectiveness of our schools. The problem is that we have been looking in the wrong places. The first generation of school effectiveness research just looked at outputs: Some schools achieved good results, and others achieved less good results. The rather simplistic conclusion drawn

from this was that schools made a difference, so the search commenced for features of effective schools (see, for example, Chubb & Moe, 1990). However, a second generation of school effectiveness research showed that demographic factors accounted for most of the differences between schools. Most of the schools getting good results were in affluent areas, and most of the schools with low student achievement were in areas of poverty. The conclusion based on this research was that maybe schools did not make so much of a difference, and the reasons for low student achievement were demographic (see, for example, Thrupp, 1999).

More recently, a third generation of school effectiveness research has looked not only at the outputs of schools, but at the difference in what students knew when they started at the school compared to when they left—the so-called value added. What this research shows is that it does not matter very much which school students attend. What matters very much is which classrooms they are in in that school. If a student is in one of the most effective classrooms, he or she will learn in 6 months what those in an average classroom will take a year to learn. And if a student is in one of the least effective classrooms in that school, the same amount of learning will take 2 years. Students in the most effective classrooms learn at four times the speed of those in the least effective classrooms (Hanushek, 2004).

What accounts for these very different rates of learning? One obvious factor is class size, but it turns out that the effects of class-size reduction programs on student achievement are quite small, and such programs are very expensive (Hattie, 2005). Jepsen and Rivkin (2002) found that, for example, teaching 120 third-grade students in classes of 20 rather than 30 would result in just five additional students passing a standardized test, at a cost of around $120,000. In general, class-size reduction programs are most effective for younger students (kindergarten, first, and possibly second grade), and then

only if class size is reduced to 13 to 15 students (Mosteller, 1995). For most students, the effect of class-size reduction is small.

Other reform efforts have centered on the structures of our schools, such as size or different governance and funding models. Much time, effort, and money has been invested in the creation of smaller schools, and while in some cases, improvements in student engagement and attitudes have been found, there appears to be little or no impact on student achievement (Gewertz, 2006a; 2006b). The charter school movement has claimed some successes, but once the demographic factors are taken into account, the impact on student achievement seems to be small, if not negligible (Carnoy, Jacobsen, Mishel, & Rothstein, 2005; Lubienski & Lubienski, 2006).

What about teacher subject knowledge? In a study of almost 3,000 students in 115 classrooms, Hill, Rowan, and Ball (2005) found that higher levels of what they called "mathematical knowledge for teaching" *were* associated with increased student progress, but the effects were, again, small. Over a year, the students taught by the most knowledgeable teachers (in other words, the top 5%) learned about 25% faster than the students taught by the least knowledgeable teachers (those in the bottom 5%). This difference was statistically significant, and bigger than the impact of socioeconomic factors, but nowhere near the 400% speed-of-learning differential between students in the most and least effective classrooms.

It appears that the most important difference between the most and the least effective classrooms is the teacher, but the most important variable appears to be what they *do*, rather than what they *know* (Monk, 1994). If we want better teachers, we can achieve this in a number of ways: If we do it gradually, by increasing the thresholds at which teachers enter the profession, it takes many years to have any effect (Hanushek, 2004). If we try to do it quickly, for example by allowing people to teach even though they have not completed a teacher-preparation program, then these teachers turn out to be no

better than those trained in traditional ways (Darling-Hammond, Holtzman, Gatlin, & Vasquez Heilig, 2005).

To sum up the argument so far: We need to raise student achievement, because it matters for individuals and for society. To raise student achievement, we need to improve teacher quality, and the only way to do this, at least in the short to medium term, is to invest in the teachers we have already—what my colleague Marnie Thompson calls the "love the one you're with" strategy. In other words, if we are serious about improving student achievement, we have to invest in the right professional development for teachers.

This is an important point, because too often, professional development is presented as a fringe benefit—part of a compensation package to make teachers feel better about their jobs. This certainly seems to be the way teacher professional development is viewed by many outside the world of education, and also, sometimes, by policymakers.

The question is this: Can we do it? Can we improve teacher quality through professional development? Asked 20 or even 10 years ago, the answer would probably have been, "No." While investment in teacher professional development has been a feature of the educational landscape for many years, there was depressingly little evidence that it made any difference to student achievement: "Nothing has promised so much and has been so frustratingly wasteful as the thousands of workshops and conferences that led to no significant change in practice when teachers returned to their classrooms" (Fullan, 1991, p. 315). However, in recent years, it has become clear that the main reason teacher professional development has largely failed to impact student achievement is because we have not been doing what the research shows makes a difference to student learning. So now the question is this: What do we need to be doing to improve teacher quality and thus improve student achievement, and how do we go about doing it?

In recent years, there has been a growing acceptance that to be successful, teacher professional development needs to concentrate on both content *and* process (Reeves, McCall, & MacGilchrist, 2001; Wilson & Berne, 1999). In other words, we need to focus on *what* we want teachers to change, or change about what they do, and we have to understand *how* to support teachers in making these changes. However, in practice, this often gets implemented through unfocused models. Teacher coaching is one example. This strategy can be used to change things about the way teachers teach that do affect student learning, but it can also be used to change things that do not make any difference to student learning. For instance, a teacher might want to learn how to implement "jigsaw" groups, despite the absence of evidence that this is likely to make a difference to student achievement (indeed, according to Slavin, Hurley, & Chamberlain, 2003, not only is there an absence of evidence of an effect; there is, in fact, evidence of the absence of an effect). Similarly, there has been marked interest among educators in such strategies as "brain-based education" or learning styles despite the absence of evidence that these models make a difference to student achievement (see, respectively, Bruer, 1999; Adey, Fairbrother, & Wiliam, 1999).

If we are serious about improving student achievement, we have to focus relentlessly on changing those things that teachers do that are the most important to change: those things that, when we focus on them, and change them, improve student learning.

The remainder of this chapter shows that, measured in terms of impact on student achievement, the single most important thing to change in teachers' practice is the minute-to-minute and day-by-day use of assessment to adjust instruction. Because this involves sustained change in deeply ingrained practices and habits, this will necessitate a different type of teacher professional development: building-based teacher learning communities. But note that the teacher learning communities are a means to an end, not an end in themselves: content, *then* process.

Formative Assessment: The Evidence

The term "formative assessment" has been with us for 40 years (see Wiliam, 2007, for an extended account of the origins of the term), but the underlying idea—that we should use evidence of learning to adjust instruction—has been around for thousands of years. Reviews of research in this area by Natriello (1987) and Crooks (1988) were updated by Black and Wiliam (1998), who concluded that regular use of classroom formative assessment would raise student achievement by 0.4 to 0.7 standard deviations—enough to raise the United States into the top five countries in the international rankings for math achievement, for example. Subsequent longer-term implementation studies (those at least 1 year in duration), with tests that are less sensitive to instruction than those typically used in research studies, have found smaller effect sizes—typically around 0.3 standard deviations (Wiliam, Lee, Harrison, & Black, 2004)—but even these are large effects. To see how large, it is useful to compare the effects of teachers' use of formative assessment with other kinds of educational interventions.

Table 1 (page 190) shows the effect, in the number of additional months progress per year, of three different educational interventions, and the cost per classroom per year. The estimate of the effects of class size is based on the data generated by Jepsen and Rivkin (2002), and the estimate for teacher content knowledge is derived from Hill, Rowan, and Ball (2005). (The cost of increasing teacher content knowledge is unknown. We know it increases student achievement, but we do not know how to raise teachers' pedagogical content knowledge by this much.) The estimate for formative assessment is derived from Wiliam, Harrison, & Black (2004), and other small-scale studies. The data in Table 1 suggest that investing in teacher professional development is 20 to 30 times more cost-effective than class-size reduction, at least beyond the second grade.

Intervention	Extra Months of Learning Gained per Year	Classroom Cost per Year
Class-size reduction by 30% (for example, from 30 to 20 students)	3	$30,000
Increase teacher content knowledge from weak to strong (2 standard deviations)	1.5	Unknown
Formative assessment	6 to 9	$3,000

Table 1: Cost-Effect Comparisons for Three Educational Interventions

The substantial cost-effectiveness of formative assessment as a lever for school improvement has, of course, attracted considerable attention, and a number of test publishers have produced what they call "formative assessment systems." Typically, these systems provide for assessment of student progress at regular intervals (generally every 4 to 9 weeks) and provide reports that identify students or particular aspects of the curriculum that require special attention.

These systems *are* formative, in the sense that they provide evidence about student achievement that can be used to adapt instruction to better meet student learning needs. However, they are very different from the research reviewed by Natriello (1987), Crooks (1988), and Black and Wiliam (1998), which found that regular use of classroom assessment increases student achievement. Such formative assessment systems have a role to play in the effective monitoring of student progress. Indeed, I would argue that some means of tracking student progress over the medium term and taking action to address any problems identified are essential components of any comprehensive assessment system. But it is disingenuous at least, and possibly mendacious, to claim that the research literature provides evidence of the effectiveness of such systems. Quite simply, it does not (Popham, 2006; Shepard, 2007). That is not to say that

such evidence will not be forthcoming in the future—it may well be—but no such evidence has been assembled to date. So if we are serious about a relentless focus on the things that we know will work to raise student achievement, these formative assessment systems have only a marginal role to play. They will not by themselves result in substantial increases in student achievement, nor are they even necessary. Such systems have a role in supporting good management and supervision, but in terms of improving—as opposed to monitoring—student learning, they are almost irrelevant.

What is needed, rather, is a focus on what actually goes on inside the classroom (Black & Wiliam, 1998). The kinds of formative assessment practices that profoundly impact student achievement cannot wait until the end of a marking period, or even to the end of an instructional unit. If students have left the classroom before teachers have made adjustments to their teaching on the basis of what they have learned about the students' achievement, then they are already playing catch-up. If teachers do not make adjustments before students come back the next day, it is probably too late. This is why the most important formative assessments are those that occur minute-by-minute and day-by-day (Leahy, Lyon, Thompson, & Wiliam, 2005).

So the big idea of formative assessment is that evidence about student learning is used to adjust instruction to better meet student needs; in other words, teaching is *adaptive* to the student's learning needs, and assessment is done in real time. More explicitly (Thompson & Wiliam, 2007, p. 6), formative assessment is:

Students and teachers
Using evidence of learning
To adapt teaching and learning
To meet immediate learning needs
Minute-to-minute and day-by-day

An important feature of this definition is that it shares the responsibility for formative assessment between teachers and students. While this definition succeeds in explaining what effective formative assessment is not, it provides few clues about how effective assessment should be done. A careful analysis of the theoretical and empirical work in this area (Wiliam & Leahy, 2007; Wiliam & Thompson, 2007) suggests that effective formative assessment consists of five key strategies:

1. Clarifying learning intentions and sharing criteria for success
2. Engineering effective classroom discussions, questions, and learning tasks that elicit evidence of learning
3. Providing feedback that moves learners forward
4. Activating students as the owners of their own learning
5. Activating students as instructional resources for one another

(A comprehensive review of the research underlying this analysis can be found in Wiliam, 2007.)

These strategies clarify what effective assessment is, but they still provide relatively little guidance for the teacher about how to apply these strategies in the classroom. For that reason, I draw a careful distinction between *strategies* and *techniques*. These five strategies are "no-brainers"—they are always smart things to do in the classroom; however, the techniques used to implement these strategies require careful thought by the teacher. What might work for one class in one context might not be appropriate for another class—regardless of how similar they appear to be. Through work with teachers and researchers in dozens of schools, my colleagues and I have developed a list of more than 100 of these techniques across the five strategies. Five of these techniques follow, one for each of the five strategies.

Strategy 1: Clarifying Learning Intentions and Sharing Criteria for Success

Technique: Sharing Exemplars

Before asking students to write a new kind of assignment such as a lab report, the teacher gives each student four sample lab reports that represent varying degrees of quality. These samples can be student work from earlier years—with the names removed, of course—or teacher-produced samples. Students are asked to place the pieces in order of quality and identify what is good about the good ones and what is missing or weak about those that are not as good.

Strategy 2: Engineering Effective Classroom Discussions, Questions, and Learning Tasks That Elicit Evidence of Learning

Technique: Dry-Erase Boards

During a lesson on equivalent fractions, the teacher asks the class a question such as, "Write down a fraction between ⅙ and ½," and asks all students to hold up their responses on the count of three. Using this kind of "all-student response system" helps the teacher to quickly get a sense of what students know or understand while requiring all students in the class to engage in the task. If all the answers are correct, the teacher moves on. If none of the answers are correct, the teacher may choose to reteach the concept in a different way. If there are a variety of answers, the teacher can then use the information gleaned from student responses to direct the subsequent discussion.

Strategy 3: Providing Feedback That Moves Learners Forward

Technique: Find It and Fix It

Rather than checking all correct answers in an exercise and putting a check mark next to those that are incorrect, the teacher directs the student to identify them him- or herself: "Five of these are incorrect; find them and fix them." This kind of feedback requires

the student to engage cognitively in responding to the feedback, rather that just reacting emotionally to his or her score or grade.

Strategy 4: Activating Students as the Owners of Their Own Learning

Technique: Traffic Lighting

After they complete a piece of work, students go back to the learning intention or success criteria provided at the beginning of the lesson and indicate their level of understanding with a colored circle: Green means "I understand," yellow means "I'm not sure," and red means "I do not understand." Younger students can use emoticons such as ☺, 😐, and ☹. At regular intervals, the teacher provides time in class for students to move their learning forward by turning their reds to yellow and their yellows to green.

Strategy 5: Activating Students as Instructional Resources for One Another

Technique: Pre-Flight Checklist

Before students can turn in an assignment, they must trade papers with a peer. Each student then completes a "pre-flight checklist" by comparing the peer's document against a list of required elements. For example, the pre-flight checklist for a lab report might require, among other things, a title, a date, diagrams drawn in pencil and labeled, and results that are clearly separated from conclusions. Only when the peer has signed off on the checklist can the work be turned in to the teacher. (Note that the peer is not involved in grading the work; only in providing feedback for improvement.)

By using these specific techniques for formative assessment (among others), teachers can begin to integrate the five strategies into their classroom practice. But this is only the beginning of the process. Through reflection, teachers begin to develop a new conceptualization of their practice. In some ways, this teacher professional

development process echoes the remarks of Millard Fuller, the founder of Habitat for Humanity, who said, "It's easier to act your way into a new way of thinking than to think your way into a new way of acting." The traditional, workshop-based approach to professional development has two drawbacks. First, it does nothing to involve teachers in the process and tailor their work to their own students in their own schools and classrooms. Second, as any quarterback can attest, knowing what you want to do, and actually being able to do it, are two very different things.

In this new model, teachers start with the practical components, such as the five sample techniques mentioned previously. As they adopt and integrate these techniques and others into their own practice, they find a new synergy and see their own practice in new ways, which in turn leads to new thinking. In other words, rather than trying to transfer a researcher's thinking straight to the teacher, this new approach to formative assessment emphasizes content, *then* process.

Embedding Formative Assessment With Learning Communities

If the use of minute-to-minute and day-by-day formative assessment brings such benefits, and what needs to be done is so clear, why is it not common practice in every American classroom? There are two main reasons: The first is that many of the practices identified in the research as necessary to effective implementation have to go up against long-established traditions. For example, the tradition of grading student work with letter grades is so ingrained in schools that many teachers cannot envision doing business any other way. However, there are ways of assessing that provide useful information to the teacher and can still be used to assign grades at the end of a marking period (see Clymer & Wiliam, 2006/2007). Too often, the prevailing attitude is that one should not countenance change because things are bad enough as they are, and effective leadership is very difficult to deliver when superintendents hold their position

on average for less than 2½ years (Snider, 2006). The second reason is that the changes are hard to implement. These are not superficial changes like learning to teach new units or adopting different lesson structures. The adoption of minute-to-minute and day-by-day formative assessment requires deep changes in the way that teachers teach, and this is much harder than it appears.

For example, a teacher with 20 years of experience will have asked approximately half a million questions in her career. When one has done something a certain way half a million times, doing it another way is very difficult. But there is a deeper reason why change is difficult, even for inexperienced teachers: Teachers learn most of what they know about teaching before they are 18 years old. In the same way that most of us learn what we know about parenting through being parented, teachers have internalized the "scripts" of school from when they themselves were students. Even the best 4-year teacher-education programs will find it hard to overcome the models of practice their future teachers learned in the 13 or 14 years they spent in school as students.

If we are to have any chance of really changing teacher practice, we have to take seriously that implementing minute-to-minute and day-by-day formative assessment is not primarily a matter of providing teachers with new knowledge, although some knowledge will be important. The crucial thing is to change habits, and traditional teaching structures do not change habits. The adoption of coaching in many districts is an example of this, but as noted previously, you can coach teachers to change aspects of practice that have little impact on student achievement. After many false starts and blind alleys, I have become convinced that the best way to support teachers in adopting minute-to-minute and day-by-day formative assessment is through building-based teacher learning communities. This is not because of an ideological commitment to the benefits of teachers talking to each other, but because of the nature of the

changes we are seeking to produce. If we were trying to increase teacher subject knowledge, then teacher learning communities would not be a very sensible approach—it would be far better to arrange for high-quality direct instruction. But when we are trying to change deeply ingrained, routinized practices, then it seems that teacher learning communities offer the best hope.

Over the last 3 years, my colleagues and I have explored a number of different approaches to establishing and sustaining teacher learning communities, and as a result of this experimentation, it appears that five principles are particularly important: gradualism, flexibility, choice, accountability, and support (Wiliam, 2006).

Gradualism

Asking teachers to change what they do is rather like asking a golfer to change his swing in the middle of a tournament. Teachers who try to add more than two or three techniques to their teaching at the same time almost invariably find that their teaching routines fall apart, and they go back to doing what they know how to do. In the long-term, they achieve less change than teachers who take smaller steps.

Flexibility

Techniques that work in one context may not work, or may not be appropriate, in others. Only the teacher is able to judge this, so he or she must be able to make adjustments to the techniques. Sticking within the framework of the five strategies reduces the chance that a teacher's modifications weaken the power of the technique. For example, one teacher used colored disks—green on one side and red on the other—to encourage students to do "real-time" traffic lighting. At the start of the lesson, all disks are green side up, but if a student wants to ask a question, he or she must turn the disk over to show red. Another teacher tried this, but found it hard to see the disks from the front of the classroom, so she purchased sets of red,

yellow, and green cups (enough for each student to get one of each). At the beginning of the lesson, the yellow and red cups are nested inside the green cup. If a student thinks the teacher is going too fast, he or she shows the yellow cup, and if the student wants to ask the teacher a question, he or she shows the red cup. This teacher then introduced a new variation: As soon as one student shows a red cup, the teacher chooses at random from the students showing yellow and green, and the selected student has to answer the question posed by the student with the red cup. This modification by the teacher not only made the "traffic lights" more visible, but it also increased students' accountability for their learning.

Choice

The initial reaction of most teachers to being asked to adopt minute-to-minute and day-by-day formative assessment techniques is that it is scary. Putting teachers in control of choosing which techniques they will try appears to make this challenge a little less daunting, and it also allows teachers to emphasize the techniques that best suit their teaching style. For example, my colleagues and I (Black, Harrison, Lee, Marshall, & Wiliam, 2003) describe two middle-school science teachers in the same school with very different styles. Derek is a charismatic teacher with a substantial classroom presence—a "larger-than-life character," so to speak. He chose to focus on student questioning as his priority for personal development. He is now an extraordinarily skilled leader of classroom discussions—one of the best I have ever seen—but he is very firmly in the middle of everything, much like the conductor of an orchestra. Another teacher, Philip, has a very different classroom style. When you walk into his classroom, he is often hard to spot, since he is usually in the midst of a discussion with a small group of students, while the others in the classroom work purposefully at their allotted tasks. Philip chose to focus on activating students as owners of their own learning, and as instructional resources for one another, and now creates a

highly effective learning environment for his students. The important point here is that Derek would be far less successful if he were forced to emphasize peer- and self-assessment, and Philip would be far less successful if he were forced to prioritize questioning. Teaching is a highly personal activity, and choice in implementing formative assessment is essential if teachers are to integrate it into their practice.

Accountability

Although in this process teachers are free to choose how to change or adapt their techniques, they remain accountable to the teacher learning community for the changes they have made. In other words, although they are free to choose what to change, they are accountable to the teacher learning community for the change, and, just as importantly, they have to be able to justify that what they change is likely to improve student learning by being clearly related to at least one of the five key strategies. Teachers have repeatedly told us that the fact that they had promised to their colleagues to try something out is what made them prioritize this over all the other things they had to do.

Support

The flip side of accountability is support. In fact, in our work with teachers, we have used the term "supportive accountability" to indicate the learning community's effectiveness. The learning community offers both support and accountability, but with two conditions: First, the teacher learning community builds trust among its members so that members can move beyond "polite serial turn-taking" and begin genuinely to engage in each other's professional development (Grossman, Wineburg, & Woolworth, 2000). Second, the teacher learning community is genuinely a meeting of equals, at least in terms of power. In our experience, when one member of the community sets him- or herself up as the formative assessment "expert," the learning of the other members is compromised. While there is a valuable role

for those who are not currently teaching—supporting the group, running interference, providing advocacy, and so forth— they can never be full participants in such a community.

Content *Then* Process

In this chapter, I have argued that increasing student achievement must be a priority, both for our students individually and for society as a whole. In the past, policymakers have focused on quick-fix solutions, such as curriculum reform, textbook replacement, changes in governance, altering school structures and timetables, and investing large amounts of money in information technology. Even when implemented properly (as few are), these kinds of initiatives have little impact on student achievement. They encourage us to come to the plate because the team is losing to try to hit a home run. But we end up striking out instead. What the research shows us is that the only answer is continuous small improvements—"small ball," if you like. We need to worry about getting to first base before we can make it home. We need to do this not because the solution is elegant or attractive, but because there is nothing else that we currently know of that works anything near as well.

The research shows that it is what teachers do in the classroom that really matters—not having teachers meet in workshops to talk about how to assess student work or what students' scores on tests mean for the curriculum. If the research on professional development over the last 20 years has shown us anything, it is that we can change teacher thinking without changing teacher practice, and the only thing that impacts student achievement is teacher practice. So if we are serious about raising student achievement, we must focus on helping teachers change what they do in the classroom.

Moreover, we must also be careful to focus on the things that make a difference. Teacher change for the sake of change is inadequate. The research summarized in this chapter shows that changing

teachers' minute-to-minute and day-by-day formative assessment practices is the most powerful way to increase student achievement, and it is 20 to 30 times more cost-effective than, for example, class-size reduction programs.

However, being clear about what needs to be done is only the first step, although it has to be the first step: content, *then* process. Once we are clear about what needs to be done, we must figure out ways to do it, and this is where the research on teacher expertise is so important. The necessary changes are *not* changes in teacher knowledge—teachers know much of what they need to know already. The changes we need are changes in the habits and rituals of teachers' practice that have been ingrained over many years. This requires different pedagogical models. Specifically, it requires building-based teacher learning communities. If we can concentrate on doing what is right, rather than what is expedient or easy, unprecedented increases in student achievement and teacher satisfaction are within our grasp. The question is whether we have the courage to reach.

References

Adey, P. S., Fairbrother, R. W., & Wiliam, D. (1999). *A review of research related to learning styles and strategies*. London: King's College London Centre for the Advancement of Thinking.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Buckingham, UK: Open University Press.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 86*(1), 8–21.

Black, P. J., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139–148.

Bruer, J. T. (1999). In search of … brain-based education. *Phi Delta Kappan, 80*(9), 648–657.

Carnoy, M., Jacobsen, R., Mishel, L., & Rothstein, R. (2005). *The charter school dust-up: Examining the evidence on enrollment and achievement*. Washington, DC: Economic Policy Institute.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438–481.

Chubb, J. E., & Moe, T. M. (1990). *Politics, markets and America's schools*. Washington, DC: The Brookings Institution.

Clymer, J. B., & Wiliam, D. (2006/2007). Improving the way we grade science. *Educational Leadership, 64*(4), 36–42.

Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Vasquez Heilig, J. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives, 13*(42).

Fullan, M. (1991). The new meaning of educational change. London: Cassell.

Gewertz, C. (2006a, March 8). Failed breakup in Denver offering lessons. *Education Week, 25*(27), pp. 1, 18.

Gewertz, C. (2006b, August 9). Chicago's small schools see gains, but not on tests. *Education Week, 25*(44), pp. 5, 18.

Grossman, P., Wineburg, S., & Woolworth, S. (2000). *What makes teacher communities different from a gathering of teachers?* Seattle: University of Washington Center for the Study of Teaching and Policy.

Hanushek, E. A. (2004). *Some simple analytics of school quality* (Vol. W10229). Washington, DC: National Bureau of Economic Research.

Hattie, J. (2005). The paradox of reducing class size and improving learning outcomes. *International Journal of Educational Research, 43*(6), 387–425.

Jepsen, C., & Rivkin, S. G. (2002). *What is the tradeoff between smaller classes and teacher quality?* (NBER Working Paper #9205). Cambridge, MA: National Bureau of Economic Research.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371–406.

Leahy, S., Lyon, C., Thompson, M., & Wiliam, D. (2005). Classroom assessment: Minute by minute, day by day. *Educational Leadership, 63*(3), 18–24.

Levin, H. M., Belfield, C., Muennig, P., & Rouse, C. (2007). *The costs and benefits of an excellent education for all of America's children*. New York: Teachers College Press.

Lubienski, C., & Lubienski, S. T. (2006). *Charter, private, public schools and academic achievement: New evidence from NAEP mathematics data 2006*. New York: National Center for the Study of Privatization in Education.

Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review, 13*(2), 125–145.

Mosteller, F. W. (1995). The Tennessee study of class size in the early school grades. *The Future of Children (Special issue: Critical issues for children and youths), 5*(2), 113–127.

Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist, 22*(2), 155–175.

Popham, W. J. (2006). Phony formative assessments: Buyer beware! *Educational Leadership, 64*(3), 86–87.

Reeves, J., McCall, J., & MacGilchrist, B. (2001). Change leadership: Planning, conceptualization and perception. In J. MacBeath & P. Mortimore (Eds.), *Improving school effectiveness* (pp. 122–137). Buckingham, UK: Open University Press.

Shepard, L. A. (2007). Will commercialism enable or destroy formative assessment? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Slavin, R. E., Hurley, E. A., & Chamberlain, A. M. (2003). Cooperative learning and achievement. In W. M. Reynolds & G. J. Miller (Eds.), *Handbook of psychology, volume 7: Educational psychology* (pp. 177–198). Hoboken, NJ: Wiley.

Snider, J. H. (2006, January 11). The superintendent as scapegoat. *Education Week, 25*(18), 40, 31.

Thompson, M., & Wiliam, D. (2007). *Tight but loose: A conceptual framework for scaling up school reforms.* Paper presented at a Symposium entitled "Tight but loose: Scaling up teacher professional development in diverse contexts" at the annual conference of the American Educational Research Association, Chicago, IL.

Thrupp, M. (1999). *Schools making a difference: Let's be realistic!* Buckingham, UK: Open University Press.

Wiliam, D. (2006). Assessment: Learning communities can use it to engineer a bridge connecting teaching and learning. *Journal of Staff Development, 27*(1), 16–20.

Wiliam, D. (2007). Keeping learning on track: Formative assessment and the regulation of learning. In F. K. Lester, Jr. (Ed.), *Second handbook of mathematics teaching and learning* (pp. 1053–1098). Greenwich, CT: Information Age Publishing.

Wiliam, D., & Leahy, S. (2007). A theoretical foundation for formative assessment. In J. H. McMillan (Ed.), *Formative classroom assessment: Research, theory and practice*. New York: Teachers College Press.

Wiliam, D., Lee, C., Harrison, C., & Black, P. J. (2004). Teachers developing assessment *for* learning: Impact on student achievement. *Assessment in Education: Principles, Policy, and Practice, 11*(1), 49–65.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Wilson, S. M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (Vol. 24, pp. 173–209). Washington, DC: American Educational Research Association.