

It's Time to Rethink Teacher Supervision and Evaluation

The process by which most teachers are supervised and evaluated is inefficient, ineffective, and a poor use of principals' time, argues Mr. Marshall. It needs to be drastically streamlined and linked to a broader strategy for improving teaching and learning.

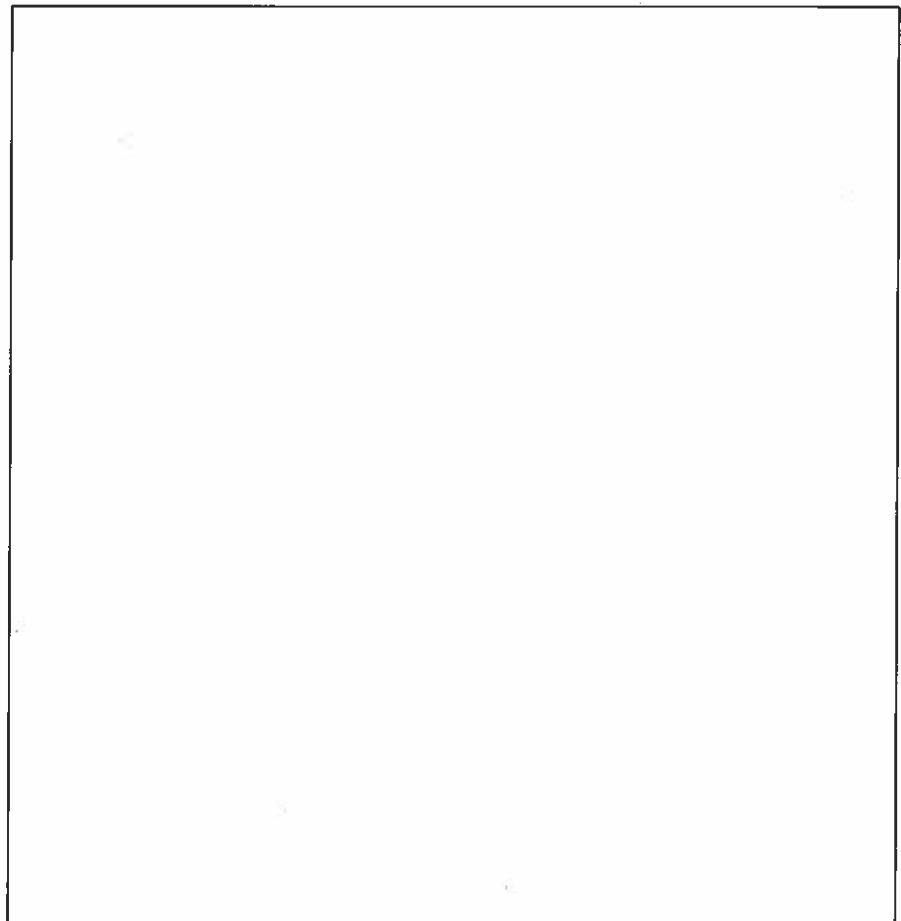
BY KIM MARSHALL

A PRINCIPAL boasts that he spends two hours a day in classrooms. And it's true — he really does visit his school's 17 teachers daily, chatting with students and occasionally chiming in on a lesson. But when teachers are asked what kind of feedback they get, they say the principal rarely talks to them about what he sees when he strolls through their classes.

* * *

A principal gets complaints from several parents about a history teacher's problems with discipline but is so overwhelmed that she rarely visits his classroom. When she does her required observation of his class, she sees a carefully planned lesson featuring an elaborate PowerPoint presentation and well-

KIM MARSHALL was a teacher, a central office administrator, and a principal in the Boston Public Schools for 32 years. He now works as a leadership coach for new principals in New York City and Newark, N.J.; teaches courses on instructional leadership; and publishes The Marshall Memo, a weekly digest of ideas and research on K-12 education (www.marshallmemo.com). He wishes to thank Roland Barth, John King, Sandy Kleinman, Jay McTighe, Douglas Reeves, Jon Saphier, Mike Schmoker, Rhoda Schneider, and Athie Tschibelu for their important contributions to this article.



behaved students. The principal feels she has no choice but to do a positive write-up of this lesson and give the teacher a satisfactory rating.

* * *

A principal spends four entire weekends in April and May completing teacher evaluations just before the deadline. He puts the evaluations into teachers' mailboxes with a cover note attached that reads, "Please let me know if you have any concerns and would like to talk. Otherwise, sign and return by tomorrow." All the teachers sign, nobody requests a meeting, and there is no further discussion.

* * *

A well-regarded veteran teacher hasn't been evaluated in five years and rarely sees the principal in her classroom. She takes this as a compliment — her teaching must be "okay." And yet she feels lonely and isolated with her students and wishes the principal would pay an occasional visit and tell her what he thinks.

* * *

A sixth-grade teacher has good classroom management and is well liked by students and parents, but his students do poorly on standardized tests. A new principal mentions the disappointing scores, and the teacher launches into a litany of complaints: he always gets the "bad class," most of his students come from dysfunctional families, and he's tired of being asked to "teach to the test." Later that day, the union representative officiously tells the principal that she can't mention test results in a teacher's evaluation.

* * *

A principal observes an elaborate hands-on math lesson in a veteran teacher's classroom and notices that the teacher is confusing the terms *mean*, *median*, and *mode*. The principal notes this error in his mostly positive evaluation, and, in the post-observation conference, the teacher suddenly begins to cry. Ten years later, at his retirement party, the principal asks the teacher what lesson she took away from this incident. "Never to take a risk," she replies.

* * *

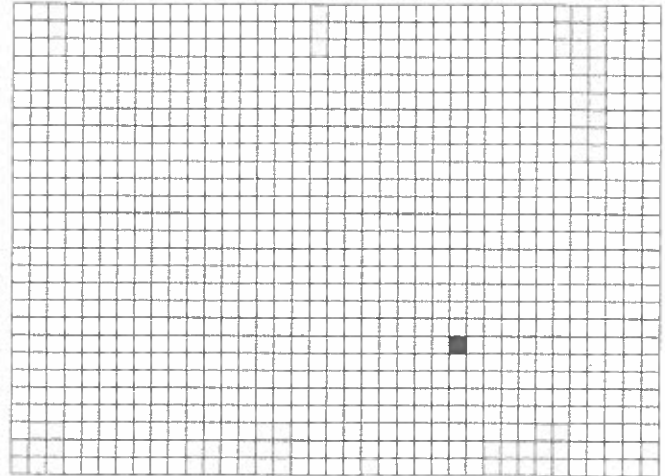
The theory of action behind supervision and evaluation is that they will improve teachers' effectiveness and therefore boost student achievement.¹ This assumption seems logical. But the vignettes above raise a troubling question: what if the theory is wrong? This article takes a close look at this possibility and explores an alternative theory of action.

WHY DO SUPERVISION AND EVALUATION OFTEN MISS THE MARK?

I believe there are 10 reasons why the conventional supervision and evaluation process is not an effective strate-

gy for improving teaching and learning.

1. *Principals evaluate only a tiny amount of teaching.* If a teacher has five classes a day, that's 900 periods each school year. A principal who formally evaluates a teacher for one full class period a year (a fairly typical scenario) sees this proportion of the teacher's lessons:



In this case, the principal evaluated 0.1% of the teacher's instruction. The other 99.9% of the time, the teacher was working with students unobserved. Even if the principal made three full-class evaluation visits a year, as required by some districts, that would still leave the teacher alone with students 99.7% of the time. No matter how observant and well trained the principal is, no matter how comprehensive the evaluation criteria are, and no matter how detailed the feedback is afterwards, this is ridiculously thin supervision of the school's most important employees. Principals who spend this little evaluative time in classrooms are basically bluffing, hoping that teachers will think they know more than they really do. Without expensive increases in administrative staffing — politically impossible in most districts — the amount of time principals spend formally observing each teacher is not going to change. Let's face it: teachers are on their own most of the time, and our schools depend heavily on their competence and professionalism.

2. *Microevaluations of individual lessons don't carry much weight.* Many school districts try to compensate for how little time principals spend in individual classrooms by requiring extremely thorough evaluations of lessons that are formally observed. Administrators are asked to script everything the teacher says and write a detailed account of exactly what happened in the class. A perceptive and well-trained principal can see a lot in a single lesson and give the teacher copious feedback on classroom management, student engagement, "accountable talk," clarity, momentum, wait time, bulletin boards, and so forth. But these elaborate write-ups don't mean a lot to most teachers; they know how little the principal sees

of their day-to-day struggles, curriculum planning, grading, work with colleagues, parent outreach, professional growth, and routine duties. Even if the evaluation is complimentary, it usually gets filed in a nanosecond. Except in extreme (and quite rare) cases when a principal gives an unsatisfactory rating, evaluation is a pro forma process that has very little influence on what teachers do on a daily basis.

3. *The lessons that principals evaluate are often atypical.* The only way that microevaluating lessons can give an accurate picture of a teacher's overall classroom performance is if the observed lessons are truly representative. But this is often not the case. When teachers have advance notice of an evaluation, they can present a glamorized lesson for the principal's benefit. Even if they don't, the presence of a top-level authority figure in the classroom usually reduces discipline problems and results in a more orderly lesson than students generally experience. These two factors can work in teachers' favor, giving the principal an unrealistically positive view of their teaching. You'd think that principals would be wise to these dynamics, but they are often so stressed and overwhelmed that they play along, treating clearly atypical teaching as typical. When this happens, teachers get an unfortunate message: it's okay to do "special" teaching when the principal visits and "ordinary" teaching for students the rest of the time.

Evaluation visits can also distort reality in a negative way: some teachers get so nervous when the principal arrives that they go to pieces. This is every teacher's nightmare — one screwed-up lesson and the other 99.9% of the year will be painted with the same evaluative brush.

Surely the principal has other sources of information to correct egregiously off-target observations, including informal visits, quick impressions of teachers interacting with students, parent comments, colleagues' impressions, and gossip. But these time-honored sources of information, even when accurate, aren't "admissible" in official evaluations. Principals have little choice but to go by the book and use the information from formal evaluation visits, even when it's bogus.

4. *Isolated lessons give an incomplete picture of instruction.* Although the lesson is the fundamental building block of teaching, it's only a small part of a teacher's effort to in-

spire students and convey knowledge and skills. To grasp the bigger picture, a principal needs to know more: What curriculum unit is this lesson part of? What are the unit's "big ideas" and "essential questions"? How does this unit align with state standards? How will students be assessed? Principals may try to ferret out these missing pieces by asking for lesson plans and conducting pre- and post-evaluation conferences with the teacher, but evaluations are still tied to the lesson that was observed.

This is a shame, because it's impossible to teach most state standards in a single lesson; it's a huge leap from big-picture goals like "understanding number sense" to planning a single lesson. *Unit* plans, which describe a teacher's game plan for teaching skills and concepts over a three- to five-week period, tell far more about whether instruction is coherent and aligned. But principals rarely see unit plans or the assessments that teachers give at the end of their units.

5. *Evaluation almost never focuses on student learning.* In virtually all school districts, teacher unions have been successful in preventing their members from being evaluated on whether students actually learn what's being taught. Unions are right to object to accountability on norm-referenced tests, since these assessments are not designed to be "instructionally sensitive."³ Before-and-after, "value-added" assessments

are better, but even their most fervent advocates don't think it's fair to use them to evaluate a teacher after only a year of instruction.

Does this mean that principals have no way of evaluating teachers on whether students are learning? Surely a principal can get a sense of how much students are picking up by walking around classrooms, looking over their shoulders, and asking them probing questions. But this approach has three problems. First, many principals are required to produce detailed narratives after each evaluation visit and can't walk around and write furiously at the same time. Second, even if principals manage to check in with a few students during classroom visits, it's hard to tell whether the whole class understands the lesson that day — let alone a few weeks later. To really know if teachers have been successful, principals need to see students' scores on good unit assessments — which they almost never do. Third, even if principals can get their hands on interim assessment results, such evidence is not admissible in evaluations.

So principals have little choice but to focus on teaching performances versus learning results, on chalkboard razzle-dazzle versus deep understanding, on beautiful bulletin boards versus demonstrated proficiency. Constrained by the supervision/evaluation process, principals overmanage the occasional lesson and undermanage the bigger picture of whether teachers are truly making a difference in student learning.

6. *High-stakes evaluation tends to shut down adult learning.* Even though many teachers don't respect the evaluation process, it still makes them nervous. Their collective bargaining agreements may provide good protection, but teachers harbor irrational fears that every time the principal walks into their classrooms, clipboard in hand, their jobs are on the line. Formal evaluations raise the level of tension and anxiety and make it more difficult to admit errors, listen, and talk openly about areas that need improvement. Any time evaluative comments are put in writing, the parties involved tighten up: the principal is less likely to tell the whole story for fear of facing a grievance, and the teacher is less likely to talk about how things are really going. In all too many evaluative interactions, teachers put on their game face and get through the process with as little authentic interaction as possible. The principal owns the feedback, not them.

This kind of process destroys a golden opportunity for professional growth. The real challenge of supervision and evaluation is to activate (or amplify) a supervisory voice inside teachers' heads that will guide them in their work with students. Conventional supervision and evaluation seldom accomplish this goal. In fact, the exact opposite may occur, with teachers waiting nervously for their principal to judge

them and putting up a wall of resistance to any criticism. Where do teachers go for helpful feedback on their teaching? Usually they turn to a colleague, a spouse, a family member, students, parents — or nobody.

An unintended consequence of this whole dynamic is the growth of a certain emptiness in the professional relationship between teachers and school leaders. If principals are rarely in classrooms, it's hard to have meaningful professional conversations with teachers. And if principals aren't setting the tone, it's less likely that assistant principals, team leaders, department heads, and colleagues will have serious conversations about teaching and learning. This kind of instructional vacuum can result in faculty lounge conversation dominated by topics outside of the school, gossip, and funny — and not-so-funny — stories about kids.³

7. *Supervision and evaluation reinforce teacher isolation.* One of the American principal's toughest challenges is counteracting two tendencies prevalent in our schools: teachers not working with their colleagues and the "educator's egocentric fallacy" — I taught it, therefore they learned it.⁴ In far too many schools, teachers who teach the same subjects at the same grade level don't work together, missing out on the synergy of collaboration and wasting precious time reinventing the wheel. Because principals evaluate teachers in private meetings and confidential documents, evaluation reinforces this isolation and is rarely a vehicle for getting teachers to talk to one another, which detracts from teachers' sense of responsibility to their grade-level or department team.

Evaluation is also an ineffective tool for countering our natural tendency to assume that if something is taught (i.e., explained or demonstrated), it is automatically learned.⁵ Because the supervision and evaluation process doesn't focus on team curriculum planning, assessment, and student learning, it doesn't prod teachers to emerge from their isolation and reflect with their colleagues on what they need to change in order for more students to succeed. Without this impetus, teachers gravitate toward the default setting: self-contained, activity-centered lessons or marching through the textbook.

8. *Evaluation instruments often get in the way.* Good teaching is extremely complex and challenging, and research tells us there is more than one way to get students to learn. It takes experience and savvy for a principal to grasp the subtleties of a classroom; it's even more demanding for a principal to capture them in writing; and it's *really* challenging to criticize a teacher's performance in a way that is heard. Some principals are good at all three — observation, write-ups, and "difficult conversations." Unfortunately, many principals are not, and the training needed to bring them up to

speed is woefully lacking. The legendary klutziness of school administrators has motivated unions to work overtime to negotiate “principal-proof” evaluation formats and procedures to protect their members from unfair evaluations. Districts, on the other hand, push for evaluation tools that make it possible to build a case to dismiss incompetent teachers. The resultant evaluation tools are rarely conducive to fostering an honest, open, and pedagogically sophisticated dialogue between principals and teachers.

9. *Evaluations often fail to give teachers “judgmental” feedback.* This seems like an odd statement, since all evaluations judge teachers. But many evaluation instruments allow principals to fudge teachers’ general status with an overall “satisfactory” rating and a lot of verbiage. These evaluations don’t tell teachers where they stand on clearly articulated performance standards, don’t give clear direction on the ways in which teachers can improve their performance, and don’t answer the question teachers really care about (and often dread): *How am I doing?* This kind of evaluation is unlikely to motivate a mediocre teacher to improve — or spur a good teacher on to excellence.

10. *Most principals are too busy to do a good job on supervision and evaluation.* Discipline and operational duties are so insistently demanding that teacher evaluation often disappears from principals’ calendars until contractual deadlines force them to get serious.⁶ When evaluation crunch time arrives, principals fall into three types — saints, cynics, and sinners. The saints go by the book, and evaluation consumes their lives for weeks at a time. I know a principal who routinely spends eight to 10 hours on each teacher evaluation: pre-observation conference, lesson observation, write-up (like a little term paper every Saturday, she says), and post-observation conference. Principals who choose to commit this amount of time (or are required to do so by their superiors) have no alternative but to shut themselves in their offices for days at a time — or spend evenings, weekends, and vacations at their desks at home. Ironically, this reduces the amount of time the saints spend in classrooms doing low-key supervision — coaching, encouraging, and gentle correction.

The second type of principal heaves a sigh, sits down at the computer, and bangs out the required evaluations as quickly as possible. Administrators in this category have grown cynical about the evaluation process and don’t believe their write-ups will produce better teaching and learning, but they feel they have no choice but to do them.

The third, more daring, group of principals *simply don’t do evaluations* (or evaluate only the occasional egregiously ineffective teacher). These sinners ignore contractual requirements and dare the system to catch them. Since evalu-

ation is in the same category as a trip to the dentist for many teachers, they tend not to complain if their principal “forgets” year after year. And principals’ superiors are often none the wiser — or choose to wink at these omissions.

So here’s the question: are the saints, who spend hours on each evaluation, more effective at improving teaching and learning in their schools than the cynics and the sinners? Shocking as it may seem, the answer in many cases is no. This is because the conventional supervision and evaluation process is not the best way to truly change what happens in classrooms. Principals need a better way to observe, support, and judge teachers — a way that is more accurate and time efficient and more closely linked to an effective strategy for improving teaching and learning.

LINKING SUPERVISION AND EVALUATION TO HIGH STUDENT ACHIEVEMENT

I’ve argued that the theory of action behind supervision and evaluation is flawed and that the conventional process rarely changes what teachers do in their classrooms. Here is an alternative theory: *The engine that drives high student achievement is teacher teams working collaboratively toward common curriculum expectations and using interim assessments to continuously improve teaching and attend to students who are not successful.* Richard DuFour, Mike Schmoker, Robert Marzano, Douglas Reeves, Jeffrey Howard, Grant Wiggins, Jay McTighe, and others believe that this approach is a critical element in high achievement. I agree, but with a proviso: if a school adopts this theory, it must change the way teachers are supervised and evaluated. If it doesn’t, the principal won’t have the time, energy, and insight to get the engine started and monitor it during each school year.

Why are the principal’s time and focus so crucial? Because teacher collaboration is countercultural in most American schools and rarely happens without impetus and support from outside the classroom. Principals are in the best position to provide the support, and rigorous state standards and high-stakes tests can provide the impetus. Standards and tests present a common challenge (a common enemy, some would say) that makes it easier for principals to get teacher teams to buy into working toward ambitious, measurable learning for students.

Of course principals still need to evaluate teachers every year or two, as required by most states, and they also need to give honest and timely feedback to ineffective teachers and have the guts to fire them if they don’t improve. But the essence of what I’m recommending is a shift away from a process owned by the principal, in which most of the en-

ergy goes into evaluating individual lessons, to a more dynamic, informal process owned by teacher teams. To make this happen, we need to shift:

- from periodically evaluating teaching to continuously analyzing learning;
- from inspecting teachers one by one to energizing the work of teacher teams;
- from evaluating individual lessons to supervising curriculum units;
- from occasional announced classroom visits to frequent unannounced visits;
- from detailed scripting of single lessons to quick sampling of multiple lessons;
- from faking it with distorted data to conducting authentic conversations based on real data;
- from year-end judgments to continuous suggestions and redirection;
- from comprehensive, written evaluations to focused, face-to-face feedback;
- from guarded, inauthentic conversations to candid give-and-take;
- from teachers saying, “Let me do it my way” to everyone asking, “Is it working?”;
- from employing rigid evaluation criteria to continuously looking at new ideas and practices;
- from focusing mainly on bad teachers to improving teaching in every classroom;
- from cumbersome, time-consuming evaluations to streamlined rubrics; and
- from being mired in paperwork to orchestrating school-wide improvement.

TWELVE STEPS TO LINKING SUPERVISION AND EVALUATION TO HIGH SCHOOL ACHIEVEMENT

These shifts will not happen by themselves. To recover from ineffective practices and to address widening achievement gaps, principals might try the following 12-step program.

1. *Make sure the basics are in place.* These include time scheduled for teacher teams (grade-level teams in elementary schools and subject-area teams in secondary schools) to meet on a weekly basis, preferably in uninterrupted 90-minute blocks; crystal-clear, end-of-year expectations for learning that are aligned with state standards; common assessments, which can be written by teacher teams or purchased, to measure learning and diagnose needs at the end of each year and at intervals during the year; common rubrics for consistently scoring student writing and open-ended responses; and exemplars of student work at the advanced,

proficient, basic, and below-basic levels.

2. *Decide on the irreducible elements of good teaching.* For principals and teachers to communicate well about what’s happening in classrooms, there must be a common language regarding the basics of effective teaching. Most evaluation checklists are way too long to remember. A handy acronym for the five elements that every classroom should have is SOTEL: *safety* — students feel physically and psychologically protected; *objectives* — the goals of the curriculum unit are evident; *teaching* — learning experiences are skillfully orchestrated; *engagement* — students are leaning forward, involved in the learning process; and *learning* — there is evidence, either during the lesson or on follow-up assessments, that students have learned what was taught.

3. *Systematically visit all classrooms on a regular basis.* Principals need to be in classrooms frequently for a reality check on how things are going. But how frequent is “frequently,” and how much time does a principal need to be in a classroom to see how things are going? The answers to these two questions are crucial because there’s a direct relationship between the length of each visit, the number of classrooms a principal can see each day, and the quality of information that is gathered. Shorter visits mean the principal can cover more classrooms, but visits that are too short yield superficial data.

Most principals make four types of classroom visits: 1) very brief, “showing the flag” appearances; 2) “walkthroughs” lasting a few minutes, with particular attention to student work on bulletin boards; 3) five- to 15-minute mini-observations focused intently on teaching and learning; and 4) full-period, formal observations with detailed note taking. All four types of visits are useful, but as I have argued previously,⁷ the third type is optimal for teachers whose basic competence is not in question. Mini-observations allow the principal to fit as many as five substantive visits into a busy day, and, if the visits are unannounced and the principal is focused and perceptive, they yield the most accurate data on how well teachers are performing.

A principal who is self-disciplined about making three to five mini-observations a day can get into all the classrooms in a medium-sized school every two weeks, systematically sampling the quality of teaching in chunks of time that can be fitted into a busy day. Using this approach, a principal can take 12 to 15 “snapshots” of every teacher’s performance in the course of the year and compile a “photo album” of each one’s overall performance. The total time the principal spends in each teacher’s classroom is not much longer than that spent in the conventional evaluation model described earlier, but the accuracy of the information gained is far superior.

There's an additional bonus for peripatetic principals: they get to know students better and pick up information that can be useful in understanding learning problems, resolving discipline situations, and talking with parents.⁸ Frequent classroom visits also convey an unmistakable message to teachers: "You never know when I'll drop in, and I expect good teaching to be going on whenever I do." If the principal sees something of concern (for example, a student being publicly humiliated), it's time to shift gears to a formal reprimand or a traditional full-lesson evaluation.

4. *Give teachers prompt, face-to-face feedback after every classroom visit.* Teachers should not be left in the dark about what the principal thinks, and personal feedback is far preferable to sending e-mails or leaving notes in teachers' mailboxes. In an informal, low-threat, private conversation, teachers are more likely to relax and engage in honest give-and-take about how things are going. These conversations go best when the principal's feedback focuses on one or two specific points — e.g., an appreciative comment about the way the teacher drew a shy student into the discussion, or a critical comment about the fact that the hands-on activities weren't focused on the unit objectives. Follow-up talks are most effective when they happen within 24 hours: "Better 120 seconds of feedback the same day than a five-page essay delivered a month later," says Douglas Reeves.⁹

In each of these follow-up conversations, principals should make a point of asking about student learning: "How is the Egypt unit coming?" "What Fountas-Pinnell levels have your lowest reading groups reached?" "How did the algebra test go?" If a principal has established a trusting climate, a teacher should be able to say, "My team just spent two weeks teaching the concept of borrowing, and the kids bombed on our quiz. Can you help us figure out what happened?" Teachers should know that their boss is keenly interested in results and should be comfortable reaching out for support.

5. *Require teacher teams to develop common unit plans and assessments.* The best way to ensure that teaching is done right the first time (versus having to provide corrective instruction for substantial numbers of students after the fact) is to have teachers work in teams to plan each curriculum unit with the end in sight.¹⁰ Before they dive into teaching, teacher teams should work backwards from the state standards to identify clear learning objectives, decide on the big ideas and essential questions of the unit, draft assessments they will use to determine whether students have learned what was taught, create a game plan and calendar for instruction, and run the plan by the principal for feedback.

The three- to six-week curriculum unit is an ideal chunk

of instruction for principals to supervise — far better than an individual lesson. A principal who has reviewed a unit plan can check out alignment in classrooms, look at how kids are responding, suggest midcourse corrections, and ask about student learning. Are examining unit plans and following up with teachers time-consuming? Yes. Are these activities a better use of a principal's time than lesson write-ups that are ignored by teachers? Absolutely!

By far the hardest part of implementing this approach is getting teachers to plan together in the first place. Teachers in the U.S. are accustomed to autonomy, and it takes a tenacious principal to foster this kind of collaboration. It's essential, though, because teams plan better than teachers working solo, and teams generate stronger ideas, provide better support, and increase the likelihood that the supervisory voice will be in each teacher's head as the unit unfolds.

6. *Require teams to give common interim assessments.* If formative assessments are of high quality — not just clones of multiple-choice end-of-the-year tests — they can give teachers valuable insights into what students are learning and not learning.¹¹ It's vital for teams to meet after each unit or quarterly assessment to look at the results and collectively answer these three questions: What percentage of students scored at the advanced, proficient, basic, or below-basic levels? In which areas did students do best, and where were they confused and unsuccessful? What is our strategy for addressing the weakest areas and helping students who are struggling?¹² A powerful enhancement to interim assessments is for teams to set SMART goals — Specific, Measurable, Achievable, Results-oriented, and Time-

bound (for example, 85% of our first-graders will be reading at Level I on the Fountas-Pinnell scale by June) — at the beginning of the year and to track progress each quarter.

Teacher ownership of this process is vital; it's better for a teacher to chair team meetings, ideally on a rotating basis, even if the principal has the time. Teachers need to have a clear mission for their meetings (experimentation, continuous improvement, and results), contractual time to score assessments, common planning time during or after the school day to analyze and discuss data, an outside facilitator (unless there is unusually strong leadership within the team), and occasional drop-in visits by the principal to give support and contribute ideas.

7. *Have teams report on student learning after each unit or quarter.* Lots of schools suffer from data overload and insufficient analysis and follow-through. The principal can help teams crystallize their thinking by asking for a *brief*, informal report on the three questions above and on one additional question: How can I help? It's crucial that these reports, which can be submitted either in person or in writing, are low-stakes, nonthreatening, and nonbureaucratic. Teams shouldn't be bogged down in paperwork and must feel they can be creative, try new things, admit mistakes, and engage in an informal give-and-take about what's working and what needs to be improved.

To summarize, let's contrast how a principal evaluates a teacher using the conventional model with the process that would be followed under the proposed model:

Conventional Model	Proposed Model
Pre-observation conference	Teacher team develops unit plan
Lesson plan	Team writes common assessments
Lesson observation	Principal gives feedback on these to team
Evaluation write-up	Team meets during unit to share ideas
Post-observation conference with teacher	Brief principal visits to classrooms
Occasional walkthroughs	Face-to-face feedback to teachers
	Team gives a common assessment
	Team analyzes unit learning results
	Team reports results to principal
	Principal discusses results with team

8. *Arrange for high-quality feedback on lessons for teachers.* Once a principal has made the shift to short, frequent classroom visits followed by face-to-face feedback and is looking at unit plans and successfully orchestrating teacher teams to focus on student results, who will give teachers feedback on full-period lessons? The principal won't have time but might arrange for instructional coaches or other teachers to do longer observations and follow-ups on les-

sons. Colleagues and coaches can give valuable feedback to teachers, especially when their input is part of a "lesson study" process. But there's a potential problem with peers observing one another — *the culture of nice*. It's hard to give critical feedback to people you eat lunch with every day. Videotape is a better medium for taking an unsparing look at a lesson. There's no better way to see the flaws in one's teaching (and appreciate the strengths) than to watch a videotape with a critical friend. Videotaping also requires much less skill than writing up a lesson observation.

The goal of all supervision, whether it comes from the principal's short visits or from a more lengthy peer or video observation, is to foster a real openness to feedback, install the supervisory voice in teachers' heads, and breed an acute consciousness of student learning results. We want individual teachers and teacher teams to be thinking constantly about whether students are learning and what can be done to get better results.

9. *Create a professional learning culture in the school.* Teachers and principals need preparation and support to improve their skills at observing classrooms; giving frank and honest feedback; and assessing unit plans, tests, and data on student learning. The principal needs to be the "chief learner" in this regard, reaching out to the knowledge base and orchestrating study groups, article and book groups, peer observations, and lesson videotapes. The goal is to create a culture in which nondefensive analysis of student learning is "the way we do things around here."

The nine steps above could be carried out within most collective bargaining agreements. The last three would probably require waivers or contract changes.

10. *Use short observation visits to write teachers' final evaluations.* Dispensing with elaborate, announced evaluations is a huge time-saver, and once a trusting climate has been established, it's the ideal scenario. When I was principal of the Mather School in Boston, teachers became so comfortable with my short visits and personal feedback that virtually all of them agreed (via individual sign-offs with the assent of the union representative) to allow me to skip formal observation visits entirely and use my 12 or so short classroom visits-with-feedback to write their final evaluations. (For teachers who were in danger of getting overall unsatisfactory ratings, I went by the book.) The Littleton Public Schools in Massachusetts are in their second year of a negotiated agreement that gives tenured teachers the choice of being evaluated using the traditional approach or using evaluations based on at least 10 short visits.¹⁹

11. *Include measures of student learning gains in teachers' evaluations.* Teachers could be asked to submit evidence of changes in student learning from the beginning

to the end of the year, using before-and-after assessment results or an analysis of portfolios and student work.

12. *Use a rubric to evaluate teachers.* Scoring guides are being used successfully to evaluate student writing and other open-ended work, and a few school districts, including Alexandria, Virginia, and the Aspire Charter Schools in California, have begun to use them for teacher evaluation. Rubrics have several advantages over conventional evaluation instruments: they are more clearly “judgmental,” forcing the principal to give the teacher clear feedback with respect to a standard; they are more informative, telling teachers where they stand on a 4-3-2-1 scale with a detailed description of what performance looks like at each level of proficiency; they counteract “grade inflation,” if it’s clear that very few teachers will be at the advanced level; and they take much less time.

CONCLUSION

Let’s return to the vignette of the teacher who wept after being told that she had mistaught an important math concept. It’s a true story; I was the principal. Looking back, I’ve done a lot of thinking about what went wrong in that situation. The teacher was clearly putting on a special lesson for my announced visit. Her nervousness about the biennial evaluation may have thrown her off her game, and the high-stakes nature of our conference undoubtedly contributed to her feeling of devastation when, in her view, I played “gotcha.” She had been working in isolation from other teachers at her grade level and was probably more focused on impressing me than on bringing her students to proficiency on a fair assessment. The lesson she drew from my criticism — to “never take a risk” — seems like the wrong one, but given the supervision and evaluation process that we were using, it was understandable.

Had this teacher been working in the kind of professional learning community I have advocated in this article, things might have gone differently. She and her teammates would have planned the math unit together, caught the error early on, and figured out a classroom strategy for teaching the concepts. The teachers would have been less concerned about what I thought, if I happened to drop in on a lesson, than on whether the kids were getting it and how they would do on their interim assessments and on the rigorous Massachusetts math test. If I did catch a teaching error during a classroom visit, I would have corrected it in an informal conversation. When their students did well on the end-of-unit assessment, the team teachers would have reported the results to me and their colleagues with real pride — even, perhaps, with tears of a different kind.

If this scenario is to occur, some changes need to be made. We need to streamline supervision and evaluation so that principals can spend their time doing what will make the most difference: quickly and efficiently keeping tabs on what is really happening in classrooms, giving teachers constant feedback, making fair judgments about teacher performance, and getting teams invested in improving student learning and focused on results. Principals need to be able to shape a creative, low-stakes, professional learning community so that teacher teams can continuously improve their students’ chances of succeeding in a high-stakes world.

Principals are ideally situated to start this team-driven “engine of improvement” and keep it humming month after month. A few maverick school leaders are already doing this kind of work on their own. Others need permission from their superiors before they take the leap of faith, let go of the current model of supervision and evaluation, and launch a more powerful learning dynamic. I would argue that liberating principals to do the right kind of work is one of the most important steps a school district can take if it wants to close the achievement gap and get all students achieving at high levels.

1. The distinction between supervision and evaluation in this article is between formative and summative assessment of teachers’ work, between coaching and judging.

2. James Popham, “A Game Without Winners,” *Educational Leadership*, November 2004, pp. 46-50.

3. Personal communication with John King on 6 January 2005.

4. Grant Wiggins and Jay McTighe, *Understanding by Design*, expanded 2nd ed. (Alexandria, Va.: Association for Supervision and Curriculum Development, 2005).

5. Graham Nuthall, “Relating Classroom Teaching to Student Learning: A Critical Analysis of Why Research Has Failed to Bridge the Theory-Practice Gap,” *Harvard Educational Review*, Fall 2004, pp. 273-306.

6. To reduce principals’ evaluation workload, Jon Saphier has suggested adopting a four-year evaluation cycle in which each teacher gets an in-depth principal evaluation once every four years and rotates through other kinds of assessment (e.g., peer evaluation, a study group, self-assessment) in the other three. Jon Saphier, *How to Make Supervision and Evaluation Really Work* (Acton, Mass.: Research for Better Teaching, 1993). This approach sounds promising, but with high rates of staff turnover, it is not without problems.

7. Kim Marshall, “How I Confronted HSPS (Hyperactive Superficial Principal Syndrome) and Began to Deal with the Heart of the Matter,” *Phi Delta Kappan*, January 1996, pp. 336-45; and idem, “Recovering from HSPS (Hyperactive Superficial Principal Syndrome): A Progress Report,” *Phi Delta Kappan*, May 2003, pp. 701-9.


8. Personal communication with John King on 6 January 2005.

9. Personal communication with Douglas Reeves on 5 January 2005.

10. Wiggins and McTighe, op. cit.

11. Douglas Reeves, *Accountability for Learning: How Teachers and School Leaders Can Take Charge* (Alexandria, Va.: Association for Supervision and Curriculum Development, 2004).

12. Mike Schmoker, *Results: The Key to Continuous School Improvement*, 2nd ed. (Alexandria, Va.: Association for Supervision and Curriculum Development, 1999).

13. For information on the Littleton, Mass., teacher evaluation contract language, contact Littleton High School principal Robert Desaulniers at desrobert@aol.com. 

File Name and Bibliographic Information

k0506mar.pdf

Kim Marshall, "It's Time to Rethink Teacher Supervision and Evaluation," Phi Delta Kappan, Vol. 86, No. 10, June 2005, pp. 727-735.

Copyright Notice

Phi Delta Kappa International, Inc., holds copyright to this article, which may be reproduced or otherwise used only in accordance with U.S. law governing fair use. MULTIPLE copies, in print and electronic formats, may not be made or distributed without express permission from Phi Delta Kappa International, Inc. All rights reserved.

Note that photographs, artwork, advertising, and other elements to which Phi Delta Kappa does not hold copyright may have been removed from these pages.

Please fax permission requests to the attention of KAPPAN Permissions Editor at 812/339-0018 or e-mail permission requests to kappan@pdkintl.org.