



A Principal Looks Back: **STANDARDS MATTER**

When Mr. Marshall became the principal of a large elementary school in Boston, he encountered 10 barriers to high student achievement. In this article he examines why he and his staff struggled with those barriers for most of his 15-year tenure at the school. The turning point, he now believes, came when Massachusetts mandated high-stakes tests.

BY KIM MARSHALL

LATE IN THE summer of 1987, when I became principal of Mather Elementary School in Boston, I was absolutely determined to bring a first-rate education to all students. My zeal came from my years as a teacher, graduate student, and central office administrator.

Fresh out of college in 1969, I taught sixth-graders in a Boston middle school and operated pretty much as a lone wolf, writing my own curriculum and at one point actually cutting the wires of the intercom speaker to silence the incessant schoolwide announcements. In my years in the classroom, I knew that my students were learning a lot, but I was never held accountable to any external standards. The whole subject of measurable student outcomes was not part of the conversation with my supervisors.

In 1980, intrigued by the effective schools research (including the work of Ron Edmonds and the British study *Fifteen Thousand Hours*), I spent a year at the Harvard Graduate School of Education and sat at the feet of Edmonds himself. I steeped myself in his research on what appeared to make some urban schools work (strong instructional leadership, high expectations, a focus on basics, effective use of test data, and a safe and humane climate) and said "Amen"

KIM MARSHALL is now working as a consultant and a mentor to new urban principals. He lives in the Boston area. This article is adapted from David T. Gordon, ed., A Nation Reformed? American Education 20 Years After A Nation at Risk (Harvard Education Press, 2003) and is reprinted with permission. The book may be ordered by phone at 800/513-0763 or on the Web at <http://secure.edletter.org/orderbooks>. ©2003, President and Fellows of Harvard College.

to his searing comment on failing urban schools: "We can, whenever and wherever we choose, successfully teach all children whose schooling is of interest to us. We already know more than we need in order to do this. Whether we do it must finally depend on how we feel about the fact that we haven't so far." I was eager to become a school leader and put these ideas to work.

But while I was in graduate school, the voters of Massachusetts passed a tax-limiting referendum that sent Boston into a budgetary tailspin and closed 27 schools. This nixed any chance I had of becoming a principal in the near future, and so I prepared to return to my classroom.

Instead, I was hired by Robert Spillane, Boston's school superintendent, to be chief architect of a new citywide curriculum. Spillane was a forceful advocate of higher student achievement and more accountable schools. This was right around the time *A Nation at Risk* came out, and I found myself in the thick of Boston's response to the "rising tide of mediocrity." Later, under Spillane's successor, Laval Wilson, I directed an ambitious systemwide strategic planning process. My colleagues and I did some useful work, but throughout my years in the central office, I felt that our efforts were often like pushing a string. Without like-minded principals pulling our initiatives into the schools, we often didn't make much of a difference.

When I finally became a principal in 1987, I had seen three aspects of the urban school challenge: 1) talented but often cussedly independent teachers working in isolation from their colleagues and external standards, 2) provocative research theories about the key factors associated with effective urban schools, and 3) the limited power of the central office to change schools that had a great deal of autonomy and very little accountability. Now that I was in the principal's office, I thought I was ideally situated to make a difference for teachers and kids. Was I right?

First, the good news. During the 15 years that I was principal, Mather Elementary School made significant gains. Our student attendance went from 89% to 95%; our staff attendance, from 92% to 98%. Our test scores went from rock bottom in citywide standings to about two-thirds of the way up the list. In 1999, Mather was recognized for having the biggest gains on the MCAS (the rigorous Massachusetts Comprehensive Assessment System) of any large elementary school in the state. And in the spring of 2001, an in-depth review gave Mather a solid B+ based on a careful inspection of the school and its standardized test scores.

The gains we made came in agonizingly slow increments and were accompanied by many false starts, detours, and regressions.

I am proud of these gains and of dramatic improvements in staff skills and training, student climate, philanthropic support, and the physical plant.

But now some more sobering news. The gains we made came in agonizingly slow increments and were accompanied by many false starts, detours, and regressions. Graphs of our students' test scores did not show the clean, linear progress I had expected. Far too many of our students scored in the bottom category on standardized tests, too few were at proficient

and advanced levels, and our student suspension rate was too high. Serious work remained to be done.

When judging schools, everyone is an expert. If student achievement at Mather had been extraordinary, people would attribute it to certain "obvious" factors: the principal's leadership, his 78-hour work week, bringing in great teachers, raising money and lots of other resources, using the research on effective schools, and so on. But our student achievement was not extraordinary. This meant that, despite a lot of hard work, some key ingredients were missing.

WITH THE benefit of hindsight, I have a theory. I think that the absence of meaningful external standards before 1998 prevented our strenuous and thoughtful efforts from having much traction. I would like to test this theory by examining 10 notorious barriers to high student achievement, by detailing Mather's struggle with each of them before the introduction of external standards, and by looking at what changed when Massachusetts finally mandated high-stakes tests.

1. Teacher isolation. In my first months as principal, I was struck by how cut off Mather teachers were from one another and from a sense of schoolwide purpose. I understood teachers' urge to close their classroom doors and do their own thing; I had done the same thing when I was a teacher. But my reading of the effective schools research and my experience in the central office had convinced me that, if Mather teachers worked in isolation, there would be pockets of excellence — but schoolwide performance would continue to be abysmal.

So I struggled to get the faculty working as a team. I circulated a daily newsletter (dubbed the "Mather Memo") and tried to focus staff meetings on curriculum and best practices. I encouraged staff members to share their successes, publicly praised good teaching, and successfully advocat-

ed for a record-breaking number of city-wide Golden Apple awards for Mather teachers. I recruited a corporate partner whose generosity made it possible, among other things, to have occasional staff luncheons and an annual Christmas party.

But morale never seemed to get out of the sub-basement. Staff meetings gravitated to student discipline problems, and, as a young principal who was seen as being too “nice” to students, I was often on the defensive. We spent very little time talking about teaching and learning and did not develop a sense of schoolwide teamwork. The result? Teachers continued to work as private artisans, sometimes expertly, sometimes with painful mediocrity. Overall student achievement continued to be very disappointing.

2. Lack of teamwork. Having failed to unite the staff as one big happy family, I decided that grade-level teams were a more manageable arena in which to work on improving collegiality. I learned how to schedule the school so that teachers at the same grade level had the same free periods. Teams began to meet at least once a week and held occasional after-school or weekend retreats (for which they were paid). A few years later, a scheduling consultant showed me how to create once-a-week 90-minute team meetings by scheduling art, computer, library, music, and physical education classes back-to-back with lunch. This gave teams even more time to meet.

After much debate, we also introduced “looping,” with all the fourth-grade teachers moving up to fifth grade with the same students. (Fifth-grade teachers looped back to fourth.) Teachers found that spending two years with the same class strengthened relationships with students and parents as well as within their grade-level teams, and a few years later the kindergarten and first-grade teams decided to begin looping.

But despite the amount of time that teams spent together, there was a strong tendency for the agendas to be dominated by field trips, war stories about troubled students, and other management issues, with all too little attention to using student work and data to fine-tune teaching. I urged the teams to use their meetings to take a hard look at student results and to use them to plan ways to improve out-

comes, and I tried to bring in trainers and coaches to work with the teams, but I had limited success shifting the agendas of these meetings. In retrospect, I probably would have been more successful if I had attended team meetings and played more of a leadership role, but I was almost always downstairs managing the cafeteria at this point in the day, and I reasoned that teachers needed to be empowered to run their own meetings.

3. Curriculum anarchy. During my early years as principal, I was struck by the fact that most teachers resisted using a common set of grade-level standards. In the central office, I had been involved in creating Boston’s city-wide curriculum goals, and I was stunned by the degree to which these standards were simply ignored. While Mather teachers (and many of their counterparts around the country) enjoyed their “academic freedom,” it caused constant problems. While teachers in one grade emphasized multiculturalism, teachers in the next grade judged students on their knowledge of traditional history facts. While one team focused on grammar and spelling, another cared deeply about style and voice. While one encouraged students to use calculators, the next wanted students to be proficient at long multiplication and division. These ragged “hand-offs” from one grade to the next were a frequent source of

unhappiness. But teachers almost never shared their feelings with the offending colleagues in the grade just below theirs. That would have risked scary confrontations over deep pedagogical disagreements, which the teachers were sure would undermine staff morale. But the absence of honest discussion — culminating in an agreed-upon grade-by-grade curriculum — doomed Mather to a deeper morale problem stemming from suppressed anger and lousy test scores.

I saw curriculum anarchy as a major leadership challenge, and I tried again and again to get teachers to buy into a coherent K-5 sequence. At one all-day staff retreat, I asked teachers at each grade level to meet with those at the grades just below and just above theirs and agree on better curriculum hand-offs. People listened politely to one another, but made very few changes in what they were teaching. Undaunted, I brought in newly written Massachusetts curriculum frameworks and national curriculum documents, but they did not match the tests our students were required to take and could therefore be ignored with impunity. When the Boston central office produced a cumbersome new curriculum in 1996, I “translated” it into teacher-friendly packets for each grade level, but these had little impact on the private curricula in many classrooms.

As a result, far too many of our students moved to the next grade with uneven preparation, and our fifth-graders, although better prepared than most Boston elementary students, entered middle school with big gaps in their knowledge and skills. It was not a pretty picture, and I was intensely frustrated that I could not find a way to change it.

4. Weak alignment. As I wrestled with the curriculum issue, it occurred to me that standardized tests might be helpful in getting teachers on the same page. But virtually all of the tests our students took were poorly aligned with the classroom curriculum (whatever that was) and were not well respected by most teachers. When I was in the central office in the 1980s, Boston’s attempt to write citywide curriculum tests was not well received, and the tests quickly fell into disuse. The classroom tests that teachers gave every Friday and at the end of each curriculum unit were of uneven quality and covered a wide variety of topics with an even wider range of expectations and criteria for excellence. The only tests that got a modicum of respect were the Metropolitan Achievement Tests, which were given in reading and math at every grade level except kindergarten, with school-by-school results helpfully published in Boston newspapers.

Sensing that teachers cared about the Metropolitan, I thought that it might be a lever for getting them on the same curriculum page and making predictable hand-offs

of skills and knowledge to the next grade. I did a careful analysis of the Metropolitan and, without quoting specific test items, told teachers at each grade level what the test covered in reading and math.

Did the teachers use my pages and pages of goals? They did not. And hard as it was for me to admit it, they had a point. Teachers were not confident that they could improve their students’ scores by teaching to the goals I had extracted from the tests — or to Boston’s curriculum, for that matter. The Metropolitan, being norm-referenced tests, were designed to spread students out on a bell-shaped curve and were not aligned to a specific set of curriculum goals. Nor were they “sensitive” to good teaching (in other words, you could work hard and teach well and not have your efforts show up in improved scores). What’s more, I was pushing the ethical envelope by briefing teachers on the standards that were covered by a supposedly secret test. If scores at Mather had skyrocketed, there might have been a major scandal.

But I had stumbled onto an important insight. The key to turning around teachers’ well-founded cynicism about the tests they were required to give and the curriculum they were supposed to teach was to make sure that tests really measured a thoughtful K-12 curriculum. We needed to find both missing elements — a clear grade-by-grade curriculum and aligned tests — at the same time. I could not persuade teachers to buy into one without the other, and, without both, I could not coax teachers out of the isolation of their classrooms. In the next few years, I spent a lot of time looking for the right curriculum-referenced tests and trying to clarify and align the curriculum, but until the late 1990s, I couldn’t find the right combination to get teachers working on a coherent K-5 curriculum.

5. Low expectations. Another barrier in my early years as principal was teachers’ pessimism about producing significant gains in student achievement. Hamstrung by the lack of aligned curriculum and tests, gun-shy about addressing their colleagues’ idiosyncratic classroom goals, discouraged by the visible results of poverty (85% of our students qualified for free and reduced-price meals, and the community around the school was plagued by unemployment and violence), and having never seen an urban school that produced very high student achievement, many teachers found it hard to believe that it could be done. They regarded themselves as hard-working martyrs in a hopeless cause.

Going for broke in my second month as principal, I brought in Jeff Howard, the charismatic African American social psychologist, and his “Efficacy” message hit home. Jeff spoke of combating our students’ lack of achievement motivation by getting them to see that people are not just

born smart, they can *get smart* by applying effective effort. He grabbed the faculty's attention with the notion that we could dramatically improve our results by directly confronting the downward spiral of negative beliefs about intelligence and effort. Over lunch, most of the staff buzzed with excitement.

But after lunch, Jeff had to go to another school, and the consultant he left in charge was swamped by defensive and increasingly angry reactions. Was he suggesting that teachers were racist? Was he saying that teachers were making the problem worse? And what did he suggest they do on Monday? By late afternoon, it was clear that my gamble to unite the staff around this approach had failed.

Licking my wounds, I took a more incremental approach over the next few years, using private conversations, team meetings, the "Mather Memo," and research articles to drive home the message that much higher student achievement was doable at Mather Elementary School. I sent small groups of teachers to Efficacy training, and eventually I brought in one of Jeff Howard's colleagues to do a three-day workshop for the whole staff. It was an uphill battle, but gradually Efficacy beliefs were accepted as part of the school's mission, and it became taboo to express negative expectations about students' potential.

But we still did not see dramatic increases in our scores on the Metropolitan Achievement Tests. Belief was not enough. We needed something more to boost achievement in every classroom. We needed a grassroots strategy that harnessed the potential energy within teacher teams to pursue clear curriculum goals using interim assessment data. And perhaps we did not need to fight the battle over low expectations up front. If we could get teachers to suspend their disbelief — to behave as *if* their students could achieve proficiency — perhaps their deeper beliefs would slowly change as students began to show their true potential.

6. Negativism. The area in which I was least effective in my early years was dealing with some very strong personalities who declared war on my goals as principal. It's been observed that inner-city schools attract and nurture strong personalities and can develop a negative culture. When a leader starts to mess around with the unspoken expecta-

tions and mores of such a culture, he or she is playing with fire. When I appeared on the scene preaching that "all children can learn," these teachers reacted with disbelief and active resistance. A parody of the "Mather Memo" was slipped into staff mailboxes ridiculing my idealism: "For Sale: Rose-Colored Glasses! Buy Now! Cheap! Get that glowing feeling while all falls apart around you."

I was aghast at the vehemence with which these teachers attacked me. Monthly confrontations with the faculty senate invariably got my stomach churning, and I took to quoting Yeats: "The best lack all conviction, while the worst / are full of passionate intensity." I tried to keep up a brave front and jokingly dubbed my antagonists the Gang of Six, but I could not hide my dismay when it was reported to me that, on the day of the first Efficacy seminar, one of these teachers was overheard to say in the rest room, "If I had a gun, I'd shoot Jeff Howard dead." I was continually off balance, and every mistake I made became a major crisis. ("People are outraged! Morale has never been worse!") On several occasions, I failed to set limits on outrageous and insubordinate behavior, didn't assert my prerogatives as principal, and lost face with the rest of the staff, who secretly wanted me to change the negative culture that had dragged down the school for years.

Over time, the most negative people realized that I wasn't a quitter and transferred out. But they had understudies. Almost every year I was principal, I had to battle (not always very skillfully) for the hearts and minds of the silent majority, and only very gradually did the school develop a more positive culture. It was only when we were confronted with a compelling external mandate that the positive folks found their voice and the remaining negative staff members finally lost their ability to poison the climate of the school.

7. A harried leader. As every busy principal knows, the hardest part of the job is to make time for instructional leadership while dealing with the myriad administrative and disciplinary challenges of running a school. The limitless number of tasks that need to be done can also serve as a very plausible excuse for not dealing with the more intractable work of improving teaching and learning. After my initial setbacks with the staff, I plunged into a major campaign to raise money for a gala 350th anniversary celebration. (Mather is the oldest public elementary school in America, and I was tipped off by a local historian that 20 October 1989 was the 350th anniversary of the founding of the school.) I was successful in sprucing up the aging and neglected building and garnering a great deal of publicity for the school. Although these improvements were important, I had no illusions that they were the heart of the matter or even the best strategy for improving student achievement.

As I got better at handling the constant stream of "over-the-transom" demands on my time, I prided myself at being able to juggle several balls at once, and I often quoted an intern's observation that I had 200 separate interactions in a single day — and that did not include greeting students in the halls. I became an "intensity junkie," addicted to being frantically busy and constantly in demand. I had fallen victim to HSPS — Hyperactive Superficial Principal Syndrome — and was spending far too little time on teaching and learning.

This realization led me to devise a plan for dropping in on five teachers a day for brief, unannounced supervisory visits.² These visits and my follow-up conversations with teachers gave me a much better handle on what was going on in classrooms, improved my rapport with the staff, and formed the basis for much more insightful performance evaluations.

But like a recovering addict, I continued to struggle with HSPS on a daily basis. I gradually accepted that I could not (as I had naively hoped) be the school's staff developer. I got better at distributing leadership, putting grade-level teams in charge of key functions, and bringing in "coaches"

in literacy, math, and science to work with teachers in their classrooms and team meetings. I stopped sending teachers off to isolated workshops and invested in training within the building. These changes greatly improved the quality of staff development for teachers, but test scores were still not improving as much as we hoped.

8. Not focusing on outcomes. I became increasingly convinced that the most important reason for our disappointing student achievement was that we were spending too little time actually looking at what students were learning. The teachers' contract allowed me to supervise classroom teaching and inspect teachers' lesson plans, but woe betide a principal who tried to evaluate a teacher based on student learning outcomes. Resistance to evaluating teachers on results is well founded at one level: unsophisticated administrators might use unsuitable measures like norm-referenced tests to unfairly evaluate teachers for failing to reach grade-level standards with students who were poorly taught the year before or had significant learning deficits.

But not looking at the results of teaching during the school year is part of a broader American tendency to "teach, test, and hope for the best." The headlong rush through the curriculum (whatever it might be) is rarely interrupted by a thoughtful look at how students did and what needs to be fixed right now or changed next year. For a principal to ask for copies of unit tests and a breakdown of student scores is profoundly countercultural. These private artifacts are none of the principal's business. Teacher teams don't use them much either. At the end of a teaching unit, team meetings are rarely devoted to looking at which teaching "moves" and materials produced the best gains, which were less successful, and which students need more help. With one notable exception (discussed below), I failed to get teachers to slow down, relax about the accountability bugaboo, and talk about best practices in the light of the work students actually produced.

9. Mystery grading criteria. Looking at student work, especially writing and other open-ended products, is virtually impossible without objective grading tools. In many schools, the criteria for getting an A are a secret locked up in each teacher's head, with top grades going to students who are good mind readers. The absence of clear, public, usable guides for scoring student work prevents kids from getting helpful feedback and robs teacher teams of the data they need to improve their performance.

In 1996, Mather Elementary School made a successful foray into the world of standards-based thinking. Spurred on by a summer workshop with Grant Wiggins, we wrote scoring rubrics for expository writing that described in a single page for each grade the specific criteria for getting

a score of 4, 3, 2, or 1 in three areas: mechanics/usage, content/organization, and style/voice.

It was striking how much higher our standards were once we had written these rubrics; now we knew what proficiency looked like! We could also say with some certainty that the same piece of student writing would get the same scores no matter who graded it. Encouraged by our success, we began to give students a “cold prompt” writing assignment (a topic they had never seen before, with no help from the teacher allowed) in September, November, March, and June. Teachers scored the papers together and then discussed the results.

This process was a breakthrough. We had found a way to score student writing objectively; we were sharing the criteria with students and parents in advance (no surprises, no excuses); we were giving “dipstick” assessments at several points each year; teachers at each grade were working as a team to score students’ work; and teachers were analyzing students’ work, giving students feedback, and fine-tuning their teaching. We began to see significant improvements in our students’ writing.

But after a few years of regular scoring meetings and charting of student progress, our efforts began to flag. Finding enough time was always an issue, especially since the scoring/data analysis meetings were hard to fit into our 90-minute team meetings and many teachers had after-school family commitments. It became clear that only very strong leadership — or some other equally powerful force — could sustain this kind of work.

10. No schoolwide plan. Over the years, we examined many different programs to turn around student achievement — Effective Schools, Efficacy, Success for All, Core Knowledge, Accelerated Schools, the James Comer plan, Schools Without Failure, Multiple Intelligences, Whole Language, Multicultural, and others — but none got the staff buy-in needed for successful implementation. As a result, we kept trying to “grow our own” — an exhausting and frustrating process. In the late 1990s, one “whole-school” reform program was mandated as part of a Boston grant program. We appreciated the help (and the money!) but felt there were crucial pieces missing and drove the program’s administrators crazy by constantly second-guessing their model and adding components of our own. Perhaps we were asking for too much. Perhaps we should have committed to a less-than-perfect program and given it a chance to work. But we were on a constant quest for a better mousetrap.

As we continued our search, two less global programs had a big impact. The first was Reading Recovery, a highly effective, low-tech, data-driven program for struggling first-

graders. What caught the attention of the whole staff was that most of the students who appeared to be doomed to school failure got back on track after 12 weeks of hard work with the highly trained Reading Recovery teachers.

After a few years of successful implementation, there was enough support to get our primary-grade teachers to buy into the Literacy Collaborative Program, which was created by Irene Fountas and Gay Sue Pinnell of Lesley University to align the way reading and writing are taught in regular classrooms with Reading Recovery. All of our K-3 teachers bought into the program and were trained by one of their colleagues through in-class coaching and a 40-hour after-school course in which teachers looked at student work and data (using a new scale of reading proficiency) and talked constantly about best practices in a low-stakes, collegial atmosphere. The program produced significant gains in our student achievement in the lower grades, so during the 2001-02 school year, we introduced the upper-grade version of Literacy Collaborative.

But these very effective literacy programs were not part of a coherent schoolwide plan for change. And this, along with all the other factors discussed above, prevented us from getting the kinds of achievement gains we wanted our students to show.

LOOKING OVER this list of 10 barriers to the success of Mather Elementary, it’s clear that there were powerful forces at work in our school (and other schools as well) that tend to stymie efforts to improve achievement and so widen the achievement gap, perpetuating the so-called Matthew effect (“To those who have, more will be given, and they will have abundance; but from those who have nothing, even what they have will be taken away,” Matt. 13:12). Children who enter school with the advantages of a middle-class home tend to do well, even if they attend ineffective schools. But disadvantaged children desperately need effective schools to teach them key life skills and launch them into success.

Unless there is strong leadership pushing back, the 10 factors will make things much worse for these children. If teachers work in isolation, if there isn’t effective teamwork, if the curriculum is undefined and weakly aligned with standardized tests, if there are low expectations, if a negative culture prevails, if the principal is constantly distracted by nonacademic matters, if there are no objective grading tools, if the school does not measure and analyze student outcomes, and if the staff lacks a coherent overall improvement plan, then students’ entering inequalities will be amplified,

and poor children will fall further and further behind, widening the achievement gap into a chasm.

This set of circumstances presents a tremendous professional — and moral — challenge to principals, because school leaders are ideally situated to influence all 10 factors. If the principal is an effective instructional leader, the forces will be pushed back (at least for the time being), and the gap will narrow. For vulnerable, school-dependent children, this is a godsend.

How did I measure up to this challenge? For more than a decade, I had limited success pushing back the powerful gap-widening forces. Mather students began to make real progress only when strong external standards were introduced, and that did not happen until Massachusetts introduced its high-stakes tests in 1998.

When we heard that 800-pound gorilla knocking on our door, the turnaround happened with amazing speed. As our fourth-graders took the first round of MCAS tests, one of our most effective teachers (who was teaching fourth grade that year) burst into tears at a staff meeting and proclaimed, “No more Lone Ranger!” She pleaded with her colleagues in kindergarten, first grade, second grade, and third grade to prepare students with the necessary building blocks so that she would never again have to watch her students being humiliated by a test for which they were so poorly prepared. It was a dramatic moment for the staff.

Some Mather teachers joined the hand-wringing across Massachusetts about making students the victims of a forced march to high standards. But in a subsequent meeting, the staff sat down and actually took portions of the MCAS and came to these conclusions: 1) although the test was hard, it really did measure the kinds of skills and knowledge students needed to be successful in the 21st century; 2) because the MCAS test items and underlying curriculum were in the public domain, it was possible to align the curriculum and study for the test (we were lucky to live in Massachusetts; some states use norm-referenced tests and keep their tests secret); 3) our students had a long way to go; but 4) most of our kids *could* reach the proficient level if the whole school taught effectively over time.

The only problem was that the Massachusetts frameworks and tests were pegged to grades 4, 8, and 10, leaving some uncertainty about curriculum goals for the other grades. But the fourth-grade tests and accompanying curriculum “bridge” documents gave us much more information than we had before. We set up committees that worked with consultants to “tease back” the standards, and we then worked as a staff (with parent input) to create 12-page booklets for each grade level, containing clear curriculum proficiency targets accompanied by rubrics and exemplars of

good student work. We also set a schoolwide achievement target four years into the future (an idea suggested by Jeff Howard), and then we spelled out SMART (Specific, Measurable, Attainable, Results-oriented, and Timebound) goals for each grade level to act as steppingstones toward the long-range target. (For example, 85% of third-graders will score proficient or above on the end-of-the-year math assessment.) We updated the SMART goals with higher and higher expectations in each subsequent year.

I believe that the rigorous, high-stakes MCAS tests had a dramatic impact on all the areas with which we had struggled for so long. The grade-by-grade MCAS-aligned targets we created put an end to curriculum anarchy and kicked off the process of locating or writing during-the-year assessments aligned with those goals. This in turn focused the curriculum and produced data that the teaching teams could sink their teeth into, giving much more substance to their meetings. The expository writing rubrics we had developed just a year before were key tools in objectively measuring students’ writing and displaying data in ways that encouraged effective team discussions on improving results. As teachers gave up some “academic freedom,” their isolation from one another was greatly reduced and grade-level teams had a common purpose.

Our staff confronted the issue of teacher expectations when we sat down and took portions of the MCAS ourselves, and there was a lot less negative energy as we united in a relentless push for *proficiency* (a term we had never used before). My work as an administrator was much more focused on student learning results, which helped in my ongoing struggle with HSPS. And finally, our search for the perfect school improvement program came full circle to a very straightforward mission: preparing students with the specific proficiencies needed to be successful at the next grade level and to graduate from fifth grade with the skills and knowledge to succeed (i.e., get on the honor roll) in any middle school.

In the summer of 2002, as almost all the key pieces fell into place at Mather Elementary School, other personal and professional factors led me to resign as principal. Looking back, I feel terrific about how far my colleagues and I brought the school, and I regret that I couldn’t be part of taking the school to the next level. In the year since I left, I have done a lot of thinking about what it takes to produce very high achievement for all students — and why it took so long for our school to approach that goal. One thing is clear: from the students’ and parents’ point of view, it is unacceptable that it should take educators 15 years to create an effective school!

We did a lot of things right. We put most of the critical

elements in place. We took full advantage of the leverage provided by external assessments. But I believe one key element was missing: a “power learning cycle” in which teacher teams use data from quarterly assessments to track the percentage of students achieving proficiency, focus on key areas of weakness, and invent new strategies to boost achievement in the next nine-week cycle. Inspired by business guru W. Edwards Deming, educators like Mike Schmoker, Richard DuFour, and Jeff Howard have advocated for this approach to get teacher teams to produce very high student achievement. I am convinced that, when teachers work this way (which is possible only when a number of other factors related to the school’s organization and culture are present), student achievement will soar.

Let’s return to the Ron Edmonds statement I cited earlier. Edmonds often said that the existence of even *one* effective urban school (and he found a number) proved that we knew how to turn around failing schools — which meant that there was no excuse for any urban school to be ineffective. With these words, Edmonds laid a colossal guilt trip on urban educators who were not getting good results. His stinging rebuke may have jolted some educators out of their fatalistic attitudes and started them thinking about ways to improve their schools. But was Edmonds right that we knew in 1978 how to turn around failing schools? Was he fair to thoughtful, hard-working school leaders? Was he perhaps a little glib about what it would take to close the gap?

From my experience as a principal, I can testify that Edmonds and his generation of researchers gave us an inspiring vision of the key elements of effective urban schools, and their lists have held up remarkably well over the years. But they did not provide a detailed road map to help a failing school get out of the woods. Without a map, success depended too much on extraordinary talent, great personal charisma, an impossibly heroic work ethic, a strong staff already in place, and luck. This allowed cynics to dismiss isolated urban successes as idiosyncratic and to claim that they proved nothing about broader school change.

But Edmonds’ much more basic and lasting contribution was in getting three key messages into the heads of people who cared about urban schools:

- Demographics are not destiny: inner-city children can achieve at high levels.
- Some specific school characteristics are linked to beating the demographic odds.
- Therefore, we need to stop making excuses and get to work.

Turning around failing schools is extraordinarily difficult. My 15-year struggle to make one school effective brought

me face to face with my own personal and professional limitations and made me a student of school effectiveness and the key factors that get people and institutions to work more successfully. I learned that the starting point has to be an almost religious belief that it can be done, and Edmonds served as high priest in that regard. I also learned how important it is to have a vision of what an effective school looks like, and Edmonds’ correlates of effective urban schools have given us a sense of the pieces that need to be in place for all children to learn at high levels. I have been further inspired by examples of schools that have been successful, and the “90/90/90” schools (90% of children live in poverty, 90% are children of color, and 90% are meeting standards on rigorous external assessments) documented by Douglas Reeves and others have sent a powerful message to all urban educators. And I also benefited from the accumulated craft knowledge about turning around schools — the trial-and-error wisdom of principals — and some helpful ideas from the business world.

If I could travel back in time to 1987 and start all over again as principal with what we know now, I believe things would have improved a lot more quickly. But even with the advantage of a decade and a half of accumulated wisdom in my head, my second shot at improving Mather Elementary School would still have come up short. And that is because a critical element was not present in 1987: strong external standards linked to high-stakes curriculum-based tests. I believe that the arrival of standards and tests in the late 1990s provided the traction that a principal needs to push back the powerful gap-widening forces that operate within all schools.

Building on the accumulated lessons of researchers and practitioners, today’s principals are in a much better position to produce results than were their peers in the 1980s. With the right support from their districts (aligned curriculum, contractual time, and training are a great start), with the passionate belief that their students can achieve proficiency, with a clear vision of what makes a school effective, with a clear understanding of the lessons of school change, and with external assessments as levers for that change, principals should be able to lead a school staff to bring a first-rate education to every child.

Ron Edmonds would have smiled about that. So should all of us.

1. Ronald Edmonds, “Effective Schools for the Urban Poor,” *Educational Leadership*, October 1979, p. 23.

2. See 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. 