

Characteristics of Highly Improved Schools: A Case Study of Selected Schools in Economically Disadvantaged Districts¹

Richard Tappan
National Center for the Improvement of Educational Assessment, Inc.
February 2003

¹Background

In July of 2000 the Center for Assessment embarked on a search for schools with recent histories of significant improvement in state assessment scores after a number of years of performance considerably below state averages. In studying a set of selected schools, the Center hoped to be able to answer four questions:

1. Are there schools with large, sustained improvements in test scores?
2. Do the apparent gains in test scores reflect real changes in learning?
3. How much can an effective school improve in real learning from one year to the next?
4. What are common characteristics of highly improved schools?

This report addresses the first, second, and fourth questions. The third question, the amount of improvement that is reasonable to expect from one year to the next, has been the subject of intense national debate. Whoever could provide a definitive answer to that question might also be able to turn lead into gold. Although we do not claim to be education's alchemists, we believe that our study sheds some light on the issue of reasonable expectations of improvement. For reflections on that issue in regard to the schools cited, please see the presentation by Gong and Tappan (2001) located on the Center for Assessment's website (www.nciea.org) under Publications, "How Much School Improvement Should State Accountability Systems Require?"

The Selection Process

It is easy to find lists of schools with significant gains in mathematics and reading scores in one or two years, but it is another matter to find schools with reliable data of low performance for three or more years prior to the first year of major improvement, and reliable data for significantly improved performance for three or more years following those initial gains. In our letters to superintendents and principals of schools with data showing unusual levels of improvement, we specified that we would like to schedule visits if the school could cite a significant intervention that they believed had led to the improved scores, and if they could provide reliable data for three years before and after the intervention.

Our criteria used in selecting the schools for study and visitation also included a high percentage of students on free/reduced lunch (all of the schools cited exceeded 50%), and generally a high percentage of minorities (only one school in the study had a minority population of less than 40%).

The Center examined lists of highly improved schools on state department of education websites, as well as lists published by the Education Trust, Blue Ribbon Schools, Coalition for Essential Schools, and a number of other sources. We then wrote to approximately 125 schools out of the initial list of approximately 500 that were listed as "highly improved" in test scores and also had a high percentage of students on free/reduced lunch and/or a high percentages of minorities. The Center asked for pre- and post-improvement data, and we also asked for a description of the interventions

¹ An extension of a presentation given at the Reidy Interactive Lecture Series, Nashua, NH, October 4, 2001.

that were the apparent cause of the sudden surge in scores. Many schools were eliminated from the study because they either did not meet the socio-economic criteria, or the administrator could not cite any specific intervention that prompted the improvement. Some said, “We’ve been working really hard,” while others acknowledged that the improvement was fundamentally a result of a major focus on test-taking strategies or a narrowing of the curriculum to provide more time in the school day to push math and/or reading scores higher.

After telephone and email follow ups with administrators, the Center narrowed the list of 125 schools to 46 that seemed to be the most promising in terms of apparent systemic changes as the likely cause of the sudden improvement in student performance scores. We also requested other data—SAT or ACT scores in the case of high schools—or other assessments that would corroborate the state mathematics and reading test scores. Time and expense prevented us from visiting all 46 of the schools with positive data and with evidence of major intervention. It is likely that the majority of those schools would have provided ample evidence of major and sustained improvement and significant systemic change. In the end we were able to visit thirteen schools from central Maine to the piedmont of North Carolina and from El Paso to Harlem.

Our Visitation Protocol

Not all of the 46 schools were open to a formal visit. Some of the schools have a constant flood of visitors, and they could not arrange for teams of teachers, parents and students to meet with us. In other cases we suspected that the school’s reluctance to schedule a visit was motivated by a concern that we might not be able to validate the claims of significant school-wide change. There is no doubt that administrators and teachers are very busy, but most of the schools on the list of 46 were willing to host us, and agreed to our visitation protocol. We requested that data for at least six years be provided in advance; we asked for a meeting with administrators to question them further about data and interventions, followed by classroom visits, free access to the school in general, and an opportunity to examine a variety of student work. In most cases two members of the Center for Assessment staff made the visits, but on some occasions there were three. We generally stayed for half of the school day. After the visits we requested considerable follow up information, and the schools cited in the report were gracious in responding promptly and specifically to our queries.

Schools Visited but not Cited

All of our visits were highly enlightening about school change, but we were not always able to confirm that systemic academic interventions were the primary or exclusive cause of the improvement in student performance. Although our study sought to exclude schools from consideration that were selective in the composition of the student body based on ability, we found that some schools were selective on other grounds.

Three remarkable public schools in New York City had striking histories of improvement in student performance. (Examining student scores from earlier schools attended contrasted with their scores during their time in the new environment.) However, students and their parents/guardians had to commit to a longer school day and school year, and to a significant level of rigor in academic expectations. If students and their parents/guardians did not agree to the “contract,” the student could transfer to another public school. This selectivity by motivation meant that these schools were not really dealing with the same general population as other schools in the neighborhood, even if they did not exclude students based on achievement.

One high school in Maine had a history of significant improvement, but the growth had not been sustained after the departure of a dynamic principal, and we witnessed inconsistent quality of instruction casting doubt on the validity of the systemic changes claimed.

An El Paso, Texas elementary school could cite improved test scores sustained for three years, but we discovered that test preparation skills had crowded out science, social studies and some other elements of the curriculum. Also, some veteran teachers expressed fears that the mathematics preparation focused on the TAAS (Texas Assessment of Academic Skills) and not on fundamentals of mathematics.

We also visited an elementary school in Charlotte, North Carolina that was nearing the end of its year with a school improvement team² in residence. Since the school was still in the process of major intervention, there were not sufficient data to determine if there were significant improvements in student performance or if systemic changes were accomplished and sustained.

The Seven Schools Cited

After our visits to thirteen schools, we were able to declare seven of them to have significant evidence of major systemic change sustained for three years. Two were high schools, one was a middle school and the remaining four were elementary schools. They were:

1. Allenbrook Elementary School, Charlotte, NC
2. Arlington Elementary School, Gastonia, NC
3. Bel Air High School, El Paso, TX
4. John F. Kennedy Elementary School, Winooski, VT
5. Piscataquis Community High School, Guilford, ME
6. Union Hill Elementary School, High Point, NC
7. Ysleta Middle School, El Paso, TX

We briefly describe each on the following pages.

² The North Carolina Department of Public Instruction assigns School Assistance Teams to a percentage of low-performing schools each year. The teams, composed of 3-5 administrators and teachers, take up residence in the designated schools for a full academic year to study, evaluate, advise and train the administration and teaching staff in strategies to improve student performance.

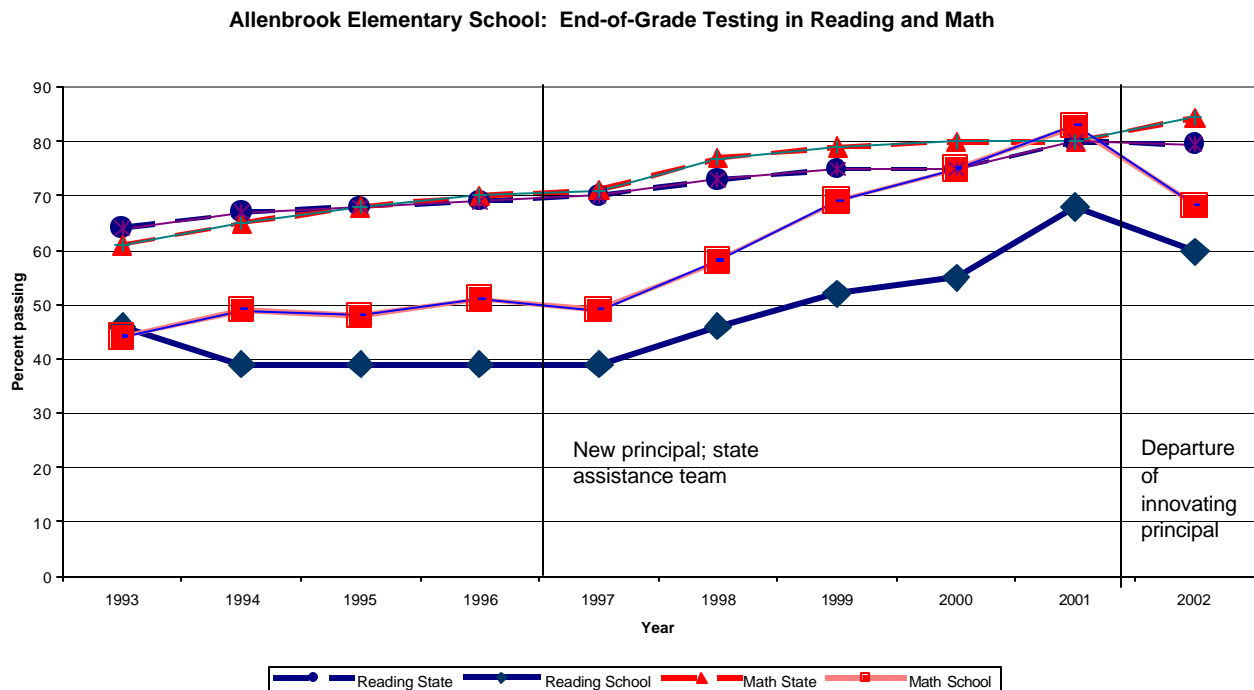
Allenbrook Elementary School, Charlotte, North Carolina

This elementary school with an enrollment of 326 (in 2001) is over 80% minority with 69% of the students on free/reduced lunch. It is a school chronically challenged by extreme student transience. The school was declared “low performing” by the North Carolina Department of Public Instruction and was assigned a school improvement team for the 1996-97 school year. Principal Catherine Hammond found a school in chaos with disorder, few library books with covers, and poor instruction.

Teacher mentoring was the keystone of the intervention, but other changes involved new rules, a major physical sprucing up of the facility, largely through community involvement, and a change in the school ethic from (at best) sympathy for the students’ difficulties to celebration of students’ potential. Within four years the school improved from twenty points below state averages in math to slightly exceeding them. (See Figure 1 below.)

Improvements in teacher morale and community involvement were widely reported during our visit; everyone from secretaries to custodians and cafeteria workers proudly referred to the remarkable change in the environment that has accompanied the rising student performance levels. We witnessed a data-driven school that faced facts and made a dramatic turn around in a very short period of time. In the spring of 2001 Principal Catherine Hammond was transferred to another elementary school in the district. Although the immediate drop in tests scores cannot necessarily be attributed to the principal’s departure, it is probably not mere coincidence that the only drop in test scores since 1997 occurred after the departure of a strong educational leader.

Figure 1

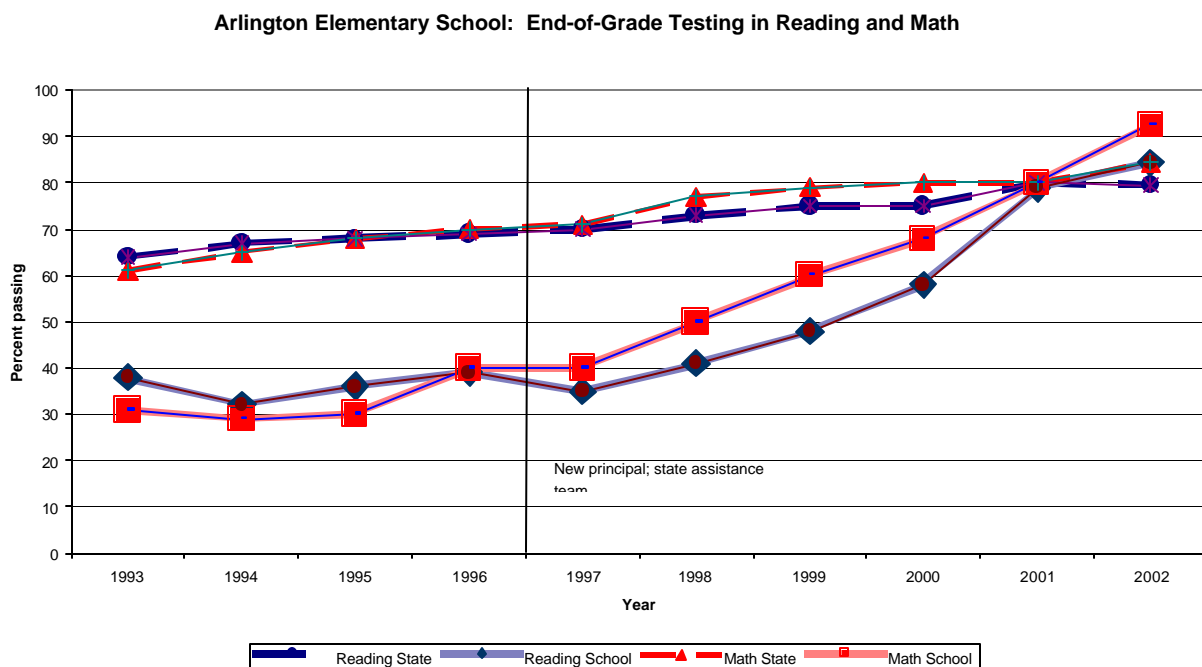


Arlington Elementary School, Gastonia, North Carolina

When Sherry Reynolds took over the principal's office in the summer of 1997, Arlington Elementary was in a sad state of affairs. The school of 400 students was far below state averages in performance, and its reputation as the worst school in the district haunted its teachers. With a tight new administration, teachers were expected to instruct students according to a strict protocol, and discipline in the building was rigorously enforced. Although the school might strike some observers as somewhat rigid, test scores rose dramatically, and over the next four years the school closed the gap between their performance and the state average, and actually exceeded state averages in mathematics and in reading in 2002. (See Figure 2.)

The pride that all school personnel and community members express in their school is a remarkable turn around for a school that was used to ridicule and scorn for many years. As part of the major intervention in uniform instruction, the principal instituted summer home tutoring for students who were behind in reading levels. Arlington teachers, led by their principal, looked the data squarely in the eye and accepted what it revealed about the school. Since 1998 the data has shown steady gains. In 2002 the school exceeded state averages in mathematics and reading for the first time.

Figure 2

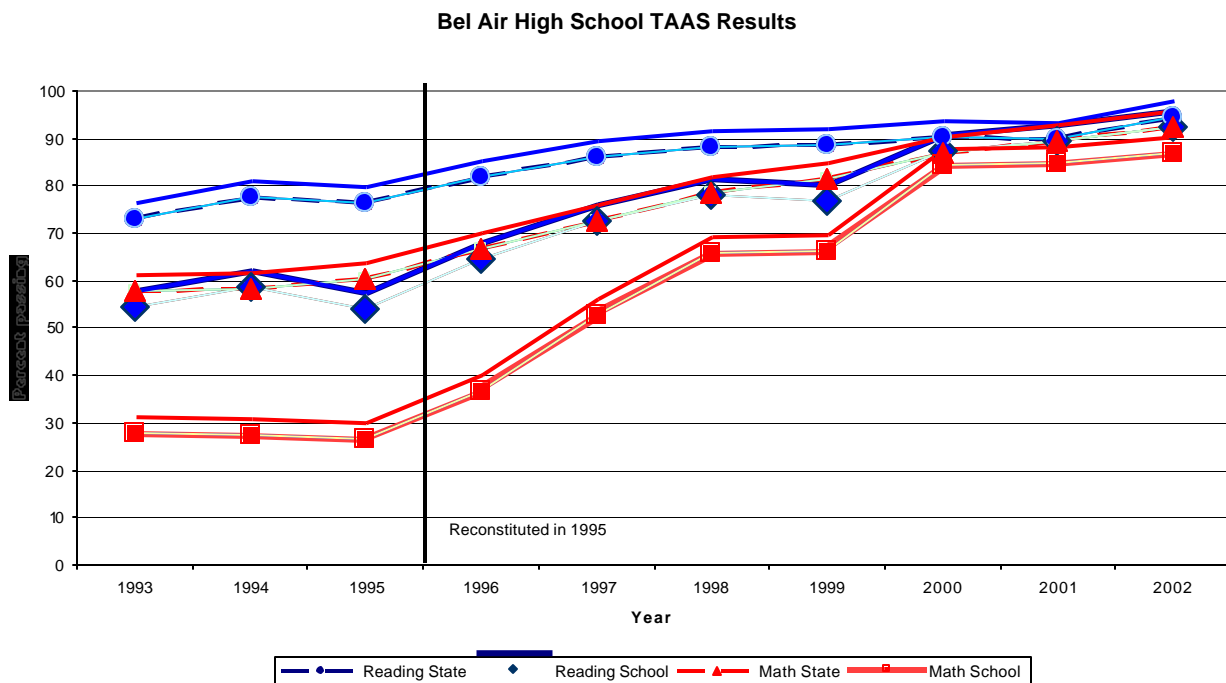


Bel Air High School, El Paso, Texas

This large urban high school of nearly 2,200 students has an Hispanic population of 95% with 83% of the students on free/reduced lunch. The school went from fewer than 30% of the students passing the TAAS mathematics to over 80% passing in a period of seven years. Under the leadership of Principal Vern Butler, now the district superintendent, Bel Air was “reconstituted” with about a 50% turnover in faculty, and a complete overhaul of program. Now, essentially all of the courses are geared to college preparation, and teachers are intensively trained and observed in cooperation with the University of Texas, El Paso. An enrichment and remedial summer school program attracts 50% of the student body, and the percentage of students taking the SAT/ACT college entrance exams rose by 135%.

We observed a very businesslike atmosphere in classes and in the hallways. Students were “dressed for success,” as many of them said; they looked us in the eye, engaged us in conversation, and talked in terms of their goals and the role that education was playing in fulfilling them. The changes were obviously systemic even though TAAS prep is part of the school day, especially as the date of the testing approached. Although the high school has doubled the percentage of students taking the SAT test, the school average combined verbal and math score in 2000 exceeded the 1993 score (when only about 48% of students took the test) by more than 50 points. There was no doubt, however, during our visit to the school that the curriculum had not been compromised for testing. The school had a rich program in music and the arts, theater and languages, and the ethic of the school has been transformed into an atmosphere of pride and hope. In July of 2001 the principal, Vernon Butler, became superintendent of schools. Bel Air sustained its improved student performance after the departure of the principal who led the school through the period of reconstitution. (See Figure 3 below.)

Figure 3



John F. Kennedy Elementary, Winooski, Vermont

This little school in an economically disadvantaged suburb of Burlington, Vermont has achieved a growth in student reading levels from 20% at or above grade level in 1993 to 72% at or above in 2001. The school was noted in the past for its huge problems with transience, disorder, parental neglect and hostility. The Winooski School District has approximately one-third of its families on food stamps compared to a state average of 10%. Over 60% of the children at JFK School are on free/reduced lunch contrasted with approximately 25% statewide.

When Bob Pequignot became principal in 1993, he prepared for a major overhaul of administrative and instructional practices to begin the 1994-95 school year. The principal instituted a continuous improvement model, adopted a school literacy profile, and began intensive staff development in instruction and assessment in cooperation with the University of Vermont.

The school's first priority was reading, but the school has since tackled writing, and has a schedule for continuing improvement in all curricular areas. Each aspect of the curriculum has targets for growth and a strategy to get to the goals. The staff and community are aware of the goals, and have a sense of progress toward reaching them. All of the staff including guidance counselors and specials participate in staff development for the focused areas and incorporate the instructional improvement efforts into their roles. Literacy has been the focus since 1993 with a major impetus in writing beginning in 1997. In 2000 JFK Elementary had the highest writing scores in the state. In 2001 mathematics became the focus of intensive staff development, and scores in the spring of 2002 showed immediate improvement (See Figures 4A and 4B, pg. 8).

The "Continuous Improvement" model may result in a dip or a flattening in other curricular areas while the school's attention is concentrated on its current curricular focus. For example, when the major effort to improve writing started in the 1998-99 school year, (Figure 4C) a dip in reading scores followed (Figure 4B). Despite some apparent negative impact on other curricular areas during intensive focus on the targeted curricular areas, JFK Elementary has kept in close range of state averages in reading while tackling other curricular areas in succession, even in the context of some of the most challenging socio-economic factors in the state.

An observer will note that the school is a model of focused academic progress with serious-minded students expecting to learn and to succeed. There is probably no school in the State of Vermont that is more data-driven in its establishment of goals, nor more convincingly transformed from an earlier history of low expectations.

Figure 4A
JFK School: Ginn Reading Series Grade Level Tests³

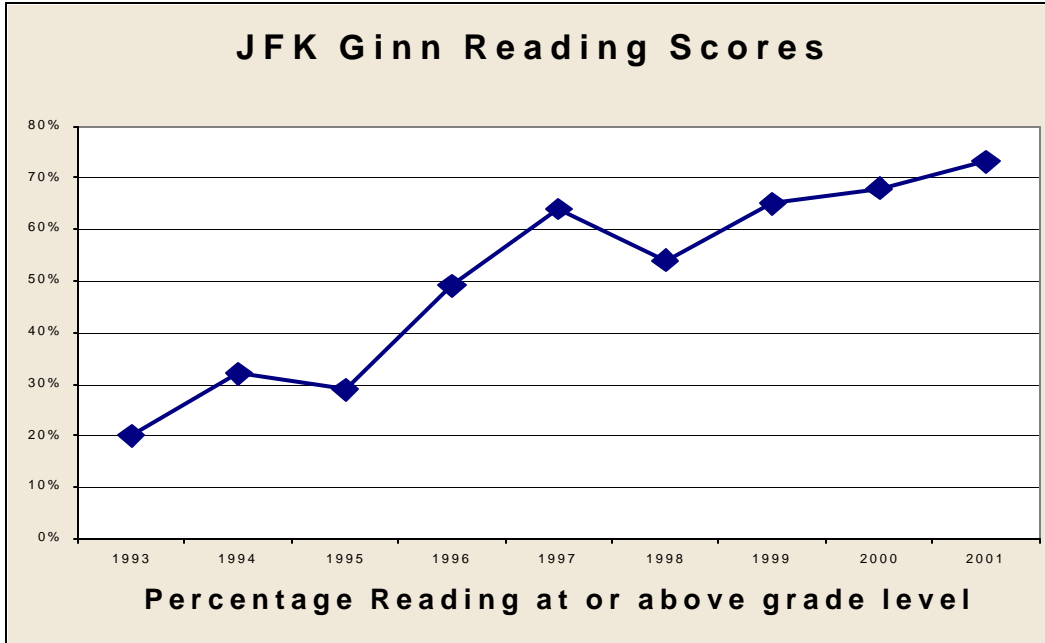
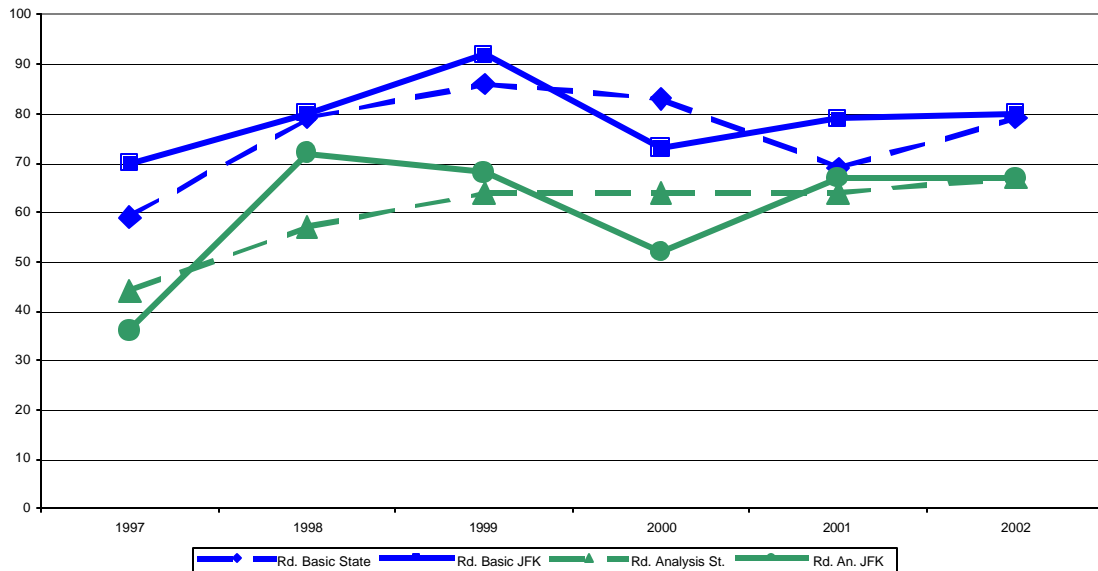


Figure 4 B
JFK School: New Standards English Language Arts Reference Examination⁴ (B)



³ JFK Elementary School initiated major interventions in reading instruction before the adoption of the current state assessment system, so the school derived data on reading improvement from tests provided by the Ginn reading program.

⁴ This State of Vermont assessment of English language arts assesses two components of reading: Basic Understanding and Analysis and Interpretation.

Figure 4C
JFK School: New Standards English Language Arts Reference Examination⁵

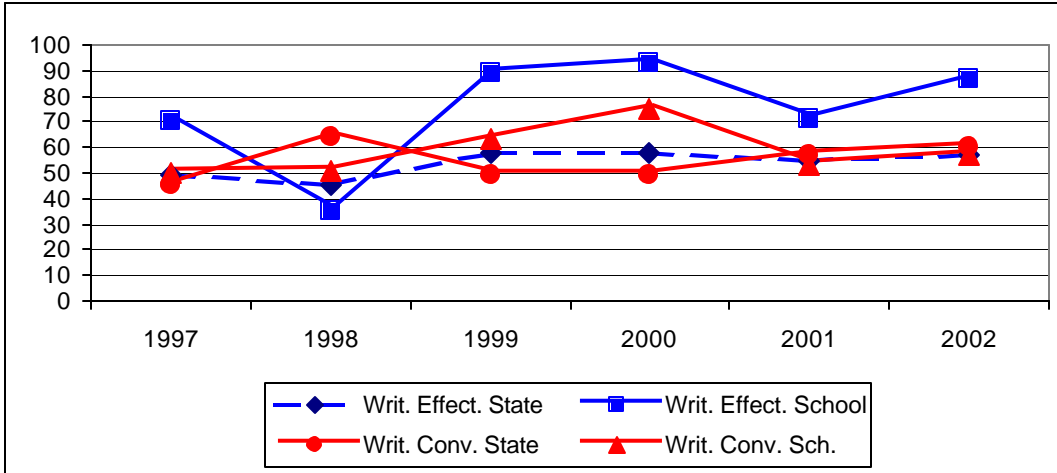
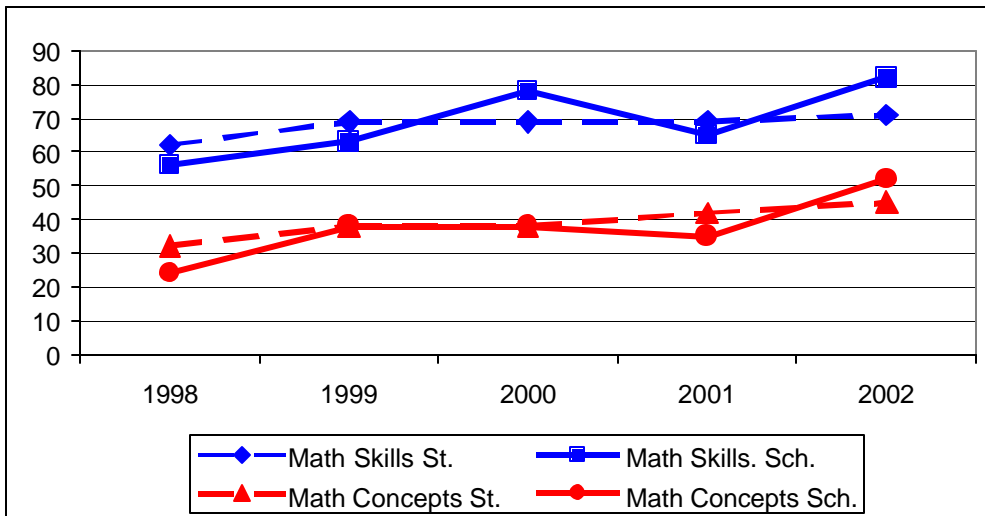


Figure 4D
JFK School: New Standards Mathematics Reference Examination



⁵ The State of Vermont Language Arts examination tests two components of writing: Effectiveness in a variety of formats, and Control of Conventions .

Piscataquis Community High School, Guilford, Maine

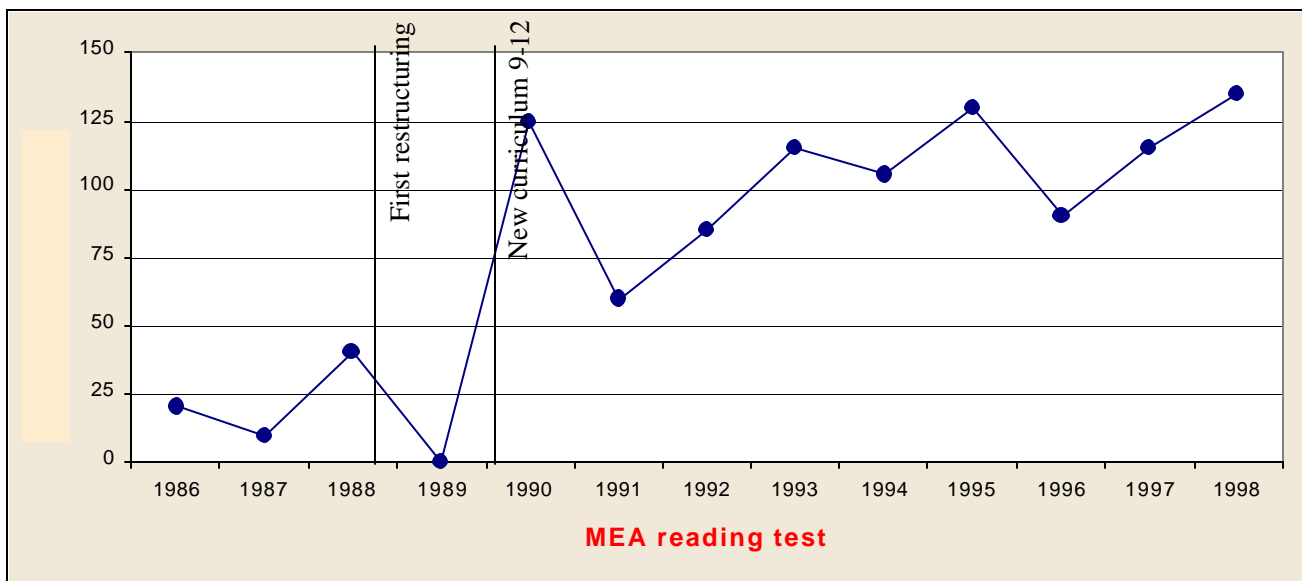
This central Maine high school ranked in the bottom 15% of the state on the Maine Educational Assessment for years. A typical mill town school with a history of low expectations for its students, Piscataquis has 52% of its student body on free/reduced lunch. Before the start of the 1988-89 school year, community members gathered at the high school to hear a startling proposal: The school's principal asked the community to support a transformation in the curriculum to treat all of the students as college material. He introduced a ninth grade core college preparatory curriculum, eliminating vocational and business programs.

The next year tracking was terminated, and the staff began introducing all students to the same fundamental curriculum. That year was a huge adjustment for staff and students alike, and the M.E.A. scores plummeted. However, scores rose after the first year of the intervention, and continued a general dramatic rise despite some bumps in the road (See Figure 5A).

Although many students needed help in their reading levels (many students use tapes to keep up with high level reading), the program provides broad exposure to material traditionally targeted to an academic elite. The results? Piscataquis quickly moved its scores to the top 15% on the M.E.A. tests, and the percentage going on to college rose from the 40 percentiles to the 50s within only a few years and has continued to increase gradually since the mid-1990s.

The school has been the subject of numerous television documentaries, and after more than a decade, it has sustained its record of high achievement. Any visitor will find a relaxed and positive atmosphere in a school where social divisions associated with traditional tracking have been largely eliminated.

Figure 5A
Piscataquis Community High School



One of the challenging aspects of the transition to a single college preparatory curriculum at Piscataquis High was the move to heterogeneous classes in mathematics. The current principal, Bruce Lundberg (who assumed his role after the current curriculum was instituted), attributed the dramatic rise in the Mathematics Educational Assessments in 1990 as a possible “Hawthorne Effect,” i.e., an increase in productivity stimulated by positive attention, not actually produced by the single variable introduced). Scores after 1990 were variable, and the relatively small size of the high school—with the variation in the make up of classes from year to year—may account for the fluctuations.

A major dip in math scores occurred in 1996. In the 1995-96 school year, PCHS introduced a modified college preparatory curriculum in mathematics with two tracks, one “theoretical,” the second, “applied.” According to Mr. Lundberg, the same curriculum is presented, but using differing pedagogical strategies. As the school worked the “bugs” out of the system, there was a significant dip in scores, but once students had been into the system for two years, scores rose again in 1997 and 1998. Since 1999 MEA reading and math scores have flattened (See Figure 5C), but they have continued to stay above state averages, especially in reading. In 2002 PCHS ranked in the top 15% of Maine’s high schools in reading scores on the MEA and in the top 19% in mathematics, approximately the same position the school has held most years since the major curricular changes of the early 1990s.

Figure 5B
Piscataquis Community High School

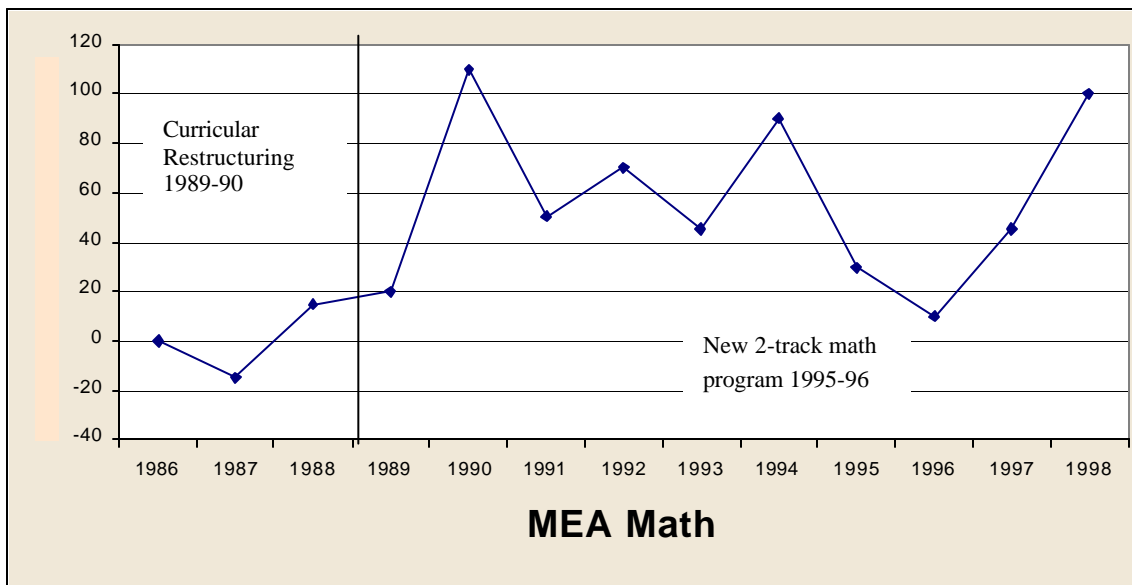


Figure 5C
Piscataquis Community High School
Maine Educational Assessment 1997-2002



Union Hill Elementary, High Point, North Carolina

A less dramatic, but still credible case of significant and sustained improvement is this High Point, North Carolina school of approximately 400 students. Union Hill has an 82% minority rate and 84% of students on free/reduced lunch. When Susan Britt assumed the principal's role in the summer of 1997, the school was officially listed as "low performing" by the North Carolina Department of Public Instruction. A school improvement team was assigned to work with Ms. Britt in making changes.

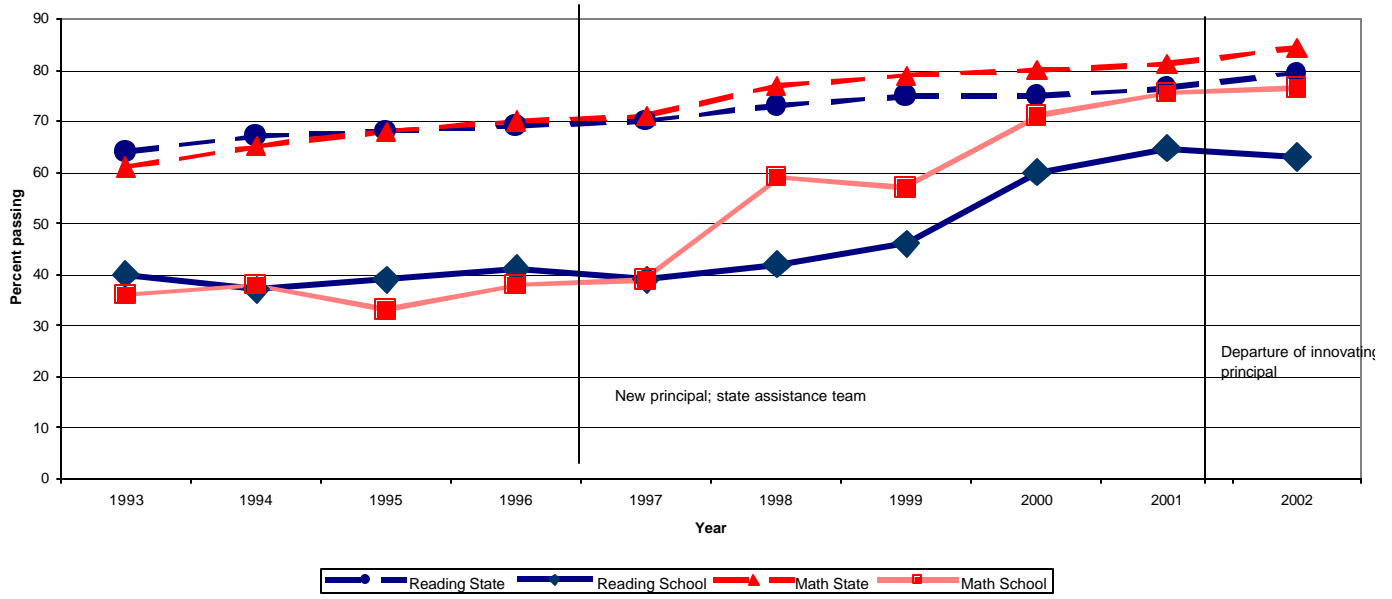
Ms. Britt said that the first order of business was to institute clear policies and rules, something that was not in evidence when she arrived. She said, "I wanted everyone to know that I say what I mean, and I mean what I say." This integrity between policy and practice was echoed by staff members—from teachers to custodians—as a real change in the school's values. As was evident in other highly improved schools, there was immediate attention to the physical appearance of the school (painting, murals, welcoming areas, neat grounds, etc.), and an expectation of achievement.

We did find during our visit some inconsistency in the quality of instruction and decorum in classrooms. For the most part classes were focused and the interaction of teachers and students was very positive. There were some exceptions. Although 60% of the teachers either resigned or transferred after the year of major intervention, we found that some of the staff had not made the "paradigm shift" evidenced in most of the school.

Another difference between Union Hill and the other two highly improved North Carolina schools we visited during April 19-22, 2001 was the expectations for further growth. At Union Hill the principal commented that she believed that "we have probably gone about as far as we can," whereas the principals at Allenbrook and Arlington spoke of reaching state averages soon and setting no limits on how far they might grow beyond state averages in years ahead. When Susan Britt was transferred to another low-performing school in the summer of 2001, there was a modest dip the reading score, but a slight improvement in the mathematics score. Union Hill improved thirty points in mathematics and twenty-six points in reading since the intervention of 1996-97, but it has not broken through state averages to date.

Figure 6

Union Hill Elementary School: End-of-Grade Testing in Reading and Math



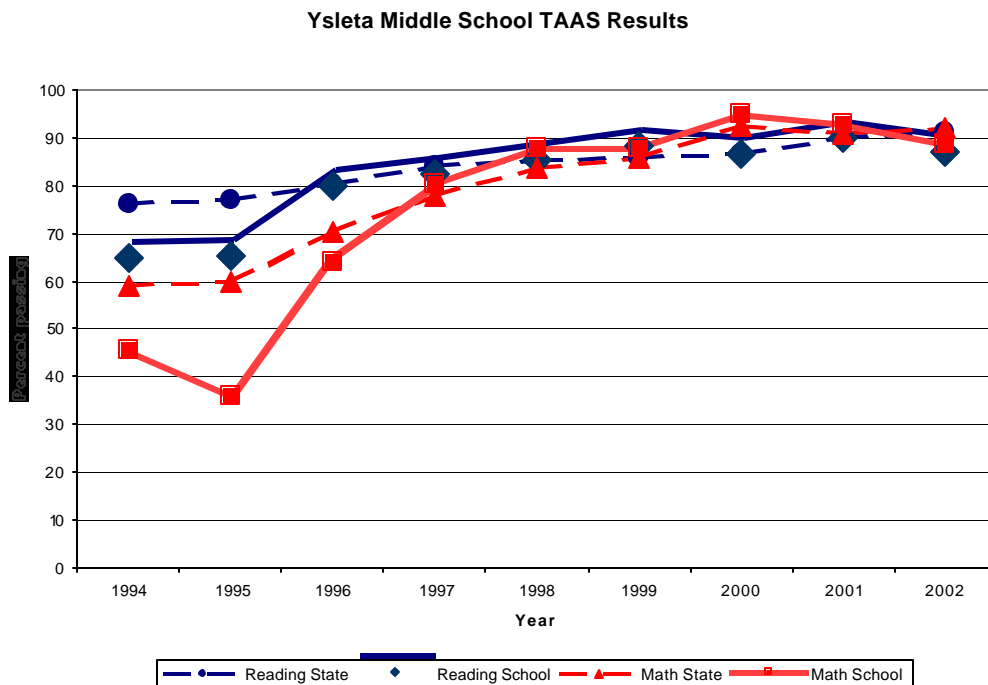
Ysleta Middle School, El Paso, Texas

Almost in sight of the Rio Grande is an unimposing, flat-roofed single story building with heavy grates over the windows. Many illegal aliens have run through Yelseta Middle School's grounds in their efforts to elude the border patrol, and the crime rate in the area forces extra security precautions on this seventh and eighth grade middle school of approximately 430 students. The school was a depressing place with poor scores and little hope when Barbara Trousdale took over as principal in 1991.

YMS is over 95% Hispanic with over 40% of the student body with limited English proficiency. The major intervention that started Ysleta on its climb out of the academic cellar was teacher mentoring and the development of professional portfolios. Teachers maintain a thick portfolio of evidence, not of their *teaching*, but of their students' *learning*. Parents are drawn into their students' education, but they are also offered a wide array of evening classes. They are assisted in their parenting and advocacy by an intense program of community involvement that promotes a college education as the expectation for all students. Although Ysleta is a middle school, the teaching and administrative staff begins working on the skills and the mindset for academic success, and that was evident in the work posted throughout the building. The positive body language of students is striking to observers. There is a level of maturity and self-esteem that is not common at the middle school level, especially in schools with a long history of poor academic results.

Ms. Trousdale and her assistants are assertive grant writers as well, and the involvement of the business community within the City of El Paso and from beyond has helped Ysleta Middle School gain a reputation as one of the most credible success stories in West Texas. Although the interventions begun in 1993 resulted in more success with mathematics than with reading, improvements in both core areas are strong. However, faculty still describe the social studies and science curricula as "weak," but we observed a rich and diverse art and music program and other signs of a diverse curriculum.

Figure 7



Common Characteristics of the Highly Improved Schools Cited

After a year of searching for schools in the top ranks in improvement, sustained for at least three years after a major intervention, we had visited and confirmed data in only seven schools. We are confident that given unlimited time and resources, the Center for Assessment could have found solid evidence of sustained, high levels of improvement in many more of the forty-six schools meeting our initial criteria. However, for our purposes, we limited our search to the schools cited here. Although we recorded comparisons on twenty criteria, there were nine criteria evident in every school cited:

- A strong, visionary principal with support
- A major outside impetus for change
- Resources focused on school-wide academic objectives
- Regular use of data to adjust instruction
- Frequent focused observation of teachers
- Inclusive “can do” atmosphere
- High correlation between vision and practice
- Community and/or parental involvement
- Professional development related to school-wide objectives

All of the schools had some form of “shock therapy” and a shakeup at the top. The shock therapy was usually in the form of a poor rating by the state department of education, but in the case of the JFK School in Winooski, Vermont, it was the principal’s own initiative to face the faculty with the facts of low student performance that began the change. All of the strong educational leaders confronted their faculties and the communities with the truth about student performance, and concentrated on changes in expectations and instruction as the heart of professional development.

In terms of other criteria we noted in comparison, not a single school had a major building renovation as the impetus of change, and three of the seven did not have policies of uniform instructional practices. Most, but not all, experienced the replacement of the majority of the teaching staff at the end of the year of intervention (exceptions were JFK School and Piscataquis Community High School). The most noteworthy change of attitude as part of the systemic change was use of data in driving instruction. Strong educational leaders are not nearly as interested in teachers’ knowledge, degrees, teaching materials or styles as they are in their results evidenced in demonstrated improvement in student performance.

What We Learned

From our investigations into hundreds of reported “highly improved” schools, we have learned several lessons that other studies confirm in much more detail:

Major change. . .

1. is possible in as few as 2-3 years, although difficult and rare;
2. can be sustained for years after the initial intervention;
3. can survive the departure of a dynamic principal if that person established ownership by stakeholders in the new vision and is succeeded by another committed leader;

4. is generally heralded by visible improvements in the physical surroundings in terms of order, maintenance, and aesthetics.
5. is accompanied by concentrated retraining in instruction, high level of communication, frequent assessment, targeted resources, and adjustment of program.
6. is evidenced in systemic culture change recognized by all stakeholders who can articulate past changes and current focus.

Although we were not able to contribute any conclusive evidence in the search for the answer to the question, “How much improvement should state accountability systems require?” we were confident that significant and sustained change in a short period of time is possible and can be sustained, even in schools with the most challenging circumstances.

About the author:

Richard Tappan taught middle school and high school language arts and social studies for thirty-one years, primarily in New Hampshire. He was a leader in curriculum development, chaired a regional committee charged with development of a professional development master plan, and has recently worked on a high school competency project for the State of Wyoming using multiple measures in the newly implemented “Body of Evidence” assessment system. Through the Center for Assessment he has conducted studies on End-of-Course Testing, State of Education reporting forms and other issues concerned with assessment practices.