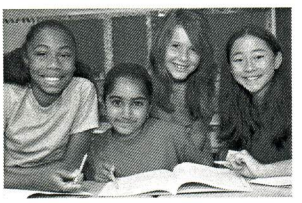
LEADERSHIP INFORMATION

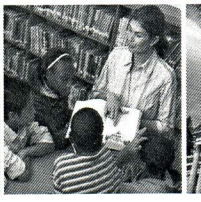
SIRS

Volume 7 No. 3 • Summer 2008

Produced by the School Information and Research Service

Achievement, Assessment, and Accountability, Part I


SCIENTIFIC RESEARCH INTO FAIRLY MEASURING K–12 SCHOOL EFFECTIVENESS

What Student and School Characteristics Impact K–12 Student Performance on the WASL

By JAY MAIDMENT and JIM LEWIS

In a spirit of full disclosure, we are not professional educators; we are scientifically oriented business analysts. For many years, we have made our living developing cost and performance standards for electric utilities and public transportation.

In the fall of 2005, our local newspaper published WASL pass rates for the six major high schools in our area. Each school's results were shown individually for reading, math, writing, and science; and separately for boys and girls.

Three things stood out to us as casual observers. First was the wide variation in pass rates between the schools with the best pass rates almost twice those of the worst. The second was the great disparity in the pass rates for the four subjects; the pass rate for reading was consistently more than twice the pass rate for science. And the third was the superior performance of girls over boys, particularly in reading and writing.

Since then, we have invested several man-years researching these issues. The research is based on the WASL pass rates and other school characteristics of over 1,000 elementary, middle, and high schools for the period 2002 to 2007. The mathematical techniques that we used have been standard practice in competitive industries around the world for decades. The advantage of good scientific analysis is that its results can be tested and replicated by others. Thus, we can offer facts, rather than opinions, about what impacts student and school performance in Washington State K–12 education.

At this point, our research has answered our question about why individual school performance varies so significantly between schools. We have not yet completely determined the reasons for the large differences in pass rates between subjects or between boys and girls. We are highly confident that if these research discoveries are used wisely, state education could be radically improved without any increase in budget. We think you will find the research described in this paper valuable as you work toward improving education.

School Pass Rates

There is an enormous difference (eighty points) in pass rates at individual schools. They range from a low of 16% to a high of 96% for the average of the four subjects over 2002–07 for the 1,000 schools. Between 2002 and 2007 there was a general improvement in overall school pass rates of about three points per year. This rate of improvement is now slowing significantly in all areas.

Our research allows us to understand and quantify the school characteristics that cause (drive) this wide variation in school pass rates. When the underlying school characteristics are known and their impact quantified, it is possible to measure individual school performance on a fair apples-to-apples basis. It should be stressed that this research applies to individual schools, not individual students or teachers.

Using the Office of the Superintendent of Public Instruction's public database, we found ten prime school characteristics that drive pass rates. In general, these characteristics drive (impact) math and science significantly more than reading and writing; and high schools more than elementary schools. We can now predict the actual pass rate at 95% of the 1,000 schools in our database to closer than plus or minus 10 points using these student and school characteristics. There is much that can be learned from the 50 "exceptional schools" (top 2.5% and bottom 2.5%) outside this range, they represent both the most and the least effective educational practices. These exceptional schools are found across the whole spectrum of schools from high poverty to rich suburban.

Scientific Research into Fairly Measuring K–12 School Effectiveness, cont.

As is well known, by far the most important characteristic that defines school pass rates is the level of student poverty at a school, as measured by the percentage of students on free or reduced-price meals. A one-point increase in meals percent causes school WASL pass rates to drop by about one-half point in all subjects.

However, it is not well known that it is poverty, not race, which primarily drives school pass rates. Ethnic minorities have a larger percentage of disadvantaged families than the population as a whole. However, our research clearly shows that disadvantaged children from ethnic minorities perform the same as economically disadvantaged white students. Similarly, student performance is the same for those who do not receive free or reduced-price meals regardless of ethnicity.

There are two exceptions to the previous finding:

- Asian students on average dramatically outperform all other students in all subjects at all grade levels.
- There is a negative impact on school pass rates at those schools that have a significantly racially mixed student body (regardless of ethnicity or poverty level). There is a threshold for this impact that starts when there are three races in a single school, each with more than 10% of the student body. This impact gets progressively worse as the number of significant racial groups and the percentages-per-group increases. This behavior is far more evident at high schools than at elementary schools.

As would be expected, there is a consistent (but only modest) negative correlation at all grade levels between students-per-teacher and pass rates. Fewer students-per-core teachers at a school results in higher pass rates. However, non-core teachers have no impact on WASL pass rates. Thus, additional emphasis by non-core teachers on including core subject skills in their classes seems worth considering.

The higher the teacher's education level, the higher the pass rates for middle and high school students. It has no impact on elementary school students. The average educational level of teachers at a school is defined by the percentage with at least a master's degree. Thus, either the master's programs for elementary school teachers need to be modified or the incentive payments for advanced degrees is not cost effective.

The average years of teacher experience at a school also impacts student performance. The greater the number of years of teaching experience, the lower the student performance for middle and high schools. On the other hand, performance in elementary schools improves dramatically with greater teacher experience. Thus, incentive payments for longevity at elementary schools are cost effective but programs to overcome the decrease in school performance at middle and high schools are needed.

There is some variation in school performance relative to the grade structure in any particular school. For instance, 8–12 high schools have higher pass rates than 10–12 high schools; 6–8 middle schools outperform 7–9; and K–5 outperforms K–6.

Total funding-per-student in Washington State is approximately $8,000 per year. Very disadvantaged school districts have high federal and low local non-tax dollars-per-student, while affluent districts are the reverse. Schools with high local non-tax funds or those with high federal funding do better than average.

School size is not a driver for student performance at elementary and middle schools. There is a correlation of small schools and smaller class size, but it is the student/teacher ratio as opposed to school size, which drives student performance. The only case where school size impacts performance are high schools with more than 2,000 students where there is some deterioration in pass rates as school size increases.

Teacher gender does not have a statistically significant impact on either boys' or girls' pass rates at any of the grades analyzed.

Importance of Reading (and Writing) Ability to Success in Math and Science

There is a consistent and appreciable difference in the pass rates for the four subjects for the thousand schools we studied. As of 2007, in descending order, the average pass rates were 76% for reading, 68% for writing, 56% for math, and 39% for science.

We were somewhat surprised at the wide disparity in the average pass rates for the four subjects. Is science actually twice as hard as reading, or is the level of instruction lower?

In addition, what is the interrelationship between the four subjects? For instance, do schools that have a high pass rate in one subject automatically have high pass rates in all the others? This question led us to an extraordinarily important finding.

There are very strong relationships between the four subjects. As one might expect, reading and writing are closely associated as are math and science. However, by far the most consistent and highly correlated relationship is the very positive relationship between reading ability and math pass rates. Furthermore, this relationship is not a straight line; it is what is known in engineering as a positive "power" curve. In non-technical language, this means that as reading performance improves on the WASL, math performance on the WASL improves significantly faster and at an accelerating rate.

Thus, the ability to read appears fundamental to the ability to learn math. Reading is the gateway skill. If students cannot read, they will not be successful in school. Helping a high percentage of students to reach grade level reading skills in the first few years of elementary school is fundamental to later academic success. Some schools with up to 90% of their students on free and reduced meals have already accomplished this goal so this is not an impossible goal.

This behavior finally explained an anomaly we had been wrestling with for some time. Two medium-sized school districts (Moses Lake and Kennewick) have exceptionally high reading performance in their elementary schools and even higher performance in math. In discussion with these districts we found they both had exceptional reading programs but had not developed specific programs to improve math performance.

It appears that if you teach a child to read properly, its mathematical ability is automatically greatly enhanced. Writing ability also has some impact on math at elementary school level, but far less than reading. However, writing ability becomes progressively more important to math and science performance as a child progresses through middle and high school.

The same behavior is equally apparent for science but analytically more complex because it is also directly impacted by math. Therefore, it would appear that if we wish to significantly improve math and science education we first need to put more effort into reading and writing before looking for improved methods of teaching math and science.

Boys versus Girls

There is a very marked difference in the performance of girls and boys at the same school. On average, approximately 35% more girls pass writing, 15% more pass reading, and 5% more pass math and science.

For years there has been international discussion on the relative learning abilities of boys and girls. The consensus appears to be that girls do slightly better at reading and writing and boys at math and science. However, nowhere have we seen it suggested that girls are 15% better at reading and 35% better at writing than boys. Not to mention that in Washington State they are also slightly better at math and science. So, we asked ourselves, why should this be?

We started by studying the writing pass rates separately for boys and girls in 2005 at each of the 1,000 individual schools. What immediately struck us was the amazing consistency with which girls outperformed boys all the way from elementary through high school.

The 35% difference varied significantly between individual schools from a very small number of schools where boys out-performed girls to schools where they underperformed by more than 50%. The next question is whether this difference is random or related to specific school characteristics.

The first thing we discovered when we tested this was that Asian boys' writing ability was at the same high level as Asian girls. This was consistent at all grade levels and certainly weakens any argument that the current dismal performance of boys is somehow due to genetic disposition. Furthermore, Hispanic and to a lesser extent African American boys also outperform their equivalent Caucasian counterparts.

We also found that poverty (as measured by free meals) impacted boys' writing ability. The poorer a school's students, the larger was the performance gap between boys and girls. This phenomenon becomes progressively more accentuated as students progress from elementary through high school.

Scientific Research into Fairly Measuring K–12 School Effectiveness, cont.

The differential between boys' and girls' reading ability is about half that for writing. Nevertheless, it is also impacted in a similar way by ethnicity and poverty. As we discussed earlier, reading and writing significantly impact math and science. This would help explain why girls are currently also slightly outperforming boys in math and science. In fact, it can be shown that when boys' and girls' reading and writing abilities are the same, boys slightly outperform girls in math and science. If boys passed reading and writing at the same levels as girls there would be a dramatic statewide improvement in all WASL subject pass rates. Is this a discipline or expectation problem?

Fair Report Cards for All Schools

Our initial question about the large differences in school pass rates has been largely answered by our research. Once pass rate expectations are adjusted on consistent, statewide differences in student and school characteristics (using multi-regression analysis) the residual differences in school pass rates is much smaller. Ninety-five percent of the schools perform closer than plus or minus 10 WASL points from where they would be expected to perform. This approach levels the playing field by normalizing for the student and school characteristics largely beyond the local school's control.

It is clearly unfair to compare school pass rates between rich suburban schools and high poverty inner city or rural schools on an equal basis. In reality, some schools with average pass rates of less than 40% are actually doing a more effective job of educating (providing education growth for) their students than rich schools with 90% of their students passing. This is not to take away from the need to increase pass rates significantly (particularly at poor schools). However, it is very important to fairly compare school performance. We should acknowledge outstanding performance for schools with significantly less favorable student and school characteristics and not overly compliment schools with favorable characteristics when they are just performing at the level they should perform.

One particularly worthwhile benefit of our research is the ability to produce fair report cards to compare all Washington State schools on an apples-to-apples basis.

Improving Education

Competitive industries collect and carefully analyze data in order to become more efficient. There is a great deal of data available for K–12 education and there needs to be an increased emphasis on quality analysis of this data. Subjective assessments of program effectiveness are not nearly as effective as scientifically derived findings. Education needs to copy industry's approach to finding what really works (at successful schools) and then broadly implement the proven programs.

The areas of research that we have covered can be used to help improve education in Washington State and nationally.

- The research into what drives test pass rates allows us to identify those few schools (and school districts) that have particularly effective programs, which could be shared with others. It also allows us to compare the cost effectiveness of changing specific school characteristics. For instance, reducing student/teacher ratio by 10% statewide would only lead to a one-point improvement in overall pass rates, not particularly cost effective. Identifying and disseminating best educational practices in a manner similar to competitive industries and medicine is clearly the fastest and least expensive approach to substantial improvements in K–12 education.
- The high correlation between reading and writing, and math and science strongly suggests that improving reading programs would be particularly effective in improving all subject pass rates. It appears that bringing reading for more than 90% of students to grade level in the first few years of elementary education is the optimal approach.
- The current high disparity between boys' and girls' reading and writing ability should be unacceptable. This whole area needs further attention.

All the above results come from our detailed research into the performance of Washington State public schools over the past six years. However, our reading of other research and initial review of other states' data, leads us to believe these results are not peculiar to Washington State or WASL performance. Conceptually, with modest differences, we expect these research results to apply nationally.

Scientific Research into Fairly Measuring K–12 School Effectiveness, cont.

Methodology: Multi-Dimensional Benchmarking

Multi-Dimensional Benchmarking is a significant advancement compared to conventional school benchmarking methods. With conventional benchmarking, schools generally try to compare their school with a small peer group. Unfortunately, conventional benchmarking often results in confusing or misleading conclusions because schools have such a wide variation in average student socio-economic situations, teacher experience, student-teacher ratio, etc.

As an example, forecasting the miles-per-gallon for a car is much more accurate if multiple dimensions (weight, engine displacement, year of manufacture, wind profile, etc.) are all used as opposed to a single dimension such as year or make of car. An additional problem with conventional benchmarking is that it is usually done on a single year's data. Although state school data is good, there are significant year-to-year variations at any school. It is necessary to have at least five years of data in order to see through these year-to-year variations so that important underlying trends become visible.

To arrive at the level of precision required by decision makers, it is necessary to progress from traditional "peer-group" comparisons to multi-dimensional analysis. Most real-world problems are multi-dimensional in nature (i.e., performance is simultaneously impacted by a number of different characteristics or variables). To improve the accuracy of performance measurement and forecasting, it is therefore necessary that this fact be taken into account.

Multi-Dimensional Benchmarking uses historic data to test for and quantify the relationships that exist between the physical aspects of a school (or any other entity) and the performance of that school or entity. Using an extensive database that we compiled on over 1,000 regular schools (240 high schools, 280 middle schools, and 480 elementary); we have identified the specific characteristics that drive (impact) school WASL pass rates. Using a combination of iterative variance analysis and complex non-linear multiple regression, we isolated and quantified the mathematical relationships that exist between these drivers and school pass rates. These mathematical relationships are subsequently used to calculate the specific benchmarks for each school for each model. Using these common mathematical relationships for every school in the database puts dissimilar schools on the same (and therefore a fair) basis; in other words, turns apples and oranges into just apples.

Finding these mathematical relationships is often referred to as Data Mining. Data mining encompasses testing as many physical, demographic, and cost data sets as reasonably available to discover the true relationships that exist between them and performance; the data defines the relationships. This approach sometimes reveals unexpected relationships, and even expected relationships are occasionally found not to apply, or to apply in a different way.

Scientific Research into Fairly Measuring K–12 School Effectiveness, cont.

How To: Guide for Using Scientifically Based Report Cards to Improve School Performance

Uses and General Benefits of Benchmarking

Benchmarking can be used to:

- Set appropriate performance measures and develop realistic target areas for improvement;
- Develop a culture of continuous improvement and a willingness to learn from outside one's own organization;
- Increase the willingness to share solutions to common problems and build a consensus about what is needed to accommodate changes;
- Provide an effective 'wake-up call' and help to make a strong case for change;
- Improve quality and productivity;
- Identify examples of best practices from other high performing organizations in the public and private sectors; and
- **Finally, it is difficult to improve what cannot be accurately measured.**

The best performance improvements result from a model of "**Plan, Do, Check, and Adjust**." Extensive use of data is necessary to make valid attainable goals. Using these goals is the basis for specific "plans." Executing the plan is the "doing." Collecting data on the performance resulting from executing the plans provides the foundation for "checking" whether the performance improvements expected actually happen. Finally, using the results of the "checking" is the basis of whether or not "adjustments" are necessary in the original plan. Continuous improvement requires a continuous cycle of *plan, do, check, and adjust.*

Fair Report Cards for Schools (Scientifically Calculated Benchmarks)

- Scientifically based benchmarks show how any individual school is performing *based on its own student and school characteristics.* Scientifically based benchmarks have a quantifiable accuracy that can be tested and replicated by others. This allows for far more accurate benchmarks than those based on small subjective peer groups (whose benchmarks cannot be tested nor accuracy determined).
- Without scientifically based benchmarks, neither schools nor districts can actually know with adequate precision and provability whether any school is performing better, as expected, or worse *than it should.*
    - Some schools with low pass rates are performing well above their scientifically based benchmarks. These school's leadership and staff should be rewarded while continuing their efforts to further increase pass rates.
    - Some schools with relatively high pass rates are performing at or even below their scientifically based benchmarks. These schools should be called on to make necessary changes to improve their effectiveness.

Steps to Obtain Maximum Value from Fair Report Cards

Districts need to make a policy decision about their performance goals for their schools. That is, is their goal to be merely state average considering their specific characteristics, in the best 25%, in the best 10%? What grade do they want on their "report card," a C, a B, or an A? This is the context in which they should look at their school's results.

Districts also need to consider the benefits and risks associated with making the report card results public so that all parents are aware of the effectiveness of district schools.

1. Instruct district and school leadership on the scientific basis of the benchmarks provided in the reports to increase their buy-in to the validity of the results.
2. Use the actual pass rate minus scientifically determined benchmarks in addition to the multi-year trends for each school in the district to find the specific areas of success or areas needing improvement. This should be done separately for school average, reading, math, writing, and science. This may be a wake-up call.
3. Based on the findings from No. 2 above, prioritize where to put leadership attention within the school district and within the individual schools.
4. Investigate and determine what's going right overall in the scientifically determined best performing schools (by subject) so these "best practices" can be shared district wide. Also, learn from the underperforming schools what needs changing.
5. As part of district wide and school strategic planning, set quantifiable targets for each school by subject (relative to their own benchmarks) that are attainable in one, five, and ten years.
6. Using the trends from No. 2 above, assess whether specific changes to leadership, curriculum, training, etc. have resulted in the expected improvements to student pass rates. Make changes as needed based on these assessments.

Extra credit exercise:

Using the actual trends at each of the district's schools, quantify the benefits to math and science pass rates resulting from increases in school reading (writing) pass rates. This will help drive home to all district personnel the critical nature of reading to success in all subjects.

About the Authors

JAY MAIDMENT has consulted with and worked for utilities using operations research to analyze the economic behavior of electric utilities for the past 20 years. His research has expanded into public transit and K–12 education in recent years. He has a B.S. degree in Physics and Applied Mathematics.

JAMES LEWIS has been since 1997 the President of J Lewis & Associates, a firm providing cost and performance benchmarking services to the electric utility and transit industries. Currently, similar services are available to K–12 educators in Washington State. Previously, he was Director of the Bonneville Power Administration's Division of Acquired Generation. He has a B.S. degree in Electrical Engineering and an MBA.