The Dallas Value-Added Accountability System

William I. Webster
Robert L. Mendro
Dallas Public Schools

Introduction

A public school system, perhaps more than any other institution, is faced with the need to determine, fairly, which schools and school personnel are effective. An urban public school system contains sufficient diversity to make such a task very difficult. The Dallas Public Schools has felt the need for a fair method for determining effectiveness for several decades. Since 1992, as an outgrowth of a decade’s work, the district has determined effective schools using a fair and equitable value-added accountability system. The system has recently been expanded to include the identification of effective teachers and to shape teacher evaluation for the district.

In 1984, the Dallas Public Schools began an effective schools ranking system using multiple regression to develop longitudinal student growth curves on norm-referenced tests and determine effectiveness by the degree to which schools exceeded their students’ predicted growth (Webster and Otto, 1988). The system remained in effect for several years, but was abandoned when a new state accountability system was mandated. In 1998, the Board of Education established the Commission for Educational Excellence which, after extensive study, recommended the development of an accountability system that was fair, that was based on variables in addition to norm-referenced test data (while retaining an array of test scores as the primary measure of effectiveness), that could be extended downward to include measures of teacher
effectiveness, and that used similar methodology to the 1984 system. The current accountability system, using a combination of multiple regression and hierarchical linear modeling is an outgrowth of the Commission report.

The accountability system is multifaceted, tying together district and campus improvement planning, principal and teacher evaluation, and school and teacher effectiveness. As such, all district elements directly related to student learning, from teaching in the classroom through campus instructional leadership through district curriculum and instruction and staff development programs, are directly related to improvement on the same set of variables. At the district level, the system sets and measures progress on absolute continuous achievement goals. Internally, at the campus and classroom levels, relative effectiveness is identified, rewarded, and used as a model for improvement. Ineffectiveness is targeted for analysis, assistance, and, if all else fails over repeated attempts, change of personnel. The entire system is focused on continuous improvement. (For a discussion of the accountability system see Webster and Menaro, 1995.)

**Identifying Effective Schools**

**Effectiveness and Fairness.** School effectiveness must be defined fairly and must be measured in terms of relevant goals. In the public schools, fairness presents a formidable technical problem. The Dallas Public Schools accountability system controls for preexisting student differences in ethnicity, gender, language proficiency, and socioeconomic status, (hereafter termed fairness variables) and prior achievement levels. Additionally, the hierarchical linear model (HLM) employed controls for school level variables including mobility, crowding, percent minority, and socioeconomic status. The
political context, however, adds a dimension in which fairness must not only be controlled, it must be capable of being shown to be controlled.

The definition of effectiveness, in turn, must be collectively and cooperatively arrived at by all stakeholders in the system. In the Dallas system, this need is met by an Accountability Task Force. This group, composed of parents, teachers, principals, and community and business representatives, serves as the final authority concerning variable selection and weighting, formulating the rules of the accountability system and the performance awards associated with it, and hearing appeals of system decisions. While the technical demands of the system narrow the necessary tools that can be used to determine effectiveness, the Accountability Task Force makes the decisions about defining the content of the system and seeing that it is implemented fairly.

No matter what is said about the goals of education, the public defines effectiveness foremost in terms of how a school's student score or tests. The Dallas system includes a variety of important educational variables including criterion and norm-referenced test scores, student attendance rates, dropout rates, student retention rates, student enrollment in honors courses and advanced diploma plans, graduation rates, and percentages of students taking college entrance tests. The ultimate measures, however, come from test results. A school with improving achievement results is held to be on the right track. Only, then do the other variables come into play. No one perceives a school with falling test scores to be effective. For this reason, the Accountability Task Force weighs test scores more heavily than other variables.

The outcome variables included in the Dallas system as of 1995 are defined in Table 1. Student level variables consist of a number of individual test measures and
individual attendance. School level variables are determined across students for the appropriate population or subpopulation. The weights included in Table 1 are the relative weights assigned to each variable by the Accountability Task Force. The Task Force approves the variables and assigns weights to them each year.

Fairness is also defined in terms of holding schools accountable only for students who were enrolled in the school long enough for the school to impact their education as well as for all students eligible for the testing programs. In the first instance, only students who are continuously enrolled in a school are used in the analyses. A continuously enrolled student is defined as a student who enrolls in the school by the end of the first six-weeks of school and remains in the school until the end of the school year. In the second instance, schools are required to test at least 95% of their eligible student populations. Therefore, schools cannot gain an advantage by withholding students from testing.

The Statistical Models. Student outcome variables are analyzed with a two-stage model. The first stage employs multiple regression to control the effects of the fairness variables and the second, a two-level hierarchical linear model, controls the effects of prior achievement or attendance and the influence of variables aggregated across the campus. School-level outcome variables are analyzed with a simple multiple regression model using two prior years of data for a school on each variable.

The two-stage model for student-level outcomes is used for two reasons. First, by regressing against the fairness variables and their interactions, leveling of existing differences associated with the fairness variables is demonstrated by computing subgroup and marginal means at the end of the first stage. The political advantages of this simple
calculation cannot be overemphasized. (No one trusts statistical manipulations unless he or she can see their effects demonstrated. It is not enough to say you have controlled for the fairness variables.) Second, the HLM used in the second stage cannot control the complex set of fairness variables. Mathematically, the school data matrices are not invertable if all variables are entered in the HLM model. However, if fairness variables are removed in the first stage, HLM proves to be quite robust. (Webster, et. al., 1995; Mendro, et. al., 1995).

The first stage of the student-level analysis regresses a combination ethnicity/language proficiency variable (4 levels: black, Hispanic, Limited-English Proficient, and other), gender (2 levels), free-lunch status (2 levels), the first and second level interactions of these first three variables, census income, census poverty, and census college attendance (each defined at the census block level) against each outcome and predictor achievement or attendance variable. The equations are given by the standard multiple regression equations in Equation 1. An outcome or predictor achievement or attendance variable is represented by the $Y$, and the fairness variables and interactions are represented by the $X$. To control for homoscedasticity of residuals and for mean residuals departing from the regression line, the predictor space is divided into 256 equal intervals after the regression line is determined and residuals are standardized to a mean of 0 and a standard deviation of 1. After the regression stage, it requires simple calculation to show that the residuals for the 16 possible subgroups formed by the crossing of the ethnicity/language, gender, and free-lunch variables have been equalized.
Linear Regression Equations
\[ Y_i = \beta_0 + \beta_1 X_1 + \ldots + \beta_j X_j + r_i \]  
where  
\[ r_i \sim N(0, \sigma^2) \]  
Residuals from the first stage are then used in the HLM in the second stage of the student analysis. The general equations for the hierarchical analysis are given in Equation 2. Here the \( Y_i \) are residuals of student outcome variables from stage one, the \( X_i \) are residuals of prior achievement or attendance variables from stage one, the gammas represent school level variables including mobility, crowdedness, percent minority, percent black, percent Hispanic, percent on free lunch, and average census variables for each school. The HLM equations are solved using empirical Bayes estimation.

Hierarchical Linear Model Equations
\[ Y_i = \beta_{0i} + \beta_{1i} X_{1i} + \ldots + \beta_{ki} X_{ki} + r_i \]  
Level 2  
\[ \beta_{0i} = \gamma_{00} + \gamma_{01} W_1 + u_{0i} \]  
\[ \beta_{1i} = \gamma_{10} + \gamma_{11} W_1 + u_{1i} \]  
\[ \vdots \]  
\[ \beta_{ki} = \gamma_{k0} + \gamma_{k1} W_1 + u_{ki} \]  
where  
\[ r_i \sim N(0, \sigma^2) \]  
and  
\[ u_{ki} \sim N(0, \tau_{ki}) \] for all \( k \).

Two types of residuals are then used in the accountability system. The first is an empirical Bayes residual for each school which is used as the measure of a school's effectiveness after fairness variables and prior values of outcomes are controlled. The empirical Bayes residuals also have the advantage of incorporating shrinkage into their
estimates. The residuals assign more weight to school estimates based on larger samples and less weight to estimates based on smaller samples.

The second type is individual student residuals obtained by solving the HLM equations back to the residual components. These student residuals are used as the basis of the teacher effectiveness indices, which in turn form the basis of the teacher evaluation system. (Although, preliminary efforts to use the student residuals have encountered some problems which are discussed in the section below on technical issues.)

As noted, effectiveness on school-level outcome variables, e.g., dropout rates, is computed using simple multiple regression. Outcome values of the school variables are predicted using the values for the previous two years of the variable with schools as the unit of analysis. The standard equations represented in Equation 1 are used with the $Y_i$ representing a school value on the outcome variable and the two $X_i$ representing the value of the variable for the previous two years. Residuals from the regression then become the effectiveness scores on the school variables.

**School Effectiveness, Performance Awards, and Penalties.** The intent of the Commission for Educational Excellence in establishing school and teacher indices was the use of indices to reward effective schools and to provide extensive help and resources to train or retrain the staff of ineffective schools. In light of this, the Dallas Public Schools has awarded $2.4 million dollars to effective school staff members each of the years from 1992 to 1994 and has increased the amount to $3 million for 1995. Ineffective schools have had increased attention ranging from additional resources to replacing administrators to restructuring the schools.
Prior to 1995, approximately 20% of the staff members in the district received performance awards. Awards went to the top schools in each of the categories K-3, 4-6, K-6, 7-8, and 9-12 with the moneys prorated by the proportion of total staff in each category. The awards were set at $1,000 for each professional staff member and $500 for each support staff member in each winning school. Part-time personnel evaluated by the campus administrator received pro-rated awards. For 1995, the district will provide the staff members of any school that exceeds prediction with an award in a two-tier system where the top tier receives a full award ($1000/$500) and the second tier a smaller award ($480/$240). Schools must meet improvement thresholds to receive awards, i.e., their students must grow at least the growth rate of the national norm group (School Performance Improvement Awards, 1994-95). Given the achievement improvements noted in preliminary analyses of 1995 tests in the system, it is expected that approximately 50% of all district staff in schools will receive awards under the tier system.

No awards are planned for top teachers outside of the effective schools. The intent of the performance awards being made at the school level is to encourage cooperation and assistance within a school building. The Accountability Task Force rejected any plan where a teacher might be encouraged to withhold information or assistance from a fellow teacher in a school. Research on effective and ineffective schools in the district has supported this position to the extent that effective schools were observed to have higher senses of community and teamwork than their ineffective counterparts (Bearden, et al, 1995).
While no formal plans existed at first for increased levels of assistance and scrutiny of ineffective schools, plans were generated each year as a result of increased interest in these schools on the part of the Board of Education. Each year, the degree of assistance and the formality of a structure to provide such assistance has increased.

Identifying Effective Teachers

As planned by the Commission for Educational Excellence and the Board of Education, identifying students above and below prediction for the measurement of school effectiveness has implied the ability to sort information within a school and thereby identify effective teachers in that school. In 1994-95, the Dallas Public Schools prepared its first set of Teacher Effectiveness Indices for teachers at the elementary and middle school levels. These indices were produced for internal school planning purposes only. The trial of teacher indices led the Board to order the district to prepare and implement a system of teacher evaluation based on the indices. That system is described in the next section.

The Teacher Effectiveness Indices were prepared from the effectiveness data resulting from the School Effectiveness Indices program. Teachers of core courses relevant to the test data available (e.g., reading test data for students of teachers of language arts or social studies test data for students of teachers of social studies) were matched to the students to whom they had given grades and the School Effectiveness Indices student residual data were computed for each teacher. The effectiveness data had all been standardized to a mean of 50 and a standard deviation of 10 across the district which made interpretation of the indices relatively easy. Teachers with mean performance above 50 were above the district mean. Standard errors of the mean ranged
from 2 to 3 for most class sizes implying that a mean Teacher Effectiveness Index in the range of 2 to 3 points above or below 50 identified significantly higher or lower student testing gains.

The actual computations of indices immediately highlighted the strengths and weakness of the indices system. In terms of strengths, indices simplified immensely the fair attribution of effect to teachers, principals had an easily interpretable guide to the relative effectiveness of staff members, and test data were readily comprehensible in terms of progress of students. Among the weaknesses of the initial system were that not all teachers had indices, not all teachers with high indices were necessarily ideal teachers, attribution of effect in situations where several teachers instructed students in an area was impossible, not all principals were able to use indices to guide their efforts to work with staff members, and, errors in grade assignment and in the student database were rapidly highlighted by the indices.

Given the aforementioned strengths and weaknesses, Teacher Effectiveness Indices generally performed as they were designed and expected to. A fair measure of relative achievement was attributable to most teachers with indices. The design of indices with a T-score scale implying a set district mean across indices facilitated interpretation. As anticipated from the beginning, the indices had to be used with caution. Not all important outcomes of education are measured by tests (although many more are than the detractors of testing imagine). Indices have to be placed in a school context and evaluated with the same degree of perception as all other data about a classroom and a teacher. Once they were, they provided a valuable source of information in estimating an individual teacher's ability to influence student learning outcomes.
Despite several hours of training, many principals were unable to make effective use of the indices. This was partly due to an initial design which provided too much information and did not simplify it sufficiently. Indices for 1995 will eliminate unnecessary information and will provide, for each class, the index, the school index on the same variable, a listing of the students on whom the index is based, and the pretest, posttest and effectiveness scores for each student. For those who deal better with pictures, graphs showing student progress will also be included. Additionally, training on the use and interpretation of the indices will be expanded.

The main fact affecting principals’ use of Teacher Effectiveness Indices is that the indices do not include, nor were they intended to include, formal diagnostic information about either the students or the teacher. While they may contain clues, indeed sometimes strong clues, to this information in some instances, the connection is not formally there. The Teacher Effectiveness Indices provide an indicator that, for nearly all teachers and students, reliably indicates that the teacher is effective or ineffective in promulgating achievement gains with those students. From there, it generally requires focused inquiry on the part of the principal or some degree of self-directed examination on the part of the teacher to determine why effectiveness is present or absent. The degree to which a principal or a teacher can focus the inquiry or, in some instances, the degree to which resources are available to help a principal or teacher do so, is directly related to the utility of the indices.

Technical Issues. Since the School Effectiveness Indices are based on a two-level HLM analysis, it is logical to assume that Teacher Effectiveness Indices might lead themselves to a three-level HLM analysis. In this analysis, student data would comprise
the first level and be nested in classrooms matched to teachers at the second level which in turn would be nested in schools at the third level. (In this analysis, the school effects would have to be added back into the teacher effect or else nesting of teachers within schools would determine teacher effectiveness relative to the school as opposed to the district, which is what was intended.) While the two-level HLM analysis showed itself to be both robust and an improvement on single level multiple regression methods when used with residual scores (Wehter, et. al., 1995; Mendro, et. al., 1995), current research on three-level HLM analyses for use in Teacher Effectiveness Indices has not shown the method to be practical, or in many instances, computable on the scale required in this system.

Preliminary investigations indicate that when conditioning variables are introduced at the second and third level, the HLM analysis is not able to analyze all classroom level data arrays and some schools and teachers must be dropped from the analysis. Research into the utility of the three-level model, compared to a two-level model, in a context where second and third level conditioning variables are not used is ongoing. The most promising option seems to be using a two-level HLM for the Teacher Effectiveness Indices in which student residual scores from a fairness analysis form the first level and classrooms form the second level.

Until the two-level (student nested in classroom) HLM investigations for Teacher Effectiveness Indices are complete, the intention was to use student residuals from the two-level School Effectiveness Indices analyses (student nested in school) for the teacher Indices. The next technical issue was encountered in using these student residuals in program evaluation and teacher indices. The residuals from the HLM analysis are not
homooscedastic and do not have identical means as assumed in the model. The residuals show a small correlation (most often ranging between .12 and .17) between the initial pretest scores and the residuals. (In the previous multiple-regression model, residuals were standardized after the second stage as they are in the first stage.) Research is being conducted to determine how to best standardize these residuals before using them in analyses.

A third technical issue is that of appropriate sample size. The initial Teacher Effectiveness Indices, which were not intended as a component of teacher evaluation, were computed whenever 6 or more students were assigned to a classroom with a minimum of 6 students being used to mask individual student identities. Since the indices were originally intended as information for both the principal and teacher, and principals received information about appropriate size for making interpretations, this arrangement was sufficient. While the teacher evaluation system will make an operational definition about sufficient sample size for an index, the issue of how large a sample is needed for valid interpretation is still an important one with no readily available answers. One safeguard, however, it to design the system in a manner that requires observations over more than one year. A second is to apply a shrinkage estimate which weights each index by the degree of associated uncertainty is being applied (Sondors, 1995). The third option solves the problem directly. If the two-level HLM model (students nested in teachers) is employed, shrinkage will be built in through the use of the model.

**Using Teacher Effectiveness Indices in Teacher Evaluation**
The Dallas Public Schools is in the midst of constructing a teacher evaluation system based on the Teacher Effectiveness indices which will be field tested on the entire district in the 1995-96 school year. The Board of Education approved a framework for the system which will be used as the basis of the elements included in the field trial (Proposal, 1995; Webster, 1995). A committee of teachers, teacher organization representatives, principals, community members, administrators, and parents, denominated the Teacher Evaluation Task Force, has been assembled and has assigned subcommittees to develop the instruments and procedures necessary to construct a system from the framework. Again, attention has been paid to systematically soliciting input from stakeholders.

The framework addresses the practical aspects of using indices as a component of teacher evaluation. Many such aspects are addressed but there are several major issues which needed to be dealt with before the framework could be constructed. The first was the issue of timing. Teacher evaluation must be completed long before indices are available for a current year. Next was the necessity to address all teachers. Approximately 30 to 40% of teachers will not have indices. As noted earlier, sufficient sample size had to be addressed. The system had to be designed to reflect reliable estimates of effectiveness. Another issue was one of meshing with other district processes already in place, in that the system must complement district and school improvement planning and principal evaluation. Finally, the issue of concentration of resources had to be addressed. That is, since the emphasis on the entire system is continuous improvement, the teacher evaluation system had to consider the number of teachers for whom the district had sufficient resources to provide extensive help.
As indicated in the discussion of the Teacher Effectiveness Indices, one problem in using indices is determining what aspects of performance to improve and how to improve them once low performance is identified. Coupling that fact with the timing problem, i.e., teacher evaluation must be completed about the time most students take tests, the framework specified the use of the Teacher Effectiveness Indices as the basis of formative system of teacher evaluation. Teacher indices, computed and available by mid-September of a current school year and based on the performance of the teacher's students in the previous year, would be used to determine the level and extent of a 'teacher's evaluation in the current year. Using the indices to guide the following year's evaluation in an effort to discover and correct the reasons underlying the performance of ineffective teachers directly focused the system on improvement and dealt with the timing issue.

The issue of teachers without indices had to be addressed next. First-year teachers, teachers of non-core areas, teachers of supplemental or remedial courses, and core teachers without specific measures in the Teacher Effectiveness Indices must also be evaluated. The intent was to also tie this evaluation more closely to student performance. A second track was proposed in the framework for these teachers. Their evaluation is primarily dependent on how well they use information about student outcomes in the classroom, particularly tests, assignments and profiles of student progress. This served to tie their evaluation to the only readily available measure of student performance for them. The focus on student outcomes was referred to as the Basic Duties of the Teacher and is based on the work of Michael Scriven (1994). Table 2 summarizes the Basic Duties of the Teacher.
Part of the answer to the concern regarding sample sizes for indices was addressed by placing teacher evaluation on a three-year cycle. This allows for the collection of sufficient data for estimation of teacher effectiveness, for estimating the stability of the teacher’s performance, for the teacher and principal to examine teacher performance and, finally, for the teacher to make substantive adjustments in performance and show improvement. The use of shrinkage estimates is being investigated as a further safeguard.

The mesh with other district processes was both built into the framework for teacher evaluation and is part of the consequences of using the School Effectiveness Indices variables and outcomes as the underlying basis of district and campus improvement planning and principal evaluation. The school effectiveness process focuses on what the Accountability Task Force defines as the most important outcomes of schooling and a school’s performance on them. The District Improvement Plan is required by the Board to address the School Effectiveness Indices outcome variables and a school’s Campus Improvement Plan must address identified weaknesses on these variables. Principals are required to focus on the same weaknesses and to focus on planned assistance for teachers with low Teacher Effectiveness Indices. Finally, teachers, as a required part of teacher evaluation, must implement changes in their teaching behavior which will improve their student’s outcomes. This singleness of focus serves to unify the three processes.

Concentration of resources was addressed by using the Teacher Effectiveness Indices and the Basic Duties of the Teacher to divide teachers into three groups for evaluation purposes. The top tier of teachers, approximately 40%, will receive a classroom observation based on a recast Texas Teacher Appraisal System (TTAS) in
the first year of each cycle and this will be the extent of their evaluation provided the observation indicates no problems and the teachers’ indices remain above average. In essence, high indices or Basic Duties of the Teacher ratings provide their evaluation in years 2 and 3 of the cycle. Thus, the level of evaluation is reduced for the most effective teachers. Middle tier teachers, approximately 30%, receive a classroom observation in each year of the cycle provided their indices remain in the middle level. This amounts to the same level of evaluation that they currently receive. The bottom tier of teachers, approximately 10% (or 750 teachers in Dallas), receive extensive evaluation including classroom observation, Basic Duties of the Teacher, and extra assistance from staff development and curriculum and instruction personnel each of the three years in the cycle until such time as their students’ performance improves. Additionally, supplemental evaluation instruments and processes are available to both the teacher and principal during the cycle to use as needed. These supplemental indicators include a formal teacher portfolio, structured formative classroom observation, a peer review process, a student survey, a parent interview, and, where necessary, a content knowledge assessment. Thus, evaluation and help is greatly increased for the least effective teachers.

The system is summarized in Table 3. It shows the indices and non-indices groups, the 3 tiers, the concentration of resources, and the formative nature of the process. As expected, the system as being designed and developed by the Teacher Evaluation Task Force within the guidelines of the framework, is and will be much more detailed and richer than the summary in Table 3. However, the table shows the essence of the system as it will be field tested in 1995-96.
Future Considerations

Several future considerations spring from the student-performance based systems defined above. This summary has not addressed changes in testing and testing policy resulting from the implementation of these systems. An important part of that policy is the analysis and detection of cheating and the more fundamental question of teaching to the test or even the yet more fundamental question of teaching only to the objectives being tested. Nor has the distinction between relative and absolute systems been addressed.

The system includes comprehensive computerized cheating analyses that are done on a classroom-by-classroom basis. Unusual increases in achievement, or outliers, are analyzed and, where deemed necessary by the Accountability Task Force, investigated. Cases of apparent inflated test scores from previous years are also investigated. Where necessary, examination of test documents and retesting is done. If evidence of cheating exists, schools are disqualified from the awards program and personnel are dealt with through the district's personnel system.

The overall accountability system contains absolute goals and specifies them for campuses and the district. However, these goals are not currently empirically determined. The systems for School Effectiveness Indices, Teacher Effectiveness Indices and teacher evaluation all are designed to identify relative performance. A pressing need for future development of the system is to determine whether absolute minimums of performance can be and should be empirically established. Is there a point in the future where many teachers or schools will reach acceptable absolute levels of performance so that their relative performance is immaterial? What technical problems must be
overcome to identify satisfactory minimal performance? As a minimum, the most effective schools and teachers should be utilized to establish goals based on best practice.

The next major task in the development of this system is the establishment of meaningful criterion-referenced objectives for students and teachers. While it will always be enlightening and useful from a training perspective to determine the most effective schools and teachers, there may come a time, if the district continues to improve, when perhaps 90% of teachers and schools will deserve recognition. When that occurs, the methodology for recognizing them must be available.
References


Proposal For a Comprehensive Teacher Evaluation System (1995). Dallas Public Schools, Department of Institutional Research, Dallas, TX.


20
<table>
<thead>
<tr>
<th>Elementary Criteria</th>
<th>Weight</th>
<th>Middle School Criteria</th>
<th>Weight</th>
<th>High School Criteria</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITBS-Reading</td>
<td>4/Grade</td>
<td>ITBS-Reading</td>
<td>4/Grade</td>
<td>TAP-Reading Grade 9</td>
<td>6</td>
</tr>
<tr>
<td>ITBS-Mathematics</td>
<td>4/Grade</td>
<td>ITBS-Mathematics</td>
<td>4/Grade</td>
<td>TAP-Mathematics Grade 9</td>
<td>6</td>
</tr>
<tr>
<td>TAAS-Reading</td>
<td>5/Grade</td>
<td>TAAS-Reading</td>
<td>5/Grade</td>
<td>TAAS-Reading Grade 10</td>
<td>12</td>
</tr>
<tr>
<td>TAAS-Mathematics</td>
<td>4/Grade</td>
<td>TAAS-Mathematics</td>
<td>4/Grade</td>
<td>TAAS-Mathematics Grade 10</td>
<td>12</td>
</tr>
<tr>
<td>TAAS-Writing Grade 5</td>
<td>5</td>
<td>TAAS-Writing Grade 5</td>
<td>5</td>
<td>TAAS-Writing Grade 10</td>
<td>12</td>
</tr>
<tr>
<td>Attendance</td>
<td>1/Grade</td>
<td>TAAS-Science Grade 8</td>
<td>1</td>
<td>ACP-Language</td>
<td>2/Grade</td>
</tr>
<tr>
<td>Promotion Rate</td>
<td>1/School</td>
<td>TAAS-Soc. Studies Grade 8</td>
<td>1</td>
<td>ACP-Mathematics</td>
<td>2/Grade</td>
</tr>
<tr>
<td>Dropout Rate</td>
<td>1/School</td>
<td>ACP-Soc. Studies 2/Grade</td>
<td>2/Grade</td>
<td>ACP-Reading, Imp. Grade 9</td>
<td>2/Grad</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SAT-Verbal Grade 12</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SAT-Mathematics Grade 12</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PSAT-Verbal 1/School</td>
<td>1/School</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PSAT-Mathematics Grade 1/School</td>
<td>1/School</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Graduation Rate 5/5/School</td>
<td>5/5/School</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>% taking SAT 5/5/School</td>
<td>5/5/School</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>% in Hrs. Courses 4/4/School</td>
<td>4/School</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>% in Adv. Dipl. Plan - Grade 9</td>
<td>3/3/School</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>% in Adv. Dipl. Plan - Grade 10</td>
<td>2/2/School</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*ITBS = Iowa Tests of Basic Skills
TAAS = Texas Assessment of Academic Skills, a state mandated criterion-referenced test
ACP = Assessment of Course Performance, a series of approximately 150 end-of-course criterion-referenced tests.
<table>
<thead>
<tr>
<th>Teachers Of Elective Courses And Others Without Indices</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier I Top 40%**</td>
<td>TTAS</td>
<td>No additional evaluation</td>
<td>No additional evaluation</td>
</tr>
<tr>
<td>Tier II Middle 50%**</td>
<td>TTAS</td>
<td>Basic Teacher Duties*</td>
<td>TTAS</td>
</tr>
<tr>
<td>Tier III Bottom 10%**</td>
<td>TTAS</td>
<td>Basic Teacher Duties</td>
<td>TTAS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Supplemental Instruments</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Teachers With Indices</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier I Teacher Indices Top 40%</td>
<td>TTAS</td>
<td>Indices</td>
<td>Indices</td>
</tr>
<tr>
<td>Tier II Teacher Indices Middle 50%</td>
<td>TTAS</td>
<td>TTAS</td>
<td>TTAS</td>
</tr>
<tr>
<td>Tier III Teacher Indices Bottom 10%</td>
<td>TTAS</td>
<td>Basic Teacher Duties</td>
<td>TTAS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Supplemental Instruments</td>
<td></td>
</tr>
</tbody>
</table>

* Basic Teacher Duties include profiling, testing, student portfolios, and curricula/instructional design.

** Percentages are determined after the first administration of BDOT and TTAS by taking all scores, ranking them, and dividing into the three groups.