

# Using Student Achievement in Teacher Appraisal<sup>1</sup>

Karen Bembry  
Donna Bearden  
Robert Mendro  
Dallas Public Schools

## Abstract

*Recent State legislation requires that student achievement be a part of any teacher appraisal system in Texas. The Dallas community has also requested that student achievement, as the paramount responsibility of the school district, be tied to all professional appraisal. In compliance with these State and local mandates, the Dallas Public Schools spent a year developing and field testing a teacher appraisal system driven by student achievement data. This system incorporates both formal and informal measures of student achievement. Multiple sources of evidence of achievement are encouraged by the system with the intent that achievement data be used for classroom planning. This paper describes the process used in developing the system; the assumptions, goals, attributes and major components of the system; the types of student achievement data and methodology involved in disseminating information to teachers; and the results of the field test. A description of the political processes involved in producing the system is also included.*

## Development of a Teacher Appraisal System

### *Timeline*

The development of a more effective teacher appraisal system evolved from the district's efforts to measure school effects fairly and to hold schools accountable for student growth. In 1990-91, the Commission for Educational Excellence, a community task force, recommended an accountability system for schools and teachers based on student progress (Commission's Final Report, 1991). For the past five years, schools in Dallas have been evaluated on a set of variables, including student performance on standardized tests, using School Effectiveness Indices (Mendro and Webster, 1993).

Through a literature search conducted in 1993-94, staff members identified current practices and standards on teacher evaluation. A joint project to investigate

---

<sup>1</sup> Paper presented at the annual meeting of American Educational Research Association, Chicago, March 24-28, 1997.

various aspects of teacher evaluation was begun with Western Michigan University. Several types of assessment instruments were piloted for possible use in teacher evaluation. In addition, staff members investigated different methodology and procedures that would allow the School Effectiveness Indices to be disaggregated to the classroom level. In 1994-95, preliminary classroom indices were produced for teachers at grades 1-8.

Based on these preliminary studies, the Dallas Board of Education mandated the development of a new teacher appraisal system using student growth and outcomes as its basis. Concurrently, the Texas Legislature passed legislation mandating that teacher evaluation "criteria must be based on observable, job-related behavior, including: teachers' implementation of discipline management procedures; and the performance of teachers' students" (Senate Bill 1, 1995). This bill allows districts to develop a local teacher evaluation system acceptable to the State if that system is approved by a district-wide committee and local Board of Education and meets the above two conditions.

### *Task Force*

District staff developed the framework for a system that would comply to the new requirements and the Board of Education asked that a districtwide pilot test be conducted during the 1995-96 school year. The General Superintendent then appointed a Teacher Evaluation Task Force to further develop the proposed system. This 66-member Task Force became the vehicle through which teachers, principals, parents, community members and central office staff participated in developing the procedures, instruments, a manual for the field test, and a manual for the system once it was revised and adopted.

At the time the Task Force was created, the District was subdivided into ten areas. From teacher representatives in each of the ten areas, two representatives were chosen to sit on the Task Force, the intent being that at least 20 teachers participate as Task Force Members. Both elementary and secondary principals were also appointed to the Task Force. Teacher associations, including the Classroom Teachers of Dallas, the Alliance of Dallas Educators, and the Dallas Association of Texas Professional Educators, and the administrator organization, the Dallas School Administrators Association, designated voting members to the Task Force.

The Task Force began meeting in June 1995. Procedures were established, goals were clarified, and the task force was divided into subcommittees to develop parameters for the Classroom Effectiveness Indices and to develop system procedures. Task Force members logged over 663 hours of activity during the development of the system. The results of the Task Force were included in a teacher evaluation manual distributed to all teachers in the district for the field test year.

In 1995-96, all components of the proposed system were pilot-tested throughout the school district, with every campus invited to participate. Classroom Effectiveness Indices were produced for teachers at grades 1-12. A classroom improvement plan (the Instructional Improvement Plan) was piloted by teachers in the district. An instrument

for assessing teachers without classroom indices developed by the task force was also piloted in the schools. In compliance with Senate Bill 1, the system was presented to the District-wide Committee in April 1996. The Committee voted unanimously to recommend adoption of the new system in May and, with this recommendation, the system was approved by school board members in June. Throughout the year, the task force continued to meet and to solicit input from various stakeholders. After the field test, the system was revised. A revised manual was issued by the Teacher Evaluation Task Force in Summer 1996. A process chart is included at the end of this paper.

### *System Assumptions, Goals and Objectives*

In response to demands for better schools and improved student performance, the entire system of professional personnel appraisal in Dallas has been redefined so that student outcomes have become the driving force behind teacher appraisal, principal appraisal, and school evaluation (Webster, 1995; Webster et. al, 1996). Prior to this redefinition, personnel evaluation was unrelated to student performance. Teachers and principals were evaluated on the basis of what they did, not how their students performed. Under this new system, measures of student performance or growth are guides for determining both a teacher's and a principal's evaluation. The system reflects the belief that all campus staff want to and can improve both their own performance and student learning. Within this campus-wide focus for improvement, the system ensures that all teachers are challenged to improve their own teaching skills as well as to improve student performance.

Educational systems must be grounded in a set of commonly held assumptions. Assumptions for the new teacher appraisal system are:

- *All school personnel are committed to student academic success,*
- *Improved teaching leads to improved student achievement,*
- *Teachers want to enhance personal knowledge and skills for the benefit of their students, and*
- *Administrators want to assist teachers in improving professional knowledge and skills.*

In addition to these assumptions, a set of goals was developed to focus the system as it continues to evolve. These goals are:

- *to continually improve instruction using student outcomes as the system's basis,*
- *to tie professional development to identified student needs,*
- *to align the system with other evaluation/planning efforts in the district,*
- *to provide a method for all teachers to engage in self-reflection and professional development.*

## Appraisal System Overview and Procedures

The appraisal system emphasizes student academic success and promotes continuous student improvement. Because of the extreme differences in instructional settings within the district, the system was designed to adapt to varying instructional contexts. It promotes collaboration on evaluation between the teacher, or appraisee, and an administrator, or appraiser. Because appraisal activities are engaged in throughout the school year, it also promotes year-long appraisal. In addition, the system reflects the Learner-Centered Proficiencies adopted by the State of Texas in 1994; the proficiencies are learner-centered knowledge, learner-centered instruction, equity in excellence for all learners, learner-centered communication, and learner-centered professional development (Texas Education Agency, 1994).

The Teacher Appraisal System is based on the premise that student learning should drive teacher evaluation. To this end, the system requires all teachers to complete an Instructional Improvement Plan which identifies student learning needs, addresses student learning needs through specific strategies, and uses degree of implementation of the strategies as the teacher's appraisal. The plan is divided into three sections: *Needs*, *Concepts/Content/Strategies*, and *Final Evaluation*. (See format at the end of this paper.)

In the *Needs* section, the teacher lists identified needs for his or her students. The needs assessment is done in collaboration with the appraiser so that the professional expertise of both teacher and appraiser are brought to bear on the process. The appraisal system requires that, where needs of both prior students and current students exist, they both be included in the needs assessment. The former provides direct evidence of the success of the teacher in the past while the latter addresses particular needs of current students. Specifically, the teacher reviews his or her success in meeting the needs of previous students through the results of the Classroom Effectiveness Indices, the *Teacher Accomplishments and Performance (ATAP)*, previous performance standards reviews, or other valid methods. (The *ATAP* is described in the section on Practical applications of the Indices.) The teacher assesses needs of current students through testing histories of these students (reconstituted test data or test data from student files and records), districtwide pretests, classroom assessments, or other valid measures. (For teachers new to the system or in new assignments, current needs may be all that are available.) The appraiser reviews the needs assessment and, by approving it, certifies that it is inclusive of all major needs of the teacher's students. In general, it is expected that the more needs identified or the more severe the needs, the more extensive the remedies specified for the needs. Conversely, if there are few needs identified or the identified needs are not serious deficiencies, the remedies specified will be few. Thus, the level of need is directly related to the extent of a teacher's appraisal.

### *Assessment of current students*

A key element in the Teacher Appraisal System is the set of strategies developed cooperatively by the teacher and appraiser and listed in the second section of the form,

*Concepts/Content/Strategies.* These strategies are individualized as they are keyed to the particular needs of that teacher's students. Thus each teacher's appraisal is unique to the needs of the teacher's students; each teacher's plan is unique to the needs of the teacher and students. If the students' needs parallel needs of other students in the school or district, they may be addressed by strategies from the Campus Improvement Plan or from districtwide initiatives. During the development of the plan and during the course of the school year, the teacher or appraiser may request additional resources as necessary to assist the teacher in developing or implementing strategies.

As part of the strategies discussion, the appraiser and teacher determine the evidence needed to assess implementation of strategies. These are listed in the documentation section of the third column. The documentation of strategies serves as a guide for both the teacher and the appraiser and becomes the basis of as-needed formative discussions throughout the year. At the end of the year, it provides the basis for the appraiser's evaluation of the teacher. Both can determine whether the agreed-upon strategies have been implemented. Under Texas law, the system also requires an assessment of the teacher's implementation of a discipline management plan. If necessary, one or more of the strategies in a teacher's Instructional Improvement Plan could be specifically related to the District's Discipline Management Plan. In many cases, the item is simply used to indicate that a teacher follows school and district procedures concerning discipline. The Evaluation section is completed at the end of the year. In a conference with the appraisee, the appraiser notes the evidence of the teacher's implementation of strategies and evaluates the teacher as "meets expectation" or "less than expectation." The appraiser also indicates whether or not the discipline management plan was implemented.

The timeline of the system is as follows. In September, teachers use student results from the previous year and information from current students to identify instructional needs of their classes. From the needs assessment, the teacher develops strategies to address identified needs. A preconference is held between teacher and appraiser to negotiate the contents of the needs assessment and strategies sections of the teacher's Instructional Improvement Plan and to specify the types of evidence needed for the summative conference. The Plan is signed by both the appraiser and the appraisee. During the year, the appraiser assesses the implementation of the strategies as well as implementation of the school's discipline management system. Classroom observations and/or walkthroughs are conducted throughout the year. Any time an appraiser observes a behavior critical to the teacher's evaluation, he or she documents it in writing and provides a copy to the teacher. At the end of the year, a summative conference is held between teacher and appraiser to review documentation and to complete the final evaluation section of the Plan. Based on the year long assessment, on evidence noted by the appraiser during the year, and on evidence accumulated by the teacher, the teacher is evaluated as "meets expectation" or "less than expectation."

## Measuring Student Achievement

There are two types of outcomes relevant to needs assessment: those based on standardized outcome data from the Dallas Public Schools accountability system and produced as Classroom Effectiveness Indices and those based on unadjusted data. With unadjusted data, there is the possibility that other factors are confounded with student performance. In particular, socioeconomic status, language proficiency, and the prior level of student achievement may influence student outcomes and may bias a needs assessment in the direction of making good student outcomes look worse than they really are or making poor student outcomes look better than they are depending on the direction of the bias. Unadjusted outcomes can include test scores and analyses of performance on test objectives, *Teacher Accomplishments and Performance (ATAP)* results, measures against performance standards, classroom assessments, districtwide pretest information, and other data. Unadjusted measures do not give data which have been referenced against the performance of similar students, and caution must be used when examining them.

Adjusted student outcomes are represented by the Classroom Effectiveness Indices. These indices reflect student performance data appropriately matched to similar students' performance to control for the effects of major non-classroom influences. They provide objective measures of the performance of a teacher's students relative to similar students in the district. The Indices, based on indicators in the district's accountability system, are equated on the fairness variables: ethnicity, language proficiency, gender, and socioeconomic status, and on the previous achievement levels of students. Thus, they provide a fair and comparable measure of the achievement of the teacher's students. Where Indices exist for a teacher, they are included in the needs assessment, although the needs assessment is not limited to them.

The District is able to compute Classroom Effectiveness Indices for the following teachers:

<i>School Level</i>	<i>Teachers With CEIs</i>
Elementary	Language Arts and Mathematics
Middle School	Language Arts, Mathematics, 8th grade Social Studies, 8th grade Science
High Schools	Language Arts, Mathematics, all teachers who administer District end-of-course exams

### *Computing Classroom Effectiveness Indices*

Both the School Effectiveness Indices and Classroom Effectiveness Indices were computed in two stages using multiple regression-based procedures. In the first stage, outcome and possible predictor variables were regressed against the fairness variables

using multiple regression and adjusted for differences in these variables. This ensured that no school or teacher was rewarded or penalized for the type of students that attended the school or classroom. Predictor variables were chosen from variables available at the previous grade level. For example, grade 7 outcomes were predicted from grade 6 information.

In the first part of the second stage, each adjusted outcome variable was regressed against the combinations of possible adjusted predictors and the best predictors were selected based on maintaining a combination of 1) a high multiple correlation between the variables and 2) a large number of student scores. In the second part of the second stage, for the School Effectiveness Indices, the adjusted outcome and selected predictor variables were regressed against each other using a Hierarchical Linear Model (HLM) which simultaneously modeled the student data and 11 school characteristics. This guaranteed that no school had an advantage or disadvantage based on these school characteristics.

School effectiveness was determined for each outcome variable using Bayes Empirical Residuals from each HLM analysis for the school effectiveness scores. The effectiveness scores by variable were standardized and weighted, using Accountability Task Force weights, and combined with school-level variables to produce the School Effectiveness Indices. The Classroom Effectiveness Indices were then computed from student level residuals obtained from the HLM analysis after eliminating all students who did not meet the Teacher Evaluation Task Force criteria for Classroom Effectiveness Indices. The variables used at each stage are described in the following paragraphs.

#### *Fairness Variables*

In the First Stage, each outcome variable and each possible predictor variable were adjusted by the following student level fairness variables:

- ethnicity/language proficiency (with four levels: African-American, Hispanic, Limited English Proficient, and Other);
- gender;
- free/reduced lunch status (with two levels: lunch, no lunch);
- first and second order interactions between these three variables (16 possible combinations: ethnicity/language levels, gender levels, lunch levels);
- census-block level of family income by ethnic group of the student (a level of income for each census block);
- census-block level of family poverty by ethnic group of the student (a level of poverty for each census block); and,
- census-block level of college attendance by ethnic group of the student (a level of college attendance for each census block).

After the first stage, the adjusted means for the 16 combinations of the first three fairness variables were equal for every outcome and possible predictor variable.

### *Outcome and Predictor Variables*

In the first part of the second stage, outcome variables were regressed against possible predictors and predictors selected based on the size of the multiple correlation and the number of cases preserved. In the second part of the second stage, the outcome variables and the selected predictors were entered into the first level of a two level HLM analysis and 11 school level variables were used at the second level to condition the analysis and assure that the outcomes were uncorrelated with the school level variables. The school level variables used at the second level were: mobility, overcrowdedness, average income, average poverty level, average parental college attendance, percent on free lunch, percent LEP, percent African-American students, percent Hispanic students, percent minority students, and percent instructional days excluding vacancies and medical leave. The HLM analysis generated empirical Bayes residuals for each school on each variable. These residuals were standardized to a mean of 50 and standard deviation of 10 (a T-scale) and were combined across variables with the accountability task force weights.

Classroom Effectiveness Indices were determined by computing individual student residuals from the HLM analysis for each variable, filtering them through the conditions on Classroom Effectiveness Indices imposed by the Teacher Evaluation Task Force for minimum attendance and enrollment with each teacher, and calculating average residuals for each appropriate variable and class. Then, for each teacher, an overall combined Classroom Effectiveness Index was computed by averaging all student effectiveness scores for a teacher, adjusting for the number of students, and scaling to a mean of 50 and standard deviation of 2.5.

### *Classes without Classroom Effectiveness Indices.*

Where Classroom Effectiveness Indices do not exist, the needs assessment may include a teacher's previous performance on the ATAP, current student performance or work, the performance of previous students on measures not included in the district's accountability system, or other valid and relevant measures. (See sample at the end of the paper.)

## Teacher Appraisal System Field Test

### *Participants*

All campuses in the district were asked to participate in the field test, with one teacher on each campus randomly chosen from the various grade levels and content areas. The sample of field test teachers was stratified to be representative of ethnicity, gender, and subject area. One administrator from each campus was also asked to participate. Of all the campuses, 188 sent teachers to field test training. Teacher training was held on

Saturdays, with the teachers receiving an additional stipend for participating in the training. Principals were trained by cluster in regularly scheduled cluster training sessions. Of the 188 field test campuses, 123 teachers and 120 principals submitted final information at the end of the field test in May. In addition, three teachers and three administrators were interviewed for extended input concerning the new system.

*Procedures for the Teacher Appraisal Field Test:*

1. The responsibility for the collection of field test evidence was shared by the principal or his/her designee and the field test teacher:
2. The teacher and appraiser held a preconference in February to establish parameters for the evaluation and to discuss any information unique to the teacher's instructional setting or students.
3. Appraisers observed field test teachers in at least one walk-through or one formal visit to the teacher's classroom during March and April
4. Teachers prepared documentation that included the agreed upon evidence.
5. Both appraiser and teacher participated in a summative conference in May. At this conference, the final practice evaluation of the teacher was discussed, the ATAP instrument signed by both the appraiser and appraisee, and a copy of the completed instrument given to the appraisee.
6. All appraisal information was transferred to the designated form and sent to the Office of Institutional Research. All field test information remained confidential.

The field test included a principal survey, teacher survey, qualitative analysis of the Instructional Improvement Plans on randomly chosen campuses, a field test of the *Teacher Accomplishments and Performance*, teacher and principal interviews and discussion groups of district personnel. Each campus participating in the field test received a survey for the field test teacher and one for the participating administrator. Copies of the surveys are in the Appendix. Three components of the system were evaluated through the field test: the Classroom Effectiveness Indices, the Instructional Improvement Plan, and the *Teacher Accomplishments and Performance*.

The assessment of the Classroom Effectiveness Indices included the format in which the Indices were produced and sent to the campuses; the procedures for developing, distributing and correcting errors in the Indices; the technical issues concerning the Indices; and the practical applications of the information.

Format. The initial plan of the Teacher Evaluation Task Force called for composite indices scores for teachers and placement of teachers on tiers based on their composite scores. Tier designation determined the level of supervision and evaluation for

the teacher. Because the tier system was deemed to be too negative and in order to carry out a system that promoted continuous improvement for all teachers, the tier system was dropped at the end of the field test year. The original format of the Classroom Indices included the composite score and tier assignment, and each Index by variable and by class section. A list of students that comprised each Index was included, along with a previous level of achievement, current level of achievement in grade equivalents, and a graph indicating each student's growth. During the field test, teachers and principals recommended reorganizing the graph around 0 to indicate years of growth instead of charting starting and ending points. Other suggestions included renaming sections of the information for clarity.

Development, Distribution, and Correction Procedures. In order to disaggregate student achievement to the teacher level, the district utilized a second data base from the grade reporting system, since the connection between student and teacher was made through the grade reporting process. A number of errors in the second data base were discovered, including incorrect teacher social security numbers, incorrect teaching assignments, incorrect campus assignments, and instructional adjustments for individual students made at the campus level but not reported in the grading system. These errors comprised approximately 3% of the Indices reports generated. The field test production of the Classroom Effectiveness Indices allowed the district to identify and correct these data base difficulties. Institutional Research initiated a system of submitting any adjustments to the Classroom Effectiveness Indices for correction. With few minor changes in the request form, the field test procedures were adopted.

Technical Issues of the Effectiveness Indices. Classroom Effectiveness Indices need to be free from any bias that would influence a teacher's placement in the distribution of composite scores. The Effectiveness Indices were investigated for bias due to a teacher's ethnicity, gender, years of experience, or type of school. In addition, the Indices were investigated for bias due to the incoming ability of a teacher's students and the number of students comprising the composite score. Analyses indicated that the composite scores were free from bias due to ethnicity, gender, and the academic achievement of incoming students. In the years of experience analysis, first year teachers were found to have a lower composite CEI as a group. No bias was found due to the type of campus (year-round vs. traditional) on which a teacher is employed. A nonparametric correlation analysis between the number of student scores that comprise the CEIs and the Classroom Effectiveness Indices indicated a small negative correlation that was statistically significant. Shrinkage adjustment methods are being investigated which will potentially reduce this correlation.

Practical Applications of the Indices. Survey results indicate that the schools utilized Effectiveness Indices in three ways: analysis of classrooms and the campus, planning at the teacher and campus level, and communication with constituents. Teachers and principals reported identifying individual student needs as well as group needs. Planning included identifying special assistance to students, adjusting content covered and instructional strategies used, the use of team teaching, and identifying

appropriate staff development. Types of communication reported included using Indices as individual feedback for teachers during both formative and summative conferences.

A second component studied in the field test was the development and use of the Instructional Improvement Plan. Representative campuses were randomly chosen to participate in the Instructional Improvement Plan assessment. Results of the analysis indicate teachers need assistance in using student level data in the planning process. Instructional strategies listed were often vague or general, and the activities included were often not measurable. Documentation of student accomplishments was not an initial component of the plan, so field test analyses did not include indicators of student accomplishment. Information that would assist principals and teachers in writing clear, data-driven and measurable Instructional Improvement Plans was included in both the appraiser training and the teacher orientation video.

A third component in the field test was the development and refinement of the *Assessment of Teacher Accomplishments and Performance (ATAP)*, a classroom observation instrument designed for use with non-CEI teachers. Teachers and principals participating in the field test were surveyed concerning their perceptions of each item included in the *ATAP*. Overall, the thirty-four items in four domains were rated “relevant” to “very relevant.” In addition, principals and teachers were asked to use the instrument in an assessment activity. Each principal was to “rate” the field test teacher on that campus on each item. The distribution across the district would indicate whether or not the instrument distinguished among teachers of differing quality. Results from the field test indicate that although the greatest number of teachers were consistently rated a “3” on a three-point scale, there was a sufficient number of teachers rated “2” and “1” to indicate further investigation of the instrument scaling.

### *Recommendations*

Based on input received from various constituents, from the field test, and from negotiations with one of the teacher organizations, the following recommendations were adopted at the end of the field test year:

The emphasis on continuous student improvement remained the focus of the teacher appraisal system. Student achievement data drives the teacher evaluation system just as it drives the principal evaluation process, the campus improvement plan, and the district improvement plan. The alignment of all accountability systems district-wide was retained. However, how the system would utilize Classroom Effectiveness Indices for the teachers was redefined.

The initial framework for the teacher appraisal system included a composite Effectiveness Index for each teacher who administered an achievement test or end-of-course exam. This composite score was used to assign each teacher to one of three tiers. The initial purpose for establishing tiers was to focus the finite resources of the district where they were most needed, presumably in support of the bottom 10% of the teachers. However, the tier system was not well received, and it resulted in negative feelings from

teachers that interfered with the ultimate purpose of the system--continuous student improvement. The tier system was eliminated by the Task Force, and both the composite score for each teacher and the tier designation was dropped from the Effectiveness Indices format. It was recommended that additional investigation of the Classroom Effectiveness Indices be conducted during the 1996-97 school year.

The recommendation that the name of the Indices be changed from Teacher Effectiveness Indices to Classroom Effectiveness Indices was adopted. This name change made the process less intimidating to school staff members. Information received from teachers, administrators, and central office staff, indicates additional training on the practical and appropriate use of student and school data is needed.

Since the tier system was not retained and the *ATAP* was no longer needed to place all non-indices teachers on an evaluation tier, the Task Force voted that the instrument be retained as a supplemental, diagnostic instrument to be used as needed. Further assessment of the instrument scaling was recommended.

The development of concrete examples of Instructional Improvement Plans by teachers at all levels and for all subject areas was recommended. Both teacher and administrator input indicated additional training in developing the plans is also needed.

Finally, it was recommended that the Teacher Evaluation Task Force continue as the major vehicle for refining the appraisal system during the 1996-97 school year. The ultimate test of the system will be whether or not student achievement and performance improves.

### Implementation of the System

The Board of Education approved the new teacher appraisal system in June 1996 for implementation in the 1996-97 school year. Training for appraisers began in mid-July with additional training days in late July and early August. Training consisted of six hours of orientation and practice with the new system. Trainers included staff from Personnel who addressed legal personnel evaluation issues, Curriculum and Instruction who identified specific examples of Instructional Improvement Plans for various content areas and instructional situations, Staff Development who developed the training components, and Research and Evaluation who identified various data sources for teachers and administrators, and how to utilize data gathered. By the beginning of the 1996-97 school year, 557 appraisers were trained.

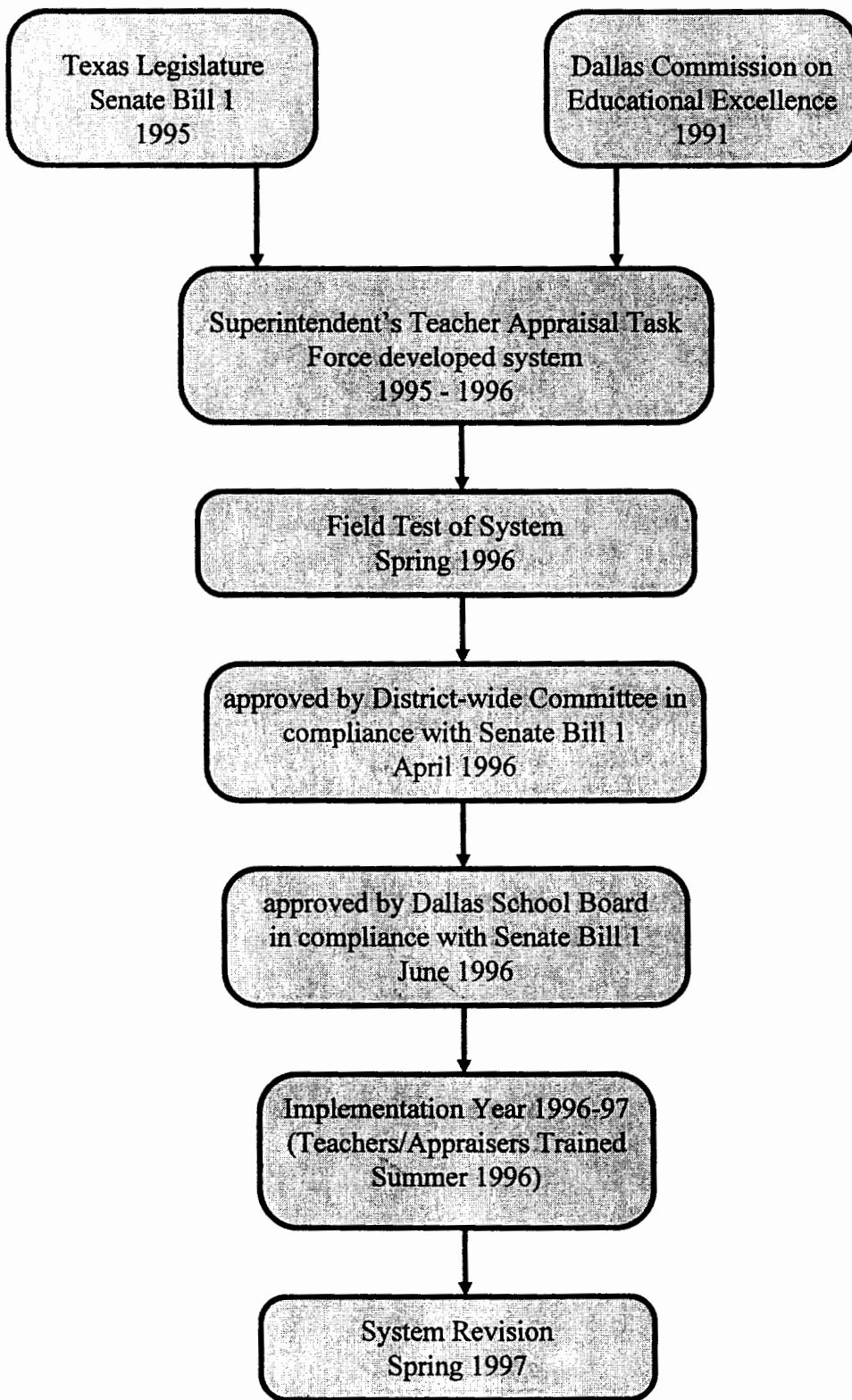
Information obtained from the field test indicated that both teachers and principals believed one of the most effective forms of training occurs at the campus level. Therefore, campus administrators were supplied with several resources to use to train their teachers. Enough copies of the *Revised Teacher Appraisal System Manual* were printed (11,000) for each teacher in the district to receive one. In addition, an hour-long orientation video was developed and a copy sent to each campus. Principals also received a packet of transparency masters to use in discussions with teachers. Examples

of Instructional Improvement Plans developed by content area specialists were also supplied to each campus.

It was recommended that the scheduled staff development day in August, called "job-alikes," be used to address the development of the Instructional Improvement Plan in the specific content areas. Cluster Offices also scheduled specific training for their teachers. In addition, teacher organizations scheduled training for both members and non-members.

### Impact of the New System

With the new teacher appraisal process in place, all appraisal and planning processes are now aligned (teacher appraisal, principal appraisal, and campus and district improvement planning). All use the same three-part format of assessing needs, devising strategies to meet needs, and evaluating implementation of strategies. All systems are based on student needs as identified in the same data. The diagnostic/prescriptive approach to evaluating teachers based on student growth could have an enormous impact on the district. Diagnosing student needs on a more consistent basis will give teachers the opportunity to adjust instruction, eliminate repetition and allow for individualization. The new appraisal system provides principals opportunities to support teachers in the instructional setting in very specific ways, strengthening the role of principal as instructional leader on the campus. It encourages Curriculum and Instruction staff members and support personnel to identify practical and effective instructional strategies that are successful in meeting very specific needs of local students. Staff development can also be focused at the individual campus and teacher level, reducing the possibility of spending staff development money without having a direct impact on student learning. Finally, the new teacher appraisal system is designed to allow all district personnel concrete ways to support and improve student achievement and learning.



# DALLAS INSTRUCTIONAL IMPROVEMENT PLAN FOR 1996-97

<p style="text-align: center;"><b>NEEDS</b></p> <p style="font-size: small;">(List needs and critical data documenting needs.)</p>	<p style="text-align: center;"><b>CONCEPTS / CONTENT / STRATEGIES</b></p> <p style="font-size: small;">(Itemize specific actions designed to address needs. Where extra resources are needed, specify these. List pertinent dates when essential to the strategy.)</p>	<p style="text-align: center;"><b>FINAL EVALUATION</b></p> <p style="font-size: small;">(Assess available proof that strategies were completed.)</p>
	<p style="text-align: center;"><b>DOCUMENTATION</b></p>	<p style="text-align: center;"><b>EVALUATION</b></p> <p style="text-align: center;">Discipline Management Plan Implemented _____</p> <p style="text-align: center;">Meets expectation _____</p> <p style="text-align: center;">Less than expectation _____</p>
<p style="text-align: center;">Signatures:</p> <p style="text-align: center;">Appraiser _____</p> <p style="text-align: center;">Appraisee _____</p> <p style="text-align: center;">Date _____</p>		

**Note:** Senate Bill 1, enacted in 1995, mandates that teacher appraisal information be confidential and not subject to the Open Records Act of Texas. Accordingly, all participants are cautioned to treat information regarding appraisals as confidential in nature.

Signatures (Final Evaluation):

Appraiser \_\_\_\_\_

Appraisee \_\_\_\_\_

Date \_\_\_\_\_

Campus \_\_\_\_\_

## DALLAS INSTRUCTIONAL IMPROVEMENT PLAN FOR 1996-97

<b>NEEDS</b> (List needs and critical data documenting needs.)	<b>CONCEPTS / CONTENT / STRATEGIES</b> (Itemize specific actions designed to address needs. Where extra resources are needed, specify these. List pertinent dates when essential to the strategy.)	<b>FINAL EVALUATION</b> (Assess available proof that strategies were completed.)
Potential Sources:  Classroom Effectiveness Indices  Diagnostic Skills Profile  Skills checklists  Teacher-made Tests  Portfolios/Student Work  Observations  Other Classroom Information  Other Formal/Informal Assessments	<div style="text-align: right; margin-bottom: 20px;"> <b>Signatures:</b>                      Appraiser _____                      Appraisee _____                      Date _____                 </div>	<div style="text-align: center; margin-bottom: 20px;"> <b>EVALUATION</b> </div> Discipline Management Plan Implemented _____ Meets expectation _____ Less than expectation _____

**Note:** Senate Bill 1, enacted in 1995, mandates that teacher appraisal information be confidential and not subject to the Open Records Act of Texas. Accordingly, all participants are cautioned to treat information regarding appraisals as confidential in nature.

## References

- Bloom, B. S., Madaus, G. F. & Hastings, J. T. (1981). *Evaluation to Improve Learning*. New York: McGraw-Hill.
- Bock, R. Darrel, (1989). *Multilevel Analysis of Educational Data*. San Diego, CA: Academic Press.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, CA: Sage Publications.
- Bryk, A. S., Raudenbush, S. W., Seltzer, M. & Congdon, R. (1988). *An Introduction to HLM: A Computer Program User's Guide* (2nd ed.) Chicago, Ill: University of Chicago.
- Commission for Educational Excellence: Final Report*. (1991). Board of Trustees, Dallas Independent School District, Dallas, TX.
- Joint Committee on Standards for Educational Evaluation. (1988). *The Personnel Evaluation Standards*. Newbury Park, CA: Sage.
- Mendro, R; L. Webster, W. J., Bembry, K. L. & Orsak, T. H. (1994). *An Application of Hierarchical Linear Modeling In Determining School Effectiveness*. Rocky Mountain Educational Research Association, Phoenix, AZ.
- Mendro, R. L., Webster, W. J. (1993). *Using School Effectiveness Indices to Identify and Reward Effective Schools*. Paper presented at the Rocky Mountain Research Association, Las Cruces, NM.
- Millman, J., & Darling-Hammond, L. (1990). *The New Handbook of Teacher Evaluation: Assessing Elementary and Secondary School Teachers*. Newbury Park, CA: Sage Publications.
- Proposal for a Comprehensive Teacher Evaluation System*. (1995). Board of Trustees, Dallas Public Schools, Dallas, Texas.
- Sanders, W; L., and Horn, S. P. (1995) The Tennessee Value-Added Assessment System (TVAAS): Mixed Model Methodology in Educational Assessment, in Shinkifield, A. J. and Stufflebeam, D. L., *Teacher Evaluation: Guide to Effective Practice*, Boston: Kluwer.
- Stufflebeam, D. L. (1992). What Fits Under the Label of Teacher Evaluation? *CREATE*, 1(2), 1-2.

- Wang, M. C., Haertel, G. D. & Walberg, H. J. (1995). *Effective Practices and Policies: An Empirical Foundation*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.
- Webster, W. J. (1995). The Connection Between Personnel Evaluation and School Evaluation. *Studies in Educational Evaluation*, 21, 227-254.
- Webster, W. J., Mendro, R. L. & Almaguer, T. (1994). Effectiveness Indices: A "Value added" Approach to Measuring School Effects. *Studies in Educational Evaluation*. 20, 113-145.
- Webster, W. J., Mendro, R. L. Bembry, K. L. & Orsak, T. H. (1995) *Alternative Methodologies for Identifying Effective Schools*. Distinguished Paper Session. American Educational Research Association, San Francisco, CA. ERIC EA 027 189.
- Webster, W. J., Mendro, R. L., Orsak, T. H. & Weerasinghe, D. (1996) *The Applicability of Selected Regression and Hierarchical Linear Models to the Estimation of School and Teacher Effects*. Paper presented at the National Council on Measurement in Education, New York, NY.
- Webster, W. J. & Olson, G. H. (1988). A Quantitative Procedure for the Identification of Effective Schools. *Journal of Experimental Education*, 56, 213-219.