

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Assessment of Teaching: Purposes, Practices,
and Implications for the Profession

Buros-Nebraska Series on Measurement and
Testing

1990

9. Teaching Assessment: The Administrator's Perspective

John R. Hoyle

Texas A&M University

Follow this and additional works at: <https://digitalcommons.unl.edu/burosassessteaching>



Part of the [Educational Administration and Supervision Commons](#), and the [Educational Assessment, Evaluation, and Research Commons](#)

Hoyle, John R., "9. Teaching Assessment: The Administrator's Perspective" (1990). *Assessment of Teaching: Purposes, Practices, and Implications for the Profession*. 11.
<https://digitalcommons.unl.edu/burosassessteaching/11>

This Article is brought to you for free and open access by the Buros-Nebraska Series on Measurement and Testing at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Assessment of Teaching: Purposes, Practices, and Implications for the Profession by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Teaching Assessment: The Administrator's Perspective

John R. Hoyle
Texas A&M University

School administrators, especially principals, are under great pressure to insure high levels of teacher competence. Because the school effectiveness research has demonstrated convincingly that effective schools begin with effective principals, Peterson and Finn (1985) drew a less than surprising conclusion by stating that "Practically never does one encounter a good school with a bad principal" (p. 42). A less pedantic east Texas superintendent put it this way, "Bad principals are like fish; you either can 'em or smell 'em for a long time." It is in the complex area of teaching assessment or teacher evaluation that principals draw the most criticism from classroom teachers and particularly from university pundits. As McLaughlin (1986), a longtime student of teacher evaluation, put it: "Teachers seldom respect principals as experts on classroom practice or as skilled classroom observers, and in the absence of principal credibility, teachers consider the evaluation an illegitimate comment on their performance and ignore the findings (p. 163). Teacher evaluation, in short, is an activity that most principals have little interest in or capacity to carry out" (p. 170). Epstein (1985) said that "Critics of current evaluation schemes complain that most are based on the principal's ratings on teach-

ers that result from infrequent (sometimes just one) observations in teachers' classrooms; on cronyism, patronage, or other prejudicial decisions; or on seniority, credentials, and accumulated credits that do not involve the evaluation of teaching skills" (p. 3).

Principals and teachers vary greatly in how they perceive the principal's performance as an evaluator, according to a survey of teachers and principals in Massachusetts (Tirrell, 1986). The respondents were asked to rate the role of the principal in evaluation according to their current perceptions and ideal expectations. Principals and teachers disagreed on 28 of 37 statements concerning current perceptions. They disagreed whether or not the principal

clearly communicates the philosophy of the evaluation program to the staff; clearly states the purpose of the evaluation in writing to the teachers; ensures that the teachers know and understand the caliber of their work; ensures that teachers are not threatened by evaluation practices; and encourages teachers to experiment with new behaviors designed to address weaknesses indicated in previous evaluations. (pp. 31, 32)

Other studies raise questions about the accuracy of measurement instruments and their criteria to distinguish the truly outstanding teacher from the average or even minimally competent one. Young (1986) identified five major faults in most observation instruments. They are as follows: (a) high inference items, (b) too many items, (c) judgments based on teacher actions, (d) low interrater reliability, and (e) lack of research support. Other research suggests that various groups disagree on the criteria they use to judge teachers. Epstein (1985) found that parents judge teachers on the basis of the degree to which the teacher communicates with the child's family, whereas principals give much less weight to this factor.

In attempting to determine whether people evaluate teaching excellence with the same criteria as they use to evaluate incompetence in teaching, Carey (1986) found that,

Unlike minimal competence ratings, it might be more difficult to achieve consensus in judgments of excellence in teaching. If this contention is supported in further research it may be that merit pay and mentor teacher plans suffer an Achilles heel that will be difficult to remediate. (p. 10)

The use of student scores on standardized achievement tests has become the major criterion used by some evaluators to judge

teacher competence. St. Louis, Missouri, teachers were told by the superintendent in 1985 that they would be rated unsatisfactory and lose their jobs unless their students reached specific levels of achievement or improvement on standardized achievement tests (Shanker, 1986, p. 3c). Other authorities, while urging evaluators to have multiple data sources for more accurate teacher evaluations, are calling for more testing to determine teacher effectiveness. According to Manatt (1986), evaluators are going to have to go,

deeper than inferences based on research on teaching. We want to look at student test data broken out by classrooms. . . . That way and only that way, can you really narrow it down to a teacher rather than saying in general that the school got these achievements for these boys and girls. (p. 12)

Most researchers and practicing administrators agree that the better teacher evaluation systems can discriminate good teachers from dreadful teachers, and adequate teachers from bad teachers. However, few knowledgeable educators believe that they can segregate the master or clearly outstanding teacher from the really good teacher. This fine line appears to be the source of much of the heat and criticism generated by teacher groups and researchers about the state-of-the art in teacher evaluation.

ADMINISTRATOR TRAINING

Graduate programs in educational administration generally require course work in staff personnel, program evaluation, and curriculum and instruction. However, few programs devote major portions of time to training in teacher evaluation. The hands-on training is left to the school districts or state departments of education after a person is appointed to a principalship. Teacher evaluation is merely one of the many knowledge and skill areas taught in graduate programs that causes critics to claim that the training is "too rigid and rule bound, on the one hand and too soft and ineffective on the other" (Peterson & Finn, 1985, p. 42). Similarly, Hoyle (1985) called attention to the shortcomings of many training programs by inferring that professors often advise students into fragmented individual courses with extreme content overlap or into courses that seem unrelated. Also, most inservice programs, institutes, and academies for administrator training make little pretense at systematic learning. The content is often fragmented,

“quick fix” information and makes little attempt at building a sequential accumulation of knowledge or skills. Hoyle also implied that some university preparation programs and training academies make better use of research evidence and examples of successful practice, but it is difficult to isolate preservice and in-service training factors from other socialization factors that determine successful administrative performance. The preparation puzzle makes it clear that principals in many instances do not have specialized knowledge of all the areas that they are expected to evaluate. The limits on their time and expertise and the haphazard way many are trained to evaluate teachers mean that principals face sizable odds in their efforts to distinguish the best from the rest.

PRESSURES FACING ADMINISTRATORS

It is obvious that many critics of principals and their training reveal considerable naivete about the increased demands, pressures, and paperwork brought in by state education reforms and demands for accountability. This naivete is most evident when evaluation reformers recommend that principals spend excessive amounts of time conducting classroom observations. Harried principals find these recommendations troubling and at times offensive. First, many principals dislike and distrust mandated procedures that give them sole control over teachers' salary increases and advances (Burke, 1982; Johnson, 1984). Second, the time it takes to evaluate each teacher according to best practice is enormous. The North Carolina State Department of Public Instruction (1986) studied this time problem and presented the following scenario. If, for example, a preobservation conference requires 30 minutes, an observation requires 60 minutes, a postobservation conference requires 45 minutes with an additional 45 minutes for the required data analysis, and the actual evaluation requires 60 minutes, the following formula will result:

Observation #1 (announced)	180
Observation #2 (unannounced)	150
Observation #3 (announced)	180
Evaluation:	<u>60</u>
	570 minutes

“If a school has 50 teachers and only one administrator who supervises teacher personnel, then about 60 days of 8 hours will be

required to complete all observations and evaluations" (North Carolina State Department, 1986, pp. 20, 21). Because these 60 days should properly fall within a 150-day time span, because observations early and late in the year will be impractical, the principal would have no time for any other management duty.

Less naive scholar critics are more attuned with the real world of principals and are aware of the many hats they wear during every school day, week, and year. Acheson (1986) recognized that not all principals can do all things and believes that it is time to redefine the roles of beleaguered principals and teachers. He believes that not all principals have the necessary range of managerial, human relations, and instructional leadership competencies to lead their schools in ways suggested by school effectiveness research and state evaluation reforms. In addition they do not have the time to perform all of these functions and roles in an exemplary manner. He acknowledged that when principals are surveyed, they list instructional leadership as their most important role, but other demands on their time relegate active leadership of the instructional program to a minor role.

Even if the principal devotes 570 minutes a year to each teacher, there is no clear evidence that it does any good. Pundits applaud the "clinical supervision" model which combines a democratically humane approach to supervision with a methodologically sound process (Cogan, 1973). The issue most raised is the question of the amount of time required to implement the process. What most pundits fail to realize is that 570 minutes to complete the clinical supervision cycle amounts to only one day in the life of a teacher. One day of the best instructional leadership displayed by the principal is hardly enough time to influence a teacher to improve instruction or increase effectiveness.

Lack of time is not the only problem. Most observers realize that the principal has the difficult task being both the evaluator of teachers and also a clinical supervisor in a collegial, constructively critical mode. Some principals are able to carry out both functions and are trusted and respected by the teachers, but most principals struggle with the evaluation versus supervision roles. Principals ask how they can evaluate teachers in order to make decisions about retention, promotion, tenure, and selection for career ladders and then turn around and work with them as a friendly critic or colleague to help develop the skills the teachers want and need to become better professionals. In some ways this is the same dilemma a parent faces when serving as both a loving counselor and a stern disciplinarian. It is indeed a delicate balance that few principals or parents perfect (Acheson, 1986).

Principals need love and respect like other professionals. They believe in family, democracy, mom, baseball, and apple pie. Most of them know that there is much rich empirical evidence derived from research in schools that indicates that schools that emphasize collaboration, good leadership, creativity, high expectations, clear goals, and open communication usually out perform schools that develop strict rules, create competition, have unclear goals, and conformity. The administrator is taught in graduate classes and workshops that the key to understanding and building an organizational culture is strengthening relationships and finding the concealed talents and inner motivations of people. The administrator is told to build trust and confidence and to create high morale in the teaching staff but is then required to pit one teacher against another by assessing their teaching effectiveness in order to determine their employment status or to dispense meager financial rewards. These small rewards, better known as merit pay or a rung on a career ladder, go only to a select few who are evaluated as "clearly outstanding" and, as a result, divisiveness builds and the "family" unity is threatened. Thus, the role conflict becomes a source of confusion for administrator and teachers.

English (1985) reported that members of an ASCD-appointed Task Force on Merit Pay and Career Ladders concluded that merit pay by itself:

1. Will not solve problems now facing schools in their efforts to reach higher levels of excellence.
2. Has shown to be ineffective and self-defeating and in fact may be a disincentive for improved performance.
3. Does not have a good track record in the private sector.
4. Represents a simplistic popular approach to the very complex problem of trying to recognize, motivate and utilize talent in schools. (p. 34)

The Task Force members also believe that the current emphasis on career ladders and merit pay contains paradoxical elements that lead to political confrontation rather than productive solutions. For example:

One view from within and without the profession indicates that since most of the profession indicates that it is impossible to pay all teachers a decent salary (because of economics and perceived public resistance to such cost increases), only some teachers should or can

be paid a respectable wage. This assumes that merit pay is a means to pay just a few teachers, preferably the best or superior ones, a competitive salary.

The opposing view is that the education profession lacks a fair and acceptable means to differentiate between teachers, given the state of present teacher evaluation systems. All attempts to differentiate are therefore considered unworkable and the result is a stalemate. (English, 1985, p. 34)

Principals then are required to use questionable carrot-and-stick methods to reward or punish teachers (punishment is not receiving merit pay or the next step on the career ladder) based on the state-of-the-art evaluation systems not suited to be used for both rewarding merit to the best teachers and improving all of the rest.

Even the casual observer can recognize the intense pressure on school principals to become "slave drivers" or "strawbosses" of teachers. Legislatures, corporations, and governors are pressuring school boards and superintendents to improve our failing school systems. The principal becomes the tool for the central administration to fix the school. The answer is to tighten the "Technical Core" (Peterson, Murphy, & Hallinger, 1987), sometimes called "teaching to the test." The principal oversees this convoluted search for excellence and is forced to tighten the lines of authority over teachers and grab control of the curriculum to drive up scores on state minimum-skills tests and standardized achievement tests. This model based on bureaucratic theory can neuter teacher creativity and initiative. As Frymier (1987), a long-term student of humane education and school climate, put it: "In the main, the bureaucratic structure of the workplace is more influential in determining what professionals do than are personal abilities, professional training, or previous experience. Therefore, change efforts should focus on the workplace, not on the teacher" (p. 10).

Most administrators acknowledge that a positive open workplace can promote a positive climate or "feeling" in a school district and in each school in the district. Administrators alone cannot create an open school climate. At best they can set the tone for their staffs to create an open climate. This tone may be described best as morale and work motivation for teachers and students. Positive morale and work motivation promotes an "ethos" that promotes higher achievement by teachers and students. Bureaucratic reform mentality, including misapplied teacher evaluation

systems, has little chance of improving schools. Again Frymier (1987) said,

there are many people in policy making roles and administrative positions who mouth pat phrases about the importance of teachers and teaching—and then proceed to undercut teachers by creating conditions of work that blunt their enthusiasm and stifle their creativity. (p. 9)

THE “ONE BEST MODEL” PROBLEM

The students I teach and the practicing administrators I know want good schools and happy teachers. School improvement is high on their professional and personnel agendas. They dislike the term *school reform* because it implies removing abuses or giving up sin or error. Administrators support school improvement and believe that almost all teachers desire to be productive and to be treated as professionals. This belief leads administrators to support a popular “One Best Model” teacher-evaluation system that rewards all teachers who reach mutually agreed upon professional growth goals. However, applying the One Best Model to all teachers and classrooms becomes problematic. Scores based on classroom observations of the teacher’s performance and student gains are only two of the ingredients to use in the “Mutual Benefit” model for judging professional competence. Peterson (1987) reported that the current practice of principal visits and reports alone, “does not promise to promote reforms for teachers or teacher educators” (p. 311). In fact Medley and Coker (1987) concluded that principals’ judgments have little to do with teachers’ effectiveness in promoting student achievement. They have discovered a number of common weaknesses in research design and instrumentation that cause the problem. One such weakness is caused by using a sample of teachers drawn from different schools in order to have enough teachers to allow a relationship to be detected. Therefore, judgments made by different principals in different schools could not be treated as interchangeable. Also, statistical procedures used to estimate the effectiveness of teachers violated important assumptions. Medley and Coker proposed a promising alternative called “measurement evaluation, which would base teacher evaluation on records of classroom performance made by observers trained to record behavior without evaluating it” (p. 140). They agreed that this alternative should be given further

study and development. Until educational researchers can find clearer links between teacher-evaluation systems using the highly acclaimed instructional models of Hunter (1986) and others and student-achievement gains, principals are going to balk at using narrowly defined criteria to evaluate teachers. Any evaluation system too narrowly defined and artificially implemented will be viewed by teachers as threatening and coercive and is against the professional administrator's better nature and training. They also know that a restrictive system will destroy teacher efficacy which Berman and McLaughlin (1977) called the single most powerful explanatory variable related to student performance.

The challenge facing policy makers and administrators is to make teacher-evaluation systems actually improve teaching performance and produce positive student outcomes. Until that happens, teachers and administrators will continue to complain that most systems do not distinguish between clearly outstanding and mediocre teaching. Moreover, the incompetent teacher remains. These complaints and other concerns about the motives behind the aforementioned evaluation systems have not turned administrators away from learning new skills to improve their supervisory roles. They realize the potential value of solid broad-based teacher evaluation. Lewis (1982) found that the overwhelming concern among administrators was how to convey that evaluations are for improvement, how to relate evaluation to learning improvement of students and how to develop a personal improvement plan for each teacher. Graduate programs in educational administration, state departments of education, and administrator in-service conducted by professional education associations are providing training to strengthen principals' supervisory, diagnostic, and prescriptive skills. This new emphasis on teacher evaluation trains principals and other staff to observe classroom practices, assess teacher solutions to classroom problems, and analyze the quality of the instructional processes. This training confirms the generally held belief about the conditional nature of teacher effectiveness and stresses individual teacher judgments within widely held categories for effective teaching (Hoyle, et al., 1985).

This phenomenon of the 1980s to retool school administrators to become instructional leaders has produced some predicted results in terms of new training and teacher evaluation procedures. Both the training methods and the evaluation procedures are strikingly similar across the United States. For instance almost all of the state and university training academies secure the services of the same consultants who bring in the same song with perhaps a little

different verse. They each stress the following features of a One Best Model successful evaluation system: (a) involvement of the teachers in the entire developmental evaluation process, (b) performance criteria based on sound research and on local needs and concerns, (c) collaborative goal setting, (d) multidimensional methods for assessing teacher's skills, (e) careful analysis of data gathered in the assessment stage, (f) development of specific job targets, and (g) inclusion of a preobservation conference to acquire background data and a postobservation conference to mutually analyze classroom data and set goals for improvement (Manatt, 1982). This "Mutual Benefit Evaluation" is an adaptation of the management-by-objective (MBO) model from business and is similar to models established by Redfern, Bolton, Manatt, and Hoyle (Hoyle, English, & Steff, 1985). Consultants to the academies and professors in graduate classes not only stress this Mutual Benefit Model but they employ many of the same teaching methods to help administrators improve skills in teacher evaluation and in the teaching process. Through simulations, role modeling, videotapes, and other devices, administrators are given extensive training in clinical observation, note taking, reporting, and conferencing skills. The participants then become mentors and coaches for other appraisers (McLaughlin, 1986).

This remarkable similarity in teacher evaluation training is seen by many as the One Best Model and has many advantages and some disadvantages. The advantages are as follows:

1. The terminology is similar, which improves communication about the process.
2. Involvement of the entire professional staff supports the time-honored notion of team work and organizational culture which, "embraces the norms that inform people what is acceptable and what is not, the dominant values that the organization cherishes above others and the basic assumption beliefs, rule and philosophy that guide the organization in dealing with its employees and its clients" (Owens 1987, pp. 29, 30).
3. The emphasis is placed on improving teachers rather than proving their incompetence.

The disadvantages of the One Best Model are as follows:

1. Excessive time is needed to conduct a thorough evaluation for each teacher if the staff is over 20 in number.

2. The process appears overwhelming to many teachers and they doubt if the system will make any difference in the way they teach.

3. Even after extensive training, appraisers remain inconsistent in assigning scores to teachers' classroom performance. This has been a troubling problem with the Texas Teacher Appraisal System when appraisers tried to determine the highest level or "exceptional quality" of a teacher's performance on each of the five domains. What is exceptional to one appraiser may be satisfactory to another.

According to Stiggins and Bridgeford (1984), who conducted case studies of teachers evaluation systems in four Pacific Northwest school districts, administrators had mixed feelings about the systems. In two districts, administrators were generally satisfied with the evaluation process, but concerned about the amount of time necessary to conduct observations. In the other two districts administrators were less satisfied. Reasons for the dissatisfaction included teachers' lack of trust in the evaluation process, the lack of clarity in the criteria, and the fact that the evaluation seemed more oriented to meeting state standards than promoting improvements. There was also disagreement about the impact of the evaluation system on teacher improvement and its link to staff development and in setting instructional priorities. Evaluation was, however, used by some administrators to help teachers identify individual goals and to specify a plan of action for the year. The completion of these plans and their effect on instruction was seldom monitored.

When asked how evaluation could be more directly related to the improvement of teaching, Stiggins and Bridgeford (1984) reported that the administrators recommended

changes in system management, including increased staff involvement in goal setting and emphasis in improvement as a district priority, improved methods of conducting observations, more time allowed for evaluation and observations, development of evaluators' skill, a stronger link between evaluation and staff development, and accountability for all principals conducting evaluations. (p. 21)

These suggestions for change parallel those concerns identified in most other national studies in teacher evaluation. Most administrators agree that evaluation could be much more effective in

diagnosing teachers' needs, improving their skills and improving student learning if changes are made in the process.

THE TEXAS TEACHER APPRAISAL SYSTEM (TTAS)

In interviews with principals and superintendents in Texas, the author has found both positive support for and calls for change in the Texas Teacher Appraisal System (TTAS, 1987).

Among the positive comments are the following

1. "(The System) standardized teacher evaluation in Texas."
2. "Many of the smaller resource-poor districts would have never instituted a system otherwise."
3. "Made the public more aware of the teacher's role."
4. "It reinforced the good teachers who were already doing these things."
5. "It has helped weak teachers to be aware of better techniques because most teachers really want to be good."
6. "Helped promote a common language about instruction and improving student learning."
7. "Increases the principals confidence to make suggestions about improving teaching effectiveness to bright professional teachers."
8. "The process helps you give more concrete suggestions to each teacher."

Some of the calls for change or negative comments were as follows:

1. "The career ladder was put in place before the teacher appraisal system was in working order."
2. "The State Board of Education and the Texas Education Agency keeps changing the rules in mid stream."
3. "There is not enough state money to pay teachers who have earned level three on the career ladder. If the state wants a merit system they should fund it."
4. "There is too much inconsistency in the appraisals from district to district in terms of the number of teachers who deserve to achieve level three of the career ladder."

5. "Other appraisers from the central office can cause problems because they frequently have no vested interest in the school and its culture, also no basis for building trust between the faculty and principal."
6. "We have only hearsay that the TTAS is improving our schools in Texas. The same districts have the same successes or failures."
7. "So far the system has had no effect on the number of students who drop out."
8. "Teachers put on a good show when they are being observed because the criteria are so specific and fairly easy to follow."
9. "It is unrealistic to pretend that I have the time to do each evaluation as the TTAS calls for."
10. "We really resent the career ladder because of the competition and divisions it causes."

The only research data available to measure the impact of the TTAS and the career ladder was gathered in 1986 after the first year of its implementation. The study gathered attitudes from teachers, principals, and superintendents about Texas School Reform and included several questions about the TTAS and career ladder (Ryon, et al., 1986). The results were not positive. "Eighty percent of the teachers, 78% of the principals and 78% of the superintendents said that the existence of the career ladder had negatively affected teachers morale" (p. 15). When asked, "All things considered, is the career ladder more of a plus or a minus?" 79% of the principals and 77% of the superintendents regarded it as a minus" (p. 18).

The researchers concluded that:

It would be difficult to arrive at any conclusion other than that most teachers, principals, and superintendents hold views of the career ladder that are largely negative, but they seem favorably disposed toward the increased emphasis on teacher evaluation. It is important to note that the respondents evaluated the career ladder and appraisal system as they knew them in April, 1986. It is difficult to say what their responses would be if the career ladder were somehow restructured or if it were better funded. Similarly, the disposition of teachers toward evaluation procedures could become either more positive or more negative with the introduction of the new statewide teacher appraisal criteria and procedures. (p. 17)

Individual "voices" from the field over 1 year later bring more encouraging news about the TTAS and career ladder. According to Bill Kirby (1987), the Texas Commissioner of Education, the TTAS has been modified and refined in 1987 to respond to the needs and recommendations from teachers and administrators. Some of the changes include a reduction in the number of teaching indicators within certain criteria, clarifications of the use of the "exceptional quality" rating and modifying the overall scoring procedures.

The TTAS is based on the clinical model discussed earlier and is similar to systems in several other states. The system was implemented on a statewide basis in the fall of 1986 after 13,000 appraisers were trained the previous summer. Also, standards were set for the rating of teacher performance for career-ladder decisions. This rather rapid implementation of a massive evaluation activity was fraught with glitches in the system. Teachers were upset because they were all treated alike in the process. Master teachers with 20 years of experience were placed at the starting line along with 1st-year teachers. They were all classified as Level 1 teachers and all were to be observed four times for 50 minutes whether they needed or wanted it. The majority of the master teachers were subsequently promoted to Level 2 within the year, but the morale damage had been done. The appraisers needed more and better training because most of them felt that they had been handed an ill-conceived tool by a politically inspired state-education bureaucracy. Several laws suits were brought by disgruntled teachers who were not promoted because they were not convinced that the most capable teachers were being rewarded. The Legislature, the Commissioner of Education, and the State Board of Education were pressed by teacher and administrator associations to fine tune the system in some areas and "overhaul" it in other areas. Based on these suggestions and advice from educator groups, refinements are being made to the system, and training updates and proficiency checks were conducted for all appraisers during the summer of 1987.

The assumption undergirding the TTAS (1987) includes the caveat that "The state of the art of teacher evaluation is not advanced to an operational level in some areas. Instead the system has been based upon existing classroom-based research on teaching, craft knowledge and experience" (p. 4). Because the appraisal process has been designed to include principles of sound evaluation to reflect the best current practice, and efforts are underway

to correct problems the system is slowly being accepted by more Texas educators.

ADMINISTRATORS MAKING TEACHER EVALUATION WORK

In spite of the general knowledge that teacher-evaluation systems to measure teacher effectiveness are based on conflicting classroom-based research on teaching, craft knowledge, and experience and that there exists little agreement between principals' judgments of teachers' effectiveness and the amount that students learn, optimism prevails among educators. University professors and others engaged in improving the state-of-the-art in the assessment of teaching are finding examples of exemplary programs throughout the United States. Many administrators are learning and using the best techniques and processes to improve teaching performance and school districts. Roueche and Baker (1987), authors of a research report on 154 "excellent" secondary schools in the nation, said that "quality in these schools is the function of the school principal backed by the superintendent and school board." Also, they reported that "of 500 teachers in the excellent schools the word most-often used to describe the motivational techniques of the principal was 'inspirational,'" and that "good school leaders spent hours in the classroom, inspecting what they expected" (p. 1).

Pigford (1987), a former principal, made teacher evaluation work for her to improve teacher performance by working with faculty to develop seven clear, specific, and measurable objectives for the school year. By creating a strong collaborative support system using the classical clinical supervision model, major victories were won. For example Pigford reported that "one of our school wide goals was to increase by 5% the number of students who passed the statewide basic skills test in reading. Since the pass rate had risen from 38% the previous year to 48% at the end of the current year, we knew that our students had far exceeded our goal" (p. 142). This report of successful evaluation to improve teaching and student growth is similar to many others with principals who create a climate for success and "inspect what they expect." Where administrators stand firm in their belief that clear instructional goals must be taught by inspired skilled teachers,

learning can take place. Not only does learning take place, but teacher evaluation is viewed as a valuable activity by teachers.

LEADERSHIP BY OBJECTIVE AND RESULTS LBO/R MODEL

Based on previous discussion about the unstable history of teacher-evaluation systems and the frayed thread holding principal observation and teacher effectiveness together this writer believes that administrators can use the following system to make teacher evaluation the key to teacher effectiveness and student learning. The system is a version of the One Best Model and is described elsewhere (Hoyle, et al., 1985). The system is called "The Leadership by Objectives and Results Model (LBO/R)" (see Figure 9.1).

The following four points should be considered in applying the LBO/R Model.

1. Each person to be evaluated meets with the evaluator in August. Together, a few specific areas of the job that relate to the goals of the system are selected. Teacher and evaluator agree on specific objectives for the teacher and on dates for classroom visits.

2. Evaluators concentrate on observable skills during classroom visitation. All new teachers and others viewed as needing assistance should have at least two different observers visit *at least three times for a 50-minute period*. Master teachers may need a formal evaluation every other year.

3. The teacher is given a copy of evaluator's comments at the performance follow-up conference following each visit. Both help write new objectives and growth plans. Both sign the evaluation form and indicate agreement or disagreement with the assessment.

4. The final evaluation conference informs the teacher of recommendations concerning employment, and new growth targets are mutually identified. The LBO/R Model should be used for instructional improvement and as a basis for dismissal of ineffective and marginal teachers. The model should be used to determine merit-pay or career-ladder status with much caution.

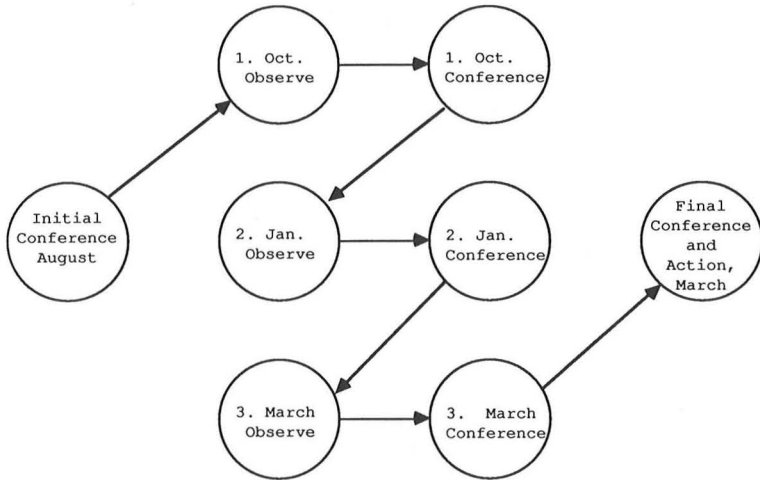


FIG. 9.1. LBO/R Teacher Evaluation Model

The LBO/R model contains the flexibility and the processes to help improve performance through improved supervision, helps plan for individual growth and development, provides information to identify marginal, average and outstanding performance, and identifies special teaching talents, creativity and skills. Note that the LBO/R model should primarily be used for instructional improvement and as a basis for dismissal of marginal teachers. It is not recommended that the LBO/R be used as the sole means to determine merit-pay or career-ladder status.

INDICATORS OF TEACHING EFFECTIVENESS

The LBO/R teacher-evaluation model can be very valuable for beginning teachers and teachers who need assistance. They should have at least two different observers who visit at least three times for a 50-minute period. The primary observer is the principal and the second could be an instructional supervisor, associate superintendent, assistant principal, or a retired teacher. Each observer must be trained to do the following things:

- Identify strengths and weaknesses of the teacher and provide assistance.

- Recognize “best practice” teaching performance.
- Use the vocabulary of staff evaluation and instructional management.
- Use motivation skills to inspire teachers to do their best.

Current research show that students are more attentive in classrooms that are businesslike and task oriented. A key to this businesslike classroom is not time-on-task alone but “academic learning time” (ALT). ALT is the amount of time students actually spend on an appropriate learning activity in which they are achieving at a high rate of success (90% or better). Researchers have found that in more effective schools teachers waste less time in starting and ending instructional activities and they select appropriate curricular materials that match the students abilities. Also, these teachers build high expectations for each learner and for themselves. Therefore, any teacher-evaluation form should include the following indicators:

- motivates students to achieve
- uses academic learning time effectively
- demonstrates proficiency in subject areas
- demonstrates command of the language
- promotes student academic growth
- learning objectives are clear
- learning strategies are based on objectives
- testing is based on objectives

Obviously, there are other important observable and nonobservable behaviors that contribute to the overall assessment of a teacher’s performance. However, if the aforementioned indicators are not measured or present then the other factors hold little value in determining a teacher’s ability.

The LBO/R or any other approach to teacher evaluation is only as effective as the administrators and teachers involved. If the administrator is protecting an image of total authority over his or her teachers, then the best evaluation model and instrument will be useless. Likewise, if the teacher feels that he or she needs no supervision, chooses to ignore school policy, and views the principal or any supervisor as the enemy, then any system regardless of its claims of a “mutual approach” is of little help. If the “mutual

approach" is taken seriously and becomes embedded into the cultural fabric of the school, then there is little doubt that administrators can make teacher evaluation work much more effectively.

CONCLUSIONS

Administrators in Texas and other states with comparable programs are trying to make the teacher-evaluation systems work to improve schools. Although there appears to be a sense of accomplishment about standardizing the process, reinforcing good teaching, improving weak teachers, promoting a common language about teaching, and improving the instructional leadership image of administrators, gnawing problems remain. The major problems appear to be the weak research base linking teacher-evaluation systems to improved student achievement and the inconsistencies in state-by-state and district-by-district applications of evaluation systems. Another problem appears to be the two-edged use of teacher evaluation. Many observers believe that removing the poorly conceived career ladders and merit pay from the appraisal process is the only answer to making the system operational to improve schools and schooling. Others have doubts about the validity of the One Best Model because it is heavily influenced by popular instructional models with specific inflexible steps that teachers and principals must follow to drive up student test scores. Although time constraints and expertise to help all teachers will remain as obstacles for principals, they want to increase their skills to help teachers teach and students to learn. Workable teacher-evaluation models are available if, and only if, the teachers and the evaluators view them as a positive process to achieve intrinsic rewards of professional growth. It should seem obvious to education policy makers that true professional educators need open, threat-free workplaces that nurture self expression and respect. Any teacher-evaluation process that restricts these rights will fail. Perhaps, time and the kind of disciplined inquiry engendered in this Buros-Nebraska Symposium will raise new questions and lead to better answers for administrators charged with insuring high levels of teacher performance. We must have the courage to try and try again—so much depends on our struggle.

In closing perhaps this story illustrates the kind of persistence we need to improve teacher evaluation. A little 9-year-old boy not endowed with much athletic ability was cut from a Little League

team for the second year in a row. In deep despair he went home and told his Dad about the great failure in his life. After a big hug and a popsicle the young lad decided to try again. He headed to the back yard with a ball and bat. He threw the ball in the air and swung—he missed the ball by at least a foot; he tried again and said “strike two” and again, “strike three.” Then without missing a breath he yelled, “Man, I’m a great pitcher.” So, man, we are great teacher evaluators. If we think we can, we will be.

REFERENCES

- Acheson, K. A. (1986). *It is time to share the responsibility for instructional leadership with others*. (ERIC Document Reproductive Service No. ED 267 510), 1–30.
- Berman, P., & McLaughlin, M. W. (1977). *Federal programs supporting educational change: Vol. VII Factors affecting implementation and continuation*. Santa Monica, CA.: The RAND Corporation.
- Burke, B. T. (1982). Merit pay for teachers: Round Valley may have the answer. *Phi Delta Kappan*, 64, 265–266.
- Carey, N. B. (1986, April). *Conceptions of excellence and incompetence in teaching: Implications for teacher evaluation*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.
- Cogan, M. L. (1973). *Clinical supervision*. Boston: Houghton Mifflin.
- English, F. W. (1985). Still searching for excellence: A preliminary report from the ASCD task force on merit pay and career ladders. *Educational Leadership*, 42, 34–35.
- Epstein, J. L. (1985). A question of merit: Principal’s and parents’ evaluations of teachers. *Educational Researcher*, 14, 3–10.
- Frymier, J. (1987). Bureaucracy and the neutering of teachers. *Phi Delta Kappan*, 69, 9–14.
- Hoyle, J. R. (1985). Programs in Educational Administration and the AASA preparation guidelines, *Educational Administration Quarterly*, 21, 71–93.
- Hoyle, J. R., English, F. W., & Steffy, B. E. (1985). *Skills for successful school teachers*. Arlington, VA: American Association of School Administrators.
- Hunter, M. (1986). Knowing teaching and supervising. *Using what we know about teaching*. Alexandria, VA: Association for Supervision and Curriculum Development, pp. 175–176.
- Johnson, S. M. (1984). *Merit pay plans: Effects on teachers and their work*. Paper presented at the 1984 National Conference on merit pay for teachers. American Center for Management Development, Sarasota.
- Kirby, B. (1987). Comments in the foreword of the 1987–1988 *Teacher orientation manual for the teacher appraisal system*. Austin, TX: Texas Education Agency.
- Lewis, A. C. (1982). *Evaluating educational performance*. Arlington, VA: American Association of School Administrators.
- Manatt, R. (1982, January). Teacher performance evaluation: Practiced application of research. Ames, IA: Iowa State University, 82–81.
- Manatt, R. (1986). How well can we truly evaluate teachers? An interview in *The School Administrator*, 43, 10–12.
- McLaughlin, M. W. (1986). Teacher evaluation and school improvement. In a

- Lieberman (Ed.), *Rethinking School Improvement* (pp. 162–175). New York: Columbia University Press.
- Medley, D., & Coker, H. (1987). How valid are principals' judgments of teacher effectiveness. *Phi Delta Kappan*, 69, 138–140.
- North Carolina State Department of Public Instruction. (1986). *Teacher performance appraisal system training: A report of outcomes*. Raleigh, North Carolina.
- Owens, R. (1987). *Organizational behavior in education* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Pigford, A. B. (1987). Teacher evaluation: More than a game that principals play. *Phi Delta Kappan*, 69, 141–142.
- Peterson, K. D. (1987). Teacher evaluation with multiple and variable lines of evidence. *American Educational Research Journal*, 24, 311–317.
- Peterson, K. D., & Finn, C. E. (1985). Principals, superintendents, and the administrator's art. *The Public Interest*, 79, 42–62.
- Peterson, K. D., Murphy, J., & Hallinger, P. (1987). Superintendents' perception of the control and coordination of the technical core in effective school districts. *Educational Administration Quarterly*, 23, 79–95.
- Roueche, J., & Baker, G. (1987, September). *Profiling excellence in America's schools*. Distinguished lecture presented at the Annual meeting of the Texas Association of School Administrators and the Texas Association of School Boards, San Antonio, TX.
- Ryon, J. H., Cross, R., Gonzales, M. A., Guerrero, T., & Pearson, D. (1986). *Voices from the schools: Texas view on public school reform*. Report sponsored by The Texas Coastal Bend Chapter of Phi Delta Kappa and Corpus Christi State University. Corpus Christi, Texas.
- Shanker, A. (1986, November 14). No way to grade teachers. *St. Louis Post Dispatch*, p. 36.
- Stiggins, R. J., & Bridgeford, N. J. (1984). *Performance assessment for teacher development*. Washington, DC: ERIC clearinghouse on teacher education. (ERIC Document Reproductive Service No. ED 242 717).
- Texas Teacher Appraisal System Teacher Orientation Manual. (1987). Austin, TX: Texas Education Agency.
- Tirrell, F. J. (1986). Reality lags behind ideal in teacher evaluation. *The School Administrator*, 43, 31–32.
- Young, R. (1986). Five reasons why your teacher evaluation form may be faulty. *The School Administrator*, 43, 31.

