1985

A Study of the Congruency between Teacher Evaluation Practices Preferred by Teachers and Teacher Evaluation Practices in Use by Elementary Principals

Kenneth E. Upshaw
Loyola University Chicago

A Study of the Congruency Between Teacher Evaluation Practices Preferred by Teachers, and Teacher Evaluation Practices in Use By Elementary Principals

by

Kenneth E. Upshaw

A Dissertation Submitted to the Faculty
Of the Graduate School of Loyola
University of Chicago in Partial Fulfillment
of the Requirements for the Degree
of Doctor of Education

December

1985

Kenneth E. Upshaw

Loyola University of Chicago

A STUDY OF THE CONGRUENCY BETWEEN TEACHER EVALUATION PRACTICES PREFERRED BY TEACHERS, AND TEACHER EVALUATION PRACTICES IN USE BY ELEMENTARY PRINCIPALS

Problem: The study explored whether teacher evaluation methods implemented by elementary principals are congruent with identified teacher-preferred practice.

Procedure: One hundred thirty-eight elementary principals were asked to indicate the frequency with which they include each of 28 teacher preferred evaluation characteristics in their teacher evaluation programs, and their opinion of the importance of each characteristic to an ideal evaluation system. Eighty-one principals responded.

Responses were analyzed by a Scale of Congruence. The differences between present and ideal practice were analyzed by a Multivariate Analysis of Variance. The effects of 6 intervening variables were analyzed using a Multiple Discriminant Analysis.

Results: 1. Suburban elementary principals, regardless of years as a teacher, a principal, or an educator, and regardless of the size of school, highest degree earned, or economic conditions of the district, tended to include evaluation procedures which are congruent with characteristics of evaluation preferred by teachers. Congruence was not consistent across all characteristics, and there were examples of very high and very low congruence.

2. Principals' concepts of ideal evaluation systems were significantly different than evaluation systems they reported operating. The difference

occured primarily among evaluation characteristics associated with a humanistic, collegial system. Moreover, principals' concepts of ideal evaluation systems were closer to teacher-preferred characteristics than the systems reportedly in operation.

Conclusions: 1. Principals tended to be implementing most of the high-priority evaluation characteristics supported by teachers' organizations.

2. Evaluation practices which produced the greatest congruence tended to be those over which principals usually have little control.

3. Principals tended to operate evaluation systems that were more principal-dominated than collegial in nature, although principals tended to believe that teacher evaluation should be more collegial than it is.

4. Principals indicated that they believe that teachers should be involved much more in the planning and implementation of evaluation systems than at present. That involvement should include developing policies as well as setting goals and designing programs.

5. Evaluation systems reported to be operated by principals in the survey tended to be summative, and not very diagnostic, and probably not focussed strongly on improvement of instruction.

ACKNOWLEDGEMENTS

To Dr. M. P. Heller, the Director of my dissertation committee, and a supportive and patient teacher, advisor and friend for many years, my sincere appreciation. To Dr. Martha Wynne, who patiently tutored me from concept through completion, at a cost of many hours which could have been invested elsewhere, my heartful thanks. To Dr. Phil Carlin, who graciously agreed to serve on the committee, and who expressed confidence when I most needed it, my gratitude. To Carol Hankus, who was determined that it be done, and done well, and who spent untold hours working with me on the manuscript, a special thanks. To Susan, who critiqued and edited, and sometimes propped me up, my love. And to Dr. Ted Riggen, my colleague, my mentor, and my friend, who knew I could and showed me how, a personal message: Thank you, and be well, my friend.

VITA

The author, Kenneth Edwin Upshaw, is the son of Edwin and Margaurette (Allen) Upshaw. He was born August 24, 1944, in Anderson, Indiana.

His elementary and secondary education was obtained in the public schools of Anderson, Indiana, where he graduated from Anderson High School in 1962.

In September, 1962, he entered Anderson College, in Anderson, Indiana, and in June, 1967, received the degree of Bachelor of Arts Summa Cum Laude, with a major in Speech and Drama.

In September, 1969, he entered Western Michigan University, and in June, 1970, received the degree of Master of Arts in Counseling.

He was employed as an English Teacher in Beattyville, Kentucky from 1967 to 1969, and as the Director of Guidance for Paw Paw High School in Paw Paw, Michigan, from 1970 to 1972. In 1972, he was employed as a Counselor at Evanston Township High School, Evanston, Illinois, a position he left in 1973 to becomw Principal at the Oak Therapeutic School, where he remained until 1978. From 1978 until the present, he has been a Program Director for South Metropolitan Association, in Flossmoor, Illinois.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

One of the greatest reservoirs for improvement of instruction exists in the competence of excellent teachers in every school building in this nation.[1]

Historical Perspective

Teacher evaluation has been the subject of intense scrutiny by educational researchers since before the turn of the century. Known by the pseudonym "teacher appraisal", teacher evaluation was studied by a record 60 researchers in 1930.[2] Into the 1960's, the activity continued at a rate of approximately 40 studies per year. By the early 1960's, the number of published researches on teacher competence already approached 2,000, and the pace has continued unabated into the 1980's. Teacher evaluation is truly a subject of long-standing interest to researchers.

Prior to 1930, much of the research was influenced by the "scientific supervision" movement. Consequently, much of the data generated from early investigations was aimed at yielding knowledge of optimum methods to be employed by teachers.

---

[1]Gerald Bryant and Frank Haack, "Appraisal: Peer-Centered and Administration-Centered," Educational Leadership (May 1977): 609.

[2]This and other information in this paragraph are taken from: Measuring Teacher Competence: Research Backgrounds and Current Practice, by the Committee on IOTA (n.p.: California Teachers Association, n.d.), p. 7.

[3]John D. McNeil, "A Scientific Approach to Supervision," in Supervision of Teaching, ed. Thomas J. Sergiovanni (Alexandria, Va.: Association for Supervision and Curriculum Development, 1982), p. 19.

By the 1950's, however, a number of researchers had begun noting the characteristics of teachers and relating those characteristics to measures of teacher effectiveness.[1] Studies attempted to correlate such characteristics as "businesslike", "reactive", "tolerant", "positive character", and with a straight face, one hopes, the rather ambiguous "Bohemian character" with effective teaching.[2] The research activity reviewed by Barr, Eustice and Noe in 1955 led to several rather optimistic conclusions:

> The amount of reported research relative to the measurement and prediction of teacher efficiency seems to be on the increase. The research studies reported appear somewhat more sophisticated than those of a decade or so ago. There is much more awareness of the importance of criteria than a decade ago. There is much interest in student evaluation of teachers. The search continues for a single generalized pattern of qualities or behaviors that characterize good teachers (emphasis added).[3]

What seemed to many to be a promising area of research proved to be, however, a disappointment. The search for measurable indicators of teacher competence began to take on the appearance of an elusive dream. Since, through the years, most educators have had a pretty good idea of what good teaching is, quantifying teacher competence always seemed to be a goal not out of reach. As more and more research came up either empty-handed or with conflicting results, however, many educators began believing that teacher competence simply could not be measured.[4]

---

[1] Robert S. Soar, "Measures of Quality in the Classroom," in Merit Pay and Evaluation, Hot Topics Series 1983-84 (Bloomington, Ind.: Phi Delta Kappa Center on Evaluation Development and Research, 1983), p.109.

[2] Arvil S. Barr, David E. Eustice, and Edward J. Noe, "The Measurement and Prediction of Teacher Efficiency," Review of Educational Research 25 (June 1955): 262

[3] Ibid., p.266.

[4] Measuring Teacher Competence, p.7.

As a consequence, researchers of the 1960's began shifting their focus from difficult-to-measure personality characteristics to the identification of "low-inference" behaviors (e.g. verbal interactions; direct vs. indirect teaching), for which they sought correlations relating specific teaching behaviors with differences in pupil achievement or attitude.[1] With the advent of the 1970's came a demand for accountability and a return to basic educational achievement by students, and with it a renewed interest in research on teacher evaluation and teacher effectiveness. This interest led to a search for new processes and procedures, as well as criteria for teacher effectiveness. The quest for accountability resulted in borrowing from business and industry a "management by objectives" approach to evaluation. As some educators began recognizing improvement of instruction in the classroom as the primary purpose for teacher evaluation, words like "collegiality" and "participatory" began creeping into the literature, and "clinical supervision" rose to a prominent place in the field.

Three factors seem to have led to a resurgence of interest in teacher evaluation in the mid-1970's: (1) Publicity and consequent concern about teacher competence led to a painful awareness by administrators of the inadequacy of current evaluation systems. (2) Teacher unions and professional associations were increasingly interested in assuring that teacher evaluation was fair and equitable. (3) Research on teacher effectiveness pointed out the importance of certain teacher competencies in the production of learning outcomes.[2] This research, in turn, has led

---

[1] Philip L. Hosford and Jeanette V. Martin, "Historical Analysis of Videotape Usage in Predicting Teacher Effectiveness," College of Education Dialogue Series (Las Cruces, New Mexico: New Mexico State University, 1980), p.11.

[2] Freda M. Holley and Randall C. Hickman, "Research on Teacher Evaluation: Needs and Relaities," paper presented at the annual meeting of the

recently to the development of empirically based instruments for the evaluation of teachers and teaching, an example of which is the Carolina Teaching Effectiveness Rating Scale, developed at the University of North Carolina at Chapel Hill for the evaluation of provisional teachers.[1] A number of states have developed similar instruments, based on "state-of-the-art" practice.

But how well has all of this activity filtered into the classroom? As the instruments for evaluation have become more empirically based, and the processes for teacher evaluation more sophisticated (e.g. clinical supervision; objectives-based evaluation), one would expect significant changes to have occurred in the evaluation of teachers in most of the schools across the country. Such, however, seems not to have been the case.

Statement of the Problem

In 1952, Barr analyzed 39 research studies, and concluded, "No one appears to have developed a satisfactory working plan or system that can be used by personnel officers who must make judgments about teacher effectiveness."[2] Twenty years later, Lewis made the following observation:

> Administrators continue the semi-annual ritual of writing narrative reports and/or checklist evaluations on teachers. These "evaluation" devices generally not only fail to measure adequately professional competence, but also actually result in alienating the relationship between the teacher and the administrator (emphasis added).[3]

---

[1] Group for the Study of Effective Teaching, "Teaching effectiveness Evaluation Project: Final Report", (Chapel Hill, N.C., School of Education, Univ. of North Carolina at Chapel Hill, 1983), p.122.

[2] Arvil S. Barr, "The Measurement and Prediction of Teacher Efficiency," Review of Educational Research 22 (Madison, Wisconsin: Dunbar Publications, 1967): 171.

[3] James Lewis, Jr., Appraising Teacher Performance (West Nyack, N.Y.: Parker Publishing Co., 1973), p.11.

In 1977, Eckard and McElhinney published an eloquent plea to us all: "The complex and powerful tasks of teacher evaluation and accountability in education must be given increased attention by all educators." [1] But in 1984, Heuss discovered that most superintendents still did not require the practices and procedures of teacher evaluation recommended by authorities. [2]

Obviously, in spite of a long, albeit inconclusive history of research, teacher evaluation remains, in the eyes of many writers, an ineffective, even potentially damaging exercise in most school systems. Lewis writes:

> The present method of appraising the performance of educators in most schools in America appears to be dysfunctional and serves no useful purpose. Not only does it fall short of assessing adequately "true" performance; it also makes it impossible to take corrective action for professional growth, improvement and development. Furthermore, it has been a device used over the years to perpetuate the division between teachers and administrators. [3]

McNeil brands much current practice as "punishing and controlling", producing unproductive levels of anxiety. [4] Harris complains that much promising current study is handicapped, not only by the complexity of the phenomena, but also by the "old traditions of teacher evaluation as summative ritual, which keep alive the threat of dismissal while corrupting efforts at the improvement of teaching practice". [5] It would seem that the

---

[1] Pamela J. Eckard and James H. McElhinney, "Teacher Evaluation and Educational Accountability," Educational Leadership (May 1977): 618.

[2] Ronald Neal Heuss, "Teacher Evaluation Purposes, Procedures and Instruments in Texas Public Schools," (Ed.D. Dissertation, Baylor University, 1984), abstract.

[3] Lewis, Appraising Teacher Performance, p.13.

[4] John D. McNeil, "Politics of Teacher Evaluation," in Handbook of Teacher Evaluation, ed. Jason Millman (Beverly Hills, Calif: Sage Publications, 1981), p.280.

[5] Ben M. Harris, Teacher Evaluation As A Developmental System (ERIC Document Reproduction Service, ED224 800), p.2.

more things have changed in the area of teacher evaluation, the more they have remained the same.

And yet there are encouraging signs on the horizon. The Educational Research Service (ERS) found, in a 1969 survey of school systems educating 16,000 or more students, that only 17 of 235 school systems responding were without some kind of formal evaluation procedures.[1] More important, more than half of the systems were involving teachers in formulating an evaluation system. In 1974, the National School Public Relations Association (NSPRA) noted some definite positive trends in teacher evaluation, including teacher involvement in the development of instruments and procedures, evaluation which focussed on instructional improvement, and the use of more sophisticated supervisory techniques and instruments.[2]

A recent line of research has focussed on the attitudes and preferences of teachers themselves regarding teacher evaluation. The results of that research indicate that many of the trends noted by the NSPRA find favor with teachers and teachers' organizations. Specifically, teachers desire direct input into evaluation decisions, including the design of the instrument, the goals, the policies and the implementation. Teachers want evaluation to be a collegial exercise designed to improve their classroom instructional methods and to provide longer-range job targets. They want the evaluator to spend significant time in this process, but only if they can trust the evaluator's expertise, both in the criteria being evaluated

---

[1]National Education Association, "Evaluation of Teaching Competence," in NEA Research Bulletin (Washington, D.C.: National Education Assoc. 1969).

[2]National School Public Relations Association, Evaluating Teachers for Professional Growth, Current Trends in School Policies and Programs (Arlington, Va.: National School Public Relations Assoc. 1974).

and in the process being used to evaluate. Unfortunately, however, too many teachers still believe that principals are not qualified to evaluate.[1]

All of this creates a dilemma. The combination of high levels of research activity and the translation of that activity into practice has provided educators with increasingly sophisticated systems and empirically-based criteria for the evaluation of teachers. Additional research indicates that teachers have very specific ideas regarding how evaluation should be done, most of which is compatible with--even supportive of--the current best thought in the field. Yet, in spite of emerging trends to the contrary, the majority of school systems still seem to ignore best practice--and with it teachers' expressed preferences--and continue to implement teacher evaluation procedures which are ineffective and--what is worse--damaging.

Purpose of the Study

Although the disparity between best practice and current practice in the majority of school systems is well documented, this author found no study which explored whether the teacher evaluation methods implemented by principals are congruent with identified teacher-preferred practice. The present study attempted such an exploration.

---

[1]Joseph Leese, "Teacher Assessment and Consensus Preference," NASSP Bulletin 65 (November 1981): 26

Significance of the Study

> People move truer and more certainly toward excellence to the extent that they clarify their purposes, measure the impact of their action, judge it, and move on--in a few words, evaluate their progress.[1]

Most educators agree that a primary goal of teacher evaluation is the improvement of instruction. Teachers and administrators seems to concur on that point. A 1969 National Education Association (NEA) survey found that 93 percent of the teachers responding thought that the purpose of teacher evaluation should be to improve teaching competence,[2] while 98.9 percent of elementary principals surveyed in another study cited improvement of instruction as a purpose of teacher evaluation.[3] If that goal is to be realized, it follows logically that the participants in the process of teacher evaluation--the teacher and the supervisor--must, in some manner, agree on the goals, the procedures, and the criteria by which the teacher is to be evaluated.

But more than that, a number of sources indicate that certain intangible characteristics of the evaluation process are essential to successful evaluation and improvement of instructional competencies. Most of those characteristics involve the relationship between the participants. Teachers want the supervisor to show concern for them as a person, to exude warmth, respect, friendship, and honesty in a constructive, non-threatening atmosphere. They see the best evaluation as a collegial, cooperative effort

---

[1] Gene Glass, "A Paradox about Excellence of Schools and the People in Them," The Educational Researcher (March 1975): 9.

[2] National Education Association, p.70.

[3] Barbara Jean Swensen Hauge, "A Study of Teacher Evaluation Practices and Perceived Attitudes of Those Practices by Elementary School Principals in the Twin Cities Metropolitan Area" (Ed.D dissertation, University of Minnesota, 1981), abstract.

between the participants, in which the supervisor takes an active role. Most important, teachers want to trust the expertise of the evaluator.

Even under the best of circumstances, however, an effective appraisal and evaluation system may be unrealistic. [1] Sapone, for example, found little agreement between supervisors and teachers regarding the relative importance of the components of an evaluation system. Nevertheless, the closer the congruity between teacher-preferred practices and the performance of principals, the better the chances that teacher evaluation will work effectively to fulfill its stated goals of improved classroom instruction. This study sought to determine whether such congruity exists.

Definition

For the purposes of this study, teacher evaluation will be defined as "...the judgment by one or more educators, usually the immediate supervisor, of the manner in which another educator has been fulfilling his professional responsibilities to the school district over a specified period of time." [2]

Subjects

The subjects for this study were a census of elementary principals from 46 elementary districts located in the south suburbs of Cook

---

[1] Carmelo V. Sapone, "Appraisal and Evaluation Systems: Perceptions of Administrators, Teachers," NASSP Bulletin 65 (February 1981): 25-30.

[2] James Lewis, Jr., Appraising Teacher Performance (West Nyack N.Y.: Parker Publishing Co., 1973), p.23.

County, Illinois. All of these districts are members of a regional special education cooperative. Past experience predicted a high response rate from these principals to mailings sent under the letterhead of the cooperative.

Development of the Survey Instrument

A review of the literature produced 8 published studies, 3 published articles, 16 unpublished dissertation studies, 1 published and 1 unpublished review and application of research literature, and 1 published report which compiled information from school districts and teachers' associations from across the country. The conclusions of these 30 articles were organized into 6 "clusters" of preferred teacher evaluation characteristics which included:

I Teacher participation in evaluation decisions.

II Activities prior to classroom observation.

III Activities following classroom observation.

IV Training and competence of the evaluator.

V Conditions of the evaluation.

VI Purposes and criteria of the evaluation.

Each characteristic in a cluster was supported by at least 2 research studies, or was cited by the National School Public Relations Association as appearing multiple times in publications by teachers' organizations. This requirement eliminated several characteristics, such as self-evaluation, setting job targets, and group supervision by peers, which had appeared, on their face, to be valid for inclusion, but were mentioned in only 1 study. After this winnowing process, 4 of the 6 clusters included

4 teacher-preferred characteristics of evaluation. Number V, the "Conditions of evaluation", included 9 characteristics preferred by teachers, while "Activities prior to classroom observation", Number II, included only 3. From the 6 clusters, a set of questions was developed for inclusion in the survey instrument. (See Appendix A for a copy of the survey instrument.)

Mailing the Questionnaire

From the original census of 179 principals, 25 were identified by a random selection process for a subsequent request to be interviewed face-to-face. These 25, therefore, did not receive the mailed questionnaire. From the remaining 154 principals, 10 percent were selected randomly for a pilot study. An initial mailing was sent to this group of 16, with a follow-up mailing 2 weeks later. Thirteen principals responded. Following this pilot study, which confirmed the questionnaire, the instrument was mailed to the remaining 138 principals in the study group, with a follow-up mailing 3 weeks later.

Personal Interviews

Twenty-two of the 25 principals randomly selected were interviewed by this author. One principal refused to be interviewed; another insisted on a phone interview, rendering the results imcomparable; and a third rescheduled the interview twice, the final date being too late for inclusion in the results. The questions on the survey instrument were read exactly as they appeared, and the answers noted. Any unsolicited information of relevance to the study was also noted. Following the formal

interview, additional questions were asked, with the purpose of probing for information which might add other dimensions to the study. These questions appear in the Appendix B, following the survey instrument itself. The interviews were conducted for two purposes: to ascertain that the survey questions were clearly written and understandable, and that they elicited the desired information; and to provide additional information of interest to the study.

Research Questions to be Answered

This survey attempted to find answers to the following questions:

1. Do principals, in their evaluation of teachers, include procedures that are congruent with practices that have been identified in the literature as being associated with increased teacher satisfaction with evaluation?
2. Is there a difference between what a principal would ideally do and what he/she actually does in the process of evaluating teachers?
3. What effect, if any, does each of the following have on research questions 1 and 2?
    A. Number of years as a teacher.
    B. Number of years as a principal.
    C. Number of years in education.
    D. Highest degree earned.
    E. Staff size.
    F. Average income of families in the community.
4. Are district-approved teacher evaluation policies and procedures congruent with practices that have been identified in the literature as being associated with increased teacher satisfaction with evaluation?

Analysis of the Data

A frequency count of responses produced preliminary information for use in analyzing the first 2 questions. This data was then organized into a "Scale of Congruence", which allowed the results to be compared easily. (See Chapter III and Tables 3.2 and 3.3 for a full explanation of this scale.) In addition, a *t*-test was used to measure the significance of the difference between principals' ideal practice and present practice on each item of the questionnaire. Finally, a Multivariate Analysis of Variance (MANOVA) assessed the significance of differences between present practice and ideal practice by cluster.

The effects in question 3 were first analyzed using a cross tabulation, to detect any noticeable patterns among the 6 intervening variables. Next, a Multiple Discriminant Analysis was employed to determine whether significant variance occurred as a result of any of the 6 intervening variables. Question 4 was unable to be answered, as very few principals complied with the request to return district documents with their surveys.

Scope and Limitations of the Study

South Suburban Cook County, Illinois was chosen for the study because it includes a broad cross-section of socio-economic, racial, political, and cultural conditions. In addition, all of the school districts in the survey area were members of one regional special education cooperative. Moreover, the unique structure of Illinois' school systems--there are many small elementary districts--allowed the survey to focus on a relatively large number of school districts (45) and on elementary

principals only, while maintaining a manageable, yet adequate sample size.

School districts in the survey ranged from the wealthiest to the poorest in the state. Some districts were rather segregated racially, while others were generally well-integrated. Several districts included semi-rural conditions; others were quite urban. Not included in the survey were principals from exclusively rural districts, exclusively inner-city districts, or principals from smaller cities and towns. Therefore, additional research from other geographic and demographic perspectives would be required before the results of this study could be broadly generalized to these locations.

It is recognized that mailed surveys involve inherent limitations, and that face-to-face interviewing would provide the highest rate of questionnaire completion. It is also recognized that the respondents may not have answered all questions with complete veracity, since many questions focussed on rather sensitive issues regarding their professional performance.

Finally, although every attempt was made to avoid personal bias when the questions were constructed, it is assumed that some bias remains.

CHAPTER II

REVIEW OF THE LITERATURE

Teacher evaluation has been the subject of prolific writing and study. To lend clarity to the topic, therefore, this chapter is divided into three sections. The first section discusses attitudes toward evaluation which have been expressed by teachers and principals. The second section is a review of models and styles of evaluation which have been developed and tested, each with an eye toward addressing and alleviating some or all of the concerns expressed in the first section. The final section of the chapter is a collection of specific preferences about evaluation which have been expressed by teachers and teachers' organizations, and which form the basis for the present study.

Teachers' and Principals' Attitudes Toward Evaluation

Purposes of Evaluation

There seems to be general agreement between principals and teachers that improvement of instruction is the primary reason for evaluating teachers. In her dissertation study, Hauge found almost 100 percent of a group of 88 elementary principals in agreement with that statement.[1] The NEA discovered, in a nationwide sample survey of public-school classroom

---

[1]Hauge, abstract.

teachers, that 93 percent of the teachers in the survey thought the purpose of teacher evaluation should be "To assist in improving teacher competence."[1] When the National School Public Relations Association (NSPRA) asked the rhetorical question, "Why evaluate?", their answer was, "To improve teacher performance so as to provide a better education for our children."[2] Jones found that most school districts agree that the primary purpose of evaluation should be improvement of instruction.[3] Along that same line, Johnson lists three reasons for supervision (which can be translated evaluation): "(1) to protect children from incompetent teaching, (2) to administer curriculum, and (3) to assist each teacher to attain and maintain the maximum effectiveness in instruction."[4]

The second most important purpose of evaluation, according to the literature, is the facilitation of administrative decisions. For example, in a survey of administrators in the state of Arizona, Davis found that over 45 percent listed rehiring, tenure, and placement decisions as a purpose of evaluation, while more than 56 percent listed the improvement of instruction as a purpose.[5] Wolf questioned 293 teachers and found that they generally considered administrators as the most important audience for classroom evaluations, and teachers the least important audience.[6]

---

[1]National Education Association, p. 70.

[2]National School Public Relations Association, p. 8.

[3]Sheridan Davis Jones, "A Model for Identifying Evaluation Procedures That Have a Positive Influence On Teacher Attitude" (Ph.D. dissertation, Oregon State University, 1981), abstract.

[4]I. T. Johnson, "Why Supervise Teachers?" School Management 15 (October 1971): 34.

[5]Barbara Irene Davis, "The Status of Teacher Evaluation Practices in Arizona and a Proposed Model" (Ed.D. dissertation, University of Arizona, 1974), abstract.

[6]Robert L. Wolf, "How Teachers Feel Toward Evaluation," in School Evaluation: The Politics and the Proces ed. Ernest R. House (Berkeley, California: McCutchan Publishing, 1973), p. 164.

In addition to improvement of teacher performance, Holly and Hickman indicate that teacher evaluation should assist the district in the termination of staff who cannot reach minimum levels of performance, even with assistance. They also cite a third purpose: the communication of systematic expectations.[1] Crenshaw and Hoyle list improvement of instruction and administrative decisions as two goals for teacher evaluation;but they add two additional goals; to assess the overall school program, and to motivate teachers to render their highest level of professional service.[2] Jensen surveyed 46 teachers in an attempt to determine, among other things, the perceived purpose of evaluation. She found 72 percent mentioning self-growth and 63 percent listing accountability as a purpose.[3] The NSPRA would agree with that analysis. They see the impetus for teacher evaluation coming from two sources: (1) the public seeking assurance that the tax dollar is well spent (accountability), and (2) teachers seeking the security of fair, objective standards of evaluation.[4]

In a survey of administrators in independent schools, Cookson found several concerned with due process as a purpose for evaluating teachers. Evidently independent school administrators were finding teacher evaluation to be increasingly necessary, to substantiate decisions about rehiring and promoting teachers in the face of challenges to those decisions.[5]

---

[1]Holley and Hickman, p. 3.

[2]Harrison M. Crenshaw II and John R. Hoyle, "The Principal's Headache - Teacher Evaluation," NASSP Bulletin 6 (February 1981): 38.

[3]Mary A. Jensen, "How Teachers View Evaluation," Education 102 (Winter 1981): 132.

[4]National School Public Relations Association, p. 5.

[5]Peter W. Cookson, Jr., "Teacher Evaluation in Independent Schools," Independent School 34 (May 1980): 53.

Bolton, however, seems to provide the best overall statement, when he lists the purposes of teacher evaluation as follows:

1. To improve teaching, including out-of-classroom activities, as well as classroom instruction.

2. To reward superior performance.

3. To supply information for modification of assignments. (Including placement in another position, reduction of load, promotion to a leadership position or termination of employment.)

4. To protect individuals or the school system in legal matters. (Including both the protection of teachers against a capricious new administrator and the protection of the school district and children against a harmful teacher.)

5. To validate the selection process.

6. To provide a basis for career planning and individual growth and development of the teacher. (Including professional degrees and in-service training programs.)[1]

---

[1]Dale L. Bolton, Teacher Evaluation, PREP Reports, no. 21 (Washington, D.C.: HEW, Office of Education, National Center for Education Communication, 1972), p. 7.

Teacher Attitudes Toward Evaluation

"The attitude of teachers toward the evaluation of their teaching will influence their ability to profit from evaluation; those who hold favorable attitudes are more likely to benefit from evaluation than those who do not."[1] The attitudes of public school teachers toward the process of teacher evaluation is well documented, even though there seems to be very little consensus on the subject.[2] The NEA discovered that 90 percent of the teachers surveyed in 1969 thought that they should be evaluated, and 75 percent felt that both probationary and tenured teachers should be evaluated. Nearly 100 percent felt the principal should be the person doing the evaluating.[3] But in 1971, Osborne reported that in a national poll, only 16 percent of teachers favored teacher evaluation by their principal.[4] Jensen's study supports this lack of consensus. She found, for instance, that teachers differed in whom they would accept as an evaluator.[5] Some preferred a composite group of evaluators: peers, administrators, parents, children. Others preferred only the principal or a district administrator. She also discovered some rather disturbing attitudes. Of the 46 teachers interviewed, only a third were willing to participate in evaluation. An additional 24 percent were willing to participate, but only if significant qualifications were met. One-fifth were rather uncertain as to whether they were willing to participate, and 17 percent were simply unwilling.

---

[1]Wolf, p. 161.
[2]National School Public Relations Association, p. 57.
[3]National Education Association, pp. 70-71.
[4]Grace S. Osborne and Allan S. Hurlburt, "Credibility Gap in Supervision," School and Society 99 (November 1971): 415.
[5]Jensen, p. 132.

In spite of their reservations, however, teachers do seem to feel that supervision and evaluation is necessary in the schools. Of selected elementary teachers in western New York State surveyed by Young and Heichberger, four-fifths expressed that attitude.[1] However, 70 percent of the teachers also expressed the feeling that supervision is often viewed as potentially dangerous. Jensen discovered that two-thirds of the teachers she interviewed, reported feelings of discomfort when evaluated by someone else.[2] In a study by Ramsay, Tennessee teachers rated the overall evaluation process in their school system and found it to be no better than fair.[3] A study of teachers 15 years earlier in the Philadelphia area reinforces that finding, reporting that a sizable percentage of teachers considered the time they spent with their supervisors to be utterly wasteful.[4] One reason may be, according to Cookson, that teachers seem to "feel isolated, working in a professional vacuum without institutional supports to become better, more self assured teachers."[5] As a result, some teachers have felt skeptical about proposals to create formal methods of evaluation. Nevertheless, in Cookson's study, most of the in-

---

[1]James M. Young and Robert L. Heichberger, "Teachers' Perceptions of an Effective School Supervision and Evaluation Program," Education 96 (Fall 1975): 10.

[2]Jensen, p. 135.

[3]Russell G. Ramsay, "Teachers' Perceptions of the Design and Implementation of Teacher Evaluation Systems in Tennessee Public Schools" (Ed.D. dissertation, University of Tennessee, 1980), abstract.

[4]Arthur Blumberg and Edmund Amidon, "Teacher Perceptions of Supervisor-Teacher Interaction," Administrator's Notebook 14 (September 1965).

[5]Cookson, p. 53.

dependent school teachers believed that open, consistent evaluation would help them do a better job and improve faculty-administration relations, which would improve the school.[1]

After studying the issue of teacher attitudes toward evaluation, Wolf wrote a rather severe indictment of teacher evaluation:

> Teachers are not fond of evaluation. They suspect any measure designed to assess the quality of their teaching, and any appraisal usually arouses anxiety. Their opposition is far from simple obstructionism. Teachers recognize the administration's need to know, but they have a stake in evaluation too. The results are the major basis for promotions, pay raises, and dismissals. Their careers are in the appraisor's hands. If teachers are to submit to an assessment of their performance, they would probably like reassurance that the criteria and method of evaluation that are to be used would produce credible results.[2]

Nevertheless, Wolf did find that when the school climate is good, teachers' feelings about teacher evaluation seem to improve, and 99 percent of the teachers he surveyed recognized the importance of teacher evaluation, regardless of the institutional climate. In a study representing teachers of all levels in the elementary schools in five different states, Claye reached much the same conclusion. All teachers need and want supervision.[3] But, in a departure from Wolf's conclusions, Noonan decided that most school personnel find present evaluation processes, "adequate, but in need of improvement.[4] The Michigan Education Association points out that teachers who have been fearful of formalized evaluation, as something which was destructive and likely to harm them, should see evaluation as a

---

[1]Ibid.

[2]Wolf, p. 160.

[3]Clifton M. Claye, "Lola Gets What Lola Wants from Supervision," The Journal of Educational Research 56 (1963): 358.

[4]Joann Alice Noonan, "An Analysis of Perceptions of Michigan Superintendents, Principals, and Teachers in Regard to Present Versus Preferred Teacher Evaluation Systems" (Ed.D. dissertation, Western Michigan University, 1981), abstract.

process which provides "a continuous record which is the best long range protection against unjustified criticism." This continuous record "provides testimony as to the teacher's effectiveness, which may be necessary in a time of crisis."[1]

To summarize teacher attitudes toward evaluation, many seem to understand that supervision and evaluation are necessary, but there is a great deal of anxiety and a wide spread lack of confidence in present systems of evaluation.

---

[1]National School Public Relations Association, p. 54.

Attitudes: Principals Versus Teachers

Of more importance perhaps than teachers' attitudes alone is the lack of congruity between teachers' and principals attitudes. In a very long treatise on the problem which exist between principals and teachers, Blumberg indicated that many teachers see supervision and evaluation as a waste of time.[1] He also found in the course of several studies that most principals tend to view the results of their work very positively, and only a very few feel that what they do is a waste of time.[2] These results certainly indicate that a major lack of agreement exists between principals and teachers regarding the efficacy of teacher evaluation. A number of studies support this view. One study done in Florida, which surveyed over 700 teachers and nearly 550 principals, found that teachers and principals have very different perceptions of evaluation procedures and standards used by principals.[3] A similar conclusion was arrived at in a dissertation study by Davis.[4] Principals and teachers differed significantly on procedures and purposes they perceived as operating in their school's teacher evaluation programs, with principals generally perceiving the

---

[1]Arthur Blumberg, Supervisors and Teachers: A Private Cold War (Berkeley, California: McCutchen Publishing Co., 1974), p. 2.

[2]Ibid, p. 13.

[3]Florida Educational Research and Development Council, "Standards and Procedures Used for Evaluating Classroom Instruction of Annual Contract Teachers," FERDC Research Bulletin 1 (May 1965):4-6.

[4]Betty Gale Mullen Davis, "The Relationship Between School Climate and the Congruency of Perceptions of Elementary School Principals and Teachers Concerning Teacher Evaluation" (Ed.D. dissertation, University of Mississippi, 1982), abstract.

the situation to be more positive than teachers. In Cookson's study, 70 percent of the 20 independent school administrators were satisfied with the teacher evaluation process, while only 4 percent of the faculty were.[1] A similar study by Sapone discovered a wide difference in attitudes between administrators and teachers when they ranked the importance of each component of an evaluatinn model.[2] Hendrix also found significant differences in the way teachers, principals, and supervisors perceived a number of characteristics of teacher evaluation.[3] Farris discovered differences between experienced teachers and principals regarding the need, the purpose, the procedures and the results of teacher evaluation.[4] A study of teachers and principals in Tennessee found similar differences in perception between administrators and teachers. A significant lack of agreement was noted regarding the number of hours spent per teacher in the evaluation process, the number of hours per teacher which should be spent in the process, the number of observations performed per teacher and the number which should be performed, and the average length of an observation. Most significantly, teachers perceived the overall evaluation process to be significantly less acceptable than did administrators.[5]

---

[1]Cookson, p. 51.

[2]Carmelo V. Sapone, "Appraisal and Evaluation Systems: Perception of Administrators, Teachers," NASSP Bulletin 65 (February 1981): 29.

[3]Clifford L. Hendrix, Jr., "A Study of Supervisory Practices with Reference to Classroom Observations and Conferences as perceived by Teachers, Principals and Supervisors in Tennessee" (Ed.D. Dissertation, University of Tennessee, 1976), abstract.

[4]John Alvin Farris, "Teacher Evaluation Process in Evangelical Christian Schools" (Ed.D. dissertation, University of the Pacific, 1983), abstract.

[5]Samuel Lawson Houston, "An Analysis of Perceptions of Tennessee Public School Teachers, Building-Level Administrators and Central Office Administrators Toward Teacher Evaluation" (Ed.D. dissertation, University of Tennessee, 1981).

Finally, Tirrell found little agreement between teachers and principals concerning the role of the principal as an evaluator, even though there was a close philosophical agreement between the two groups regarding the ideal expectations of the principal's role in evaluation. Interestingly, neither teachers nor principals believed that principals were performing as they should in the role of evaluator.[1]

Not all of the studies indicate disagreement, however. Grant and Carvell surveyed principals and teachers in an elementary school and found strong agreement between the two groups concerning what constitutes desirable and undesirable teaching behaviors and the practicability of using such behaviors indetermining teacher evaluation criteria.[2] Moreover, when the study was expanded to 29 elementary principals and 100 teachers, the agreement concerning teacher evaluation criteria remained.[3] An additional study by Searles and Ng, which involved 22 principals and 41 biology teachers, also found agreement on the relative importance of most criteria for teacher evaluation.[4]

---

[1]Frederick John Tirrell, "An Investigation of the Current Perceptions and Ideal Expectations of Senior High School Principals and Teachers Regarding the Principals' Role in Teacher Evaluation" (Ed.D. dissertation, Boston College, 1982).

[2]Stephen Grant and Robert Carvell, "A Survey of Elementary School Principals and Teachers: Teacher Evaluation Criteria," Education 100 (Spring 1980): 226.

[3]Ibid.

[4]William G. Searles and Raymond W. M. Ng, "A Comparison of Teacher and Principal Perceptions of an Outstanding Biology Teacher," Journal of Research in Science Teaching 19 (September 1981).

Some of the differences which have been noted between teachers and principals might be bridged, according to the results of one study, if supervisors can communicate their positive attitude toward evaluation. If so, measurable positive changes in attitudes toward evaluation may be produced in teachers.[1]

---

[1]David Thomas Williard, "An Assessment of the Effects Of a Staff Evaluation Model Developed From Douglas MacGregor's Theory "Y" Upon The Attitudes of Teachers and Supervisors Toward Evaluation" (Ph.D. dissertation, St. Louis University, 1979), abstract.

Models of Teacher Evaluation

In this section, a number of specific styles and/or formal models of evaluation will be discussed. Some of these enjoy strong advocacy by one or more writers, while many have been studied in specific research projects. Each enjoys some degree of legitimacy by having appeared more than once in the professional literature. Each attempts to address one or more of the problems mentioned in the previous section.

Self-evaluation

Self-evaluation, or self assessment, can take one of several forms, including videotaped feedback, self-perception, and the self-completion of observation forms. Much of the research around self-evaluation has tended to view a single strategy as the total process of self-assessment, and therefore has been somewhat misleading.[1] A number of authors, including Musella, Waimon and Ramseyer, and Ahnell and Hawn, advocate the use of videotaping, and have studied its use as a strategy in self-evaluation.[2]

Some important elements of the process include jointly developing the videotape evaluation criteria with the teacher, providing a constant videotape image of the teacher, and concentrating supervision on

---

[1]Gerald D. Bailey, Teacher Self-Assessment: A Means For Improving Classroom Instruction (Washington, D.C.: National Education Association, 1981), p. 11.

[2]Donald Musella, "Improving Teacher Evaluation," Journal of Teacher Education 21 (Spring 1970) pp.15-21; Morton D. Waimon and Gary C. Ramseyer, "Effects Of Video Feedback On The Ability To Evaluate Teaching," Journal of Teacher Evaluation 21 (Spring 1970); I.V. Ahnell and Horace C. Hawn, "Self-Evaluation Through Videotape Recordings," in Observational Methods In The Classroom, ed. Charles W. Beegle and Richard M. Brandt (Washington, D.C.: Association for Supervision and Curriculum Development, 1973).

one task at a time, assuring that each component task is mastered before moving on. Several studies have demonstrated that judgments made from videotape correlate well with judgments that are made first hand by an observer, indicating that videotaping can be an effective observational and supervisory tool, when used correctly.[1]

According to journal articles by Irvine, who studied pre-service teachers, and by Newfield, who studied in-service teachers, and a dissertation study by Cartlidge, who studied beginning teachers, teachers' self reporting without the use of videotapes can be in agreement with the evaluation rating of supervisors and principals, and can be accurate under certain conditions, including training designed to facilitate self-assessment and collegiate relationships.[2] Other studies disagree, however, and have found a discrepancy between the perceptions of classroom observers and teachers' self-perceptions.[3]

In spite of its limitations, self-evaluation, or self-assessment, maintains a fairly loyal following. For example, Crenshaw and Hoyle list

---

[1]Hosford and Martin, p. 7.

[2]Jacqueline Jordan Irving, "The Accuracy of Pre-Service Teachers' Assessments of Their Classroom Behaviors," Journal of Research snd Development in Education 17 (Fall 1983); John Newfield, "Accuracy of Teacher Reports: Reports and Observations of Specific Classroom Behaviors," Journal of Educational Research 74 (November-December 1980): 78; and Arthur James Cartlidge, "A Study of Self and Principals' Evaluation of Beginning Teachers in Selected Mississippi Public Schools" (Ed.D. dissertation, Mississippi State University, 1983).

[3]See Newfield, p. 78; Thomas W. Jones, "The Validity of Self-Evaluation of Competence by Special Education Teacher Trainees," Texas Tech Journal of Education 9 (Winter 1982); and Bailey, p. 1.

three potential problems with self-evaluation, i.e.: teachers who are secure overrate themselves, while emotionally insecure teachers underrate themselves, and few teachers are able objectively to assess their own performance. Nevertheless, they make the statement, "Self-evaluation is the key to professionalism."[1] Bodine states that:

> Self-assessment is probably the most powerful means yet developed for a teacher to be the master of his own professional growth. Self-assessment is bold, but easy to understand, revealing and thus threatening, majestic in goal and thus giving dignity to the teaching profession."[2]

A number of other authors would place themselves in full agreement with this assessment.[3]

Evaluation By Students

During the middle ages, students set up their own committees to report on professors who failed to cover required segments of learning in the specified time. The faulting professors were fined.[4]

Although students today are not granted the power to fine their professors, many colleges and universities continue to use some form of student rating or evaluation as part of an overall faculty performance

---

[1]Crenshaw and Hoyle, p. 40.

[2]Richard Bodine, "Teachers' Self-Assessment," in School Evaluation: The Politics and Process, ed. ernest R. House (Berkeley, Calif.: McCutchen Publishing Co. 1973), p. 171.

[3]Bailey, p. 13; Gerald Bryant and Frank Haack, "Appraisal: Peer-Centered and Administration-Centered," Educational Leadership (May 1977): 609; and Daniel F. Detzner, "Teacher Evaluation: A Self-Appraisal Method," The University of Minnesota General College Studies 11 (1974-75),

[4]National School Public Relations Association, p. 23.

evaluation. Some student evaluation of teachers is also found in the high schools and elementary schools. Of all the sources of evaluative information, this perhaps is the most controversial.

In a report on teacher evaluation, the Ohio Commission on Public School Personnel Policies reported "that informational feedback from students is an effective means of influencing teacher behavior and, in fact, student feedback can sometimes be more effective in changing teacher behavior than supervisory feedback."[1] The Ohio report suggests one particular advantage of student evaluation: "It is available to teachers whenever they desire to employ it. Thus, evaluation can be an ongoing process and does not have to be dependent upon the assistance of a principal or supervisor.[2]

At the college level, where student evaluations are quite common, there seems to be little unanimity regarding the validity and the proper use of student evaluations. In analyzing 129 replies to a questionnaire sent to department chairs and authors, one study found that student evaluations were the most common method of faculty evaluation being used, but repeatedly the comments indicated that student evaluations should not be considered sufficient to evaluate faculty.[3] Another study surveyed college faculty and discovered a general feeling that student evaluations

---

[1]
[2]Ibid, p. 24.
Ibid.

[3]M. E. Schaff and B. R. Siebring, "What Do Chemistry Professors Think About Evaluation of Instruction?" Journal of Chemical Education 51 (March 1974): 154.

are not valid instruments. In spite of that, 60 percent of those surveyed used student evaluations even when optional. Nevertheless, most of the surveyed faculty felt that administrators should not have access to student evaluations unless the teacher desired it. But most significantly, the authors found that the perceptions of teachers regarding student evaluation did not, in most cases, agree with their review of research findings.[1]

It has been noted that at the college level, student evaluation of teachers seems to have become firmly entrenched. Below the college level, however, there is much less frequent use of student ratings, and even less agreement as to their usefulness, according to several surveys noted by the NSPRA.[2] In 1971, the NEA surveyed its members and found that 38 percent favored student ratings of teachers, while 31.9 percent were opposed. A similar survey by the Nation's Schools in 1970 reported over 42 percent opposed, but over 40 percent in favor. A 1973 survey by Educational Resource Services reported only 24 percent of 468 districts indicating some use of student evaluation of teachers.

A number of studies have attempted to determine the usefulness of student ratings in the evaluation of teachers. The Committee on IOTA has found that pupil ratings have little correlation with other measures of teacher effectiveness. They feel that if student evaluation of teachers is valid, then the following is assumed: (1) That what the pupil observes

---

[1]Charlotte Epstein, "Student Grade Teachers? Some Faculty Attitudes", Community and Junior College Journal 44 (April 1974): 33.

[2]The three surveys noted in this paragraph were summarized by the National School Public Relations Association, pp. 23-24.

represents a comprehensive sampling of teacher responsibilities. (2) That pupils are capable of comprehending the important goals to be achieved in the classroom and the effectiveness of various activities in achieving these goals. (3) That pupils can identify teacher behavior that is indicative of ability to adjust to these goals.[1] Morrow's study, however, of student evaluation of teachers in 103 physical education classes, does not support such assumptions.[2]

A study of students in two college courses by Marsh, Overall and Kessler did find considerable agreement between college students and faculty in their description of faculty behaviors and in their overall rating of teacher behaviors, seeming to reaffirm the validity of student evaluations.[3] A study of elementary school students in California, moreover, assessed the ability of elementary students to judge student teachers' performances, in comparison with adult observers' ratings, and found "that elementary school students can assess performance and discriminate among teaching tasks.'[4] Yet another study determined that evaluations by sixth grade students appeared to be a reliable measure of teacher behavior.[5]

---

[1]Herbert W. Marsh, J. U. Overall, and Stephen P. Kessler, "Validity of student Evaluations of Instructional Effectiveness. A Comparison of Faculty Self-Evaluations and Evaluations by their Students," Journal of Educational Psychology 71 (April 1979):149.

[2]National School Public Relations Association, pp. 23-24.

[3]Marsh, Overall, and Kessler, p. 149.

[4]Ellen Kronowitz and Victoria Finney, "Student Teachers' Performance," California Journal of Teacher Education 10 (Winter 1983)

[5]Ronald Fox, et al., "Student Evaluation of Teacher As a Measure of Teacher Behavior and Teacher Impact on Students," Journal of Educational Research 77 (September-October 1983): 21.

When considering the validity of students evaluating teachers, the possibility is very real that bias on the part of the students can invalidate the results. A few studies have examined this phenomenon, but the results have been inconclusive. Holmes found that when students received a lower grade than expected, they tended to deprecate the instructor's teaching performance.[1] Another study found a strong statistical relationship between the expected grade of the class and its evaluation of the teacher.[2] Bassin discerned a significant pattern of bias in student's evaluations of instructors,[3] while Harris, who looked at possible sex bias, discovered none, but did determine that the masculinity or femininity of a teacher's style of teaching may have some effect on how that teacher is perceived.[4] Sihota found that a variety of variables affected student evaluations, although not consistently nor significantly. The highest correlation was found between the average grade given to the student and the evaluation of the instructor.[5] Larsen discovered that a student's ratings of his professors seemed to depend on his expectations upon entering the course, rather than his experience upon leaving the course.[6]

---

[1]David S. Holmes, "Effects of Grades and Disconfirmed Grade Expectancies on Students' Evaluation of Their Instructor," Journal of Educational Psychology 63 (April 1972): 143.

[2]Rolf Mirus, "Some Implications of Student Evaluation of Teachers," Journal of Economic Education 5 (February 1973): 37.

[3]William M. Bassin, "A Note On The Biases In Students' Evaluations of Instructors," Journal of Experimental Education 43 (February 1974): 16-17.

[4]Mary B. Harris, "Sex Role Stereotypes and Teacher Evaluations," Journal of Educational Psychology 67 (December 1975): 756.

[5]Sohan S. Sohota and Ram P. Singhania, "On teaching Effectiveness," Journal of Business Education 57 (November 1981): 55.

[6]Edwin M. Larsen, "Students' Criteria For Responses To Teaching Evaluation Questionnaires," Journal of Chemical Education 51 (March 1974): 165.

Fallman, however, looked at several referents commonly used by college students in rating professors on a standard rating scale, and found no effect on the reliability or the level of the rating.[1]

Arreola found that college students could distinguish between the personality of the teacher and the content and organization of the class.[2] However, a study by Zelby seems to refute that conclusion. He showed that by teaching the same course differently in different semesters, it was possible to teach to a particular student evaluation, given a particular questionnaire.[3]

And yet, in spite of inconclusive research studies, and in spite of a lack of agreement among educators, Menges maintains that "observations by students seem essential if classroom events are one focus for evaluation of teaching."[4] He also reminds us that students are reporters only, not judges. Bolton supports that idea, advocating student ratings, but only as one source of data.[5] After statistically treating three years of ratings of professors by college students, Cornwell concluded that student evaluation is reliable enough to be a viable ingredient in an overall program of teacher evaluation.[6]

---

[1]John Fallman, et al., "Student Raters' Referents In Rating College Teaching Effectiveness," Journal of Psychology 86 (March 1974): 247.

[2]Raoul A. Arreola, "Students Can Distinguish Between Personality and Content/Organization in Rating Teachers," Phi Delta Kappan 65 (November 1983): 222-223.

[3]Leon W. Zelby, "Student-Faculty Evaluation," Science 183 (March 1974).

[4]Robert J. Menges, "The New Reporters: Students Rate Instructors" in Evaluating Learning and Teaching, ed. C. Robert Pace (San Francisco: Jossey-Bass, 1973), p. 59.

[5]Bolton, p. 72.

[6]C. D. Cornwell, "Statistical Treatment of Data from Student Teaching Evaluation Questionnaires," Journal of Chemical Education 51 (March 1974): 159.

Evaluation by Peers

The idea of evaluation by and of one's peers has a certain atractiveness to it. The concept seems to enhance the professionalism of teaching; and who should be more qualified to evaluate a teacher's performance than a colleague who, by training, is an expert at what he or she does? To the question "Who should do the assessing"? Leese replies "...teacher peers should do so."[1] He goes on to say that teachers should evaluate teachers for the following reasons:

> Fellow teachers have as much at stake in quality effort as does anyone else. They have knowledge of the content which must be intimately interwoven in the tactics used to produce meaning, develop concepts, and produce generalizations and applications. They are more familiar with the conditions, the relationships, and the reasons that underlie and effect responses and choices. They are less threatening, more likely to be helpful, and their first hand involvement contributes to moral and to their own identification of aspects and areas for self-improvement. [2]

The NSPRA quotes Garford G. Gordon, the research executive of the California Teachers Association, who offers some additional arguments and support for peer evaluation: (1) It is impossible for the principal to collect all data necessary for a valid evaluation of staff. (2) No administrator can be familiar with the latest developments in education, across all ages and subject areas. (3) Peer evaluation separates evaluative judgment from the consequences.[3]

Cookson's survey of teachers in independent schools discovered that most of the teachers who responded preferred peer observation and discussion to other methods of evaluation.[4] A study by Miller found that teachers

---

[1]Joseph Leese, "Teacher Assessment and Consensus Preference," NASSP Bulletin 65 (November 1981): 26.

[2]Ibid.

[3]National School Public Relations Association, p. 11.

[4]Cookson, p. 51.

seem to accept the concept of formally evaluating peers.[1]

In spite of the appealing nature of the idea, however, there are some potential drawbacks. For example, Stronck found that peers were not as critical as students when evaluating student teachers in a micro-teaching experience.[2] Levine reported wide differences in standards of grading when faculty members evaluated each other, leading her to question the validity of colleague observation.[3] Some districts who have attempted peer evaluation have encountered these problems: Teachers actually involved in peer evaluation seem to be reluctant to judge their fellow teachers. And, moreover, if teachers are to evaluate their peers in an acceptable manner, they must be released from class and they must be trained, which results in an expense to the district of both time and money.[4]

---

[1]Duane Dean Miller, "The Development of A Process Model To Utilize Peer Opinion for Teacher Evaluation in Selected Schools in Nebraska," (Ed. D. dissertation, University of Nebraska, 1974), abstract.

[2]David R. Stronck, "A Comparison of Peer and Pupil Evaluations of Lessons Taught by Preservice Biology Teachers," Science Education 60 (January-March 1976): 220.

[3]Judith R. Levine, "When Colleagues Judge Colleagues," Teaching of Psychology 11 (February, 1984): 38.

[4]National School Public Relations Association, p. 10.

Student Progress In The Evaluation of Teachers

When considering the evaluation of a teacher's performance, one of those logical components which comes to mind is student progress made under that teacher's tutelage. No less respected an educator than Benjamin Bloom has said that teacher effectiveness can only be measured in terms of learning outcomes.[1] It seems only reasonable when Popham asserts that "the supervisor should be most attentive not to teacher activity, but to what happens to the learners as a consequence of what the teacher does."[2] But, as with all other aspects of teacher evaluation, there are no simple answers; and, in this case, what seems the most straight-forward may be the most complicated. A number of major problems become evident as soon as teaching effectiveness is measured by student progress.

One problem identified by the NSPRA is that standardized tests as used in the schools have not been set up to evaluate the progress of students; rather, they were established to differentiate between students to establish a continumum which often forms a bell curve. One possible answer to this problem, however, is the use of criterion-referenced tests which would measure minimum levels of competence.[3] But other problems have been noted by numerous authors: (1) Establishing standards in each subject area, at each grade, requires setting up thousands of progress indicators throughout the curriculum, not to mention obtaining agreement of those indicators. (2) There is the danger of freezing teaching into a frigid

---

[1]N. L. Gage, "Desirable Behaviors of Teachers," Urban Education 1 (Winter, 1965): 379-397.

[2]W. James Popham, Evaluating Instruction (Englewood Cliffs, N.J.; Prentice-Hall, 1973), p. 47.

[3]National School Public Relations Association, pp. 13-14.

mold. (3) The teacher must have a pretest score on each student before he or she enters the class. (4) How can it be determined that measured pupil changes can be wholly attributed to the classroom teacher? (5) Can desired pupil changes be objectively defined and adequately measured?

In an effort to determine the usefulness of student achievement in the evaluation of teaching, Brophy studied teachers conducting their own classes over a three year period. He was successful in identifying teachers in grades two and three who were consistent in their overall relative effectiveness. However, he discovered wide individual differences and only moderate stability from year to year, and unique "class effects" were observed despite statistical controls. His conclusion was that the use of measured student gain, or general achievement tests, for assessing teacher accountability is inappropriate and unfair to many teachers.[1] Musella drew much the same conclusions, finding it nearly impossible to determine a cause-effect link between teaching effectiveness and student growth criteria.[2]

---

[1]Jere E. Brophy, "Stability of Teacher Effectiveness," American Educational Research Journal 10 (Summer 1973): 251.

[2]Donald Musella, "Open-Closed-Mindedness As Related To The Rating of Teachers By Elementary School Principals," Journal of Experimental Education 35 (Spring, 1967): 16.

Rating Scales and Evaluation Forms

Probably the most common method of evaluating teacher performance is to rate a teacher on some type of scale. In a survey of the 60 largest school districts in the country, the Pittsburgh Public Schools found that of 53 districts responding, 50 were using some type of rating scale to measure teacher performance.[1] Rating forms used by school districts seem to fall into five categories: (1) Personal characteristics, (2) Clasroom management, (3) Relationships with community and staff, (4) Relationships with pupils, and (5) Lesson planning and presentation.[2]

Some educators and researchers, however, find that the use of rating scales for the evaluation of teaching effectiveness is faulty. Woody states two objections to the use of rating scales: (1) The scales are much too general to be applicable to all types of teaching. (2) The use of a rating scale reflects the educational values of the rater rather than the efficiency of the teacher in achieving the values which seem worthwhile to him.[3] Research by Fagan would support those conclusions. The results of Fagan's study indicate that background attributes,

---

[1]Glenn Queer, An Analysis Of Teacher Rating Scales: A National Survey (Washington D.C.: U.S. Department of Health, Education and Welfare, June 1969), p. 1.

[2]Richard A. King, "Reliable Rating Sheets: A Key To Effective Teacher Evaluation," NASSP Bulletin 62 (December 1978): 23.

[3]Clifford Woody, "Some Observations Of Methods Of Research In The Appraisal Of Teaching," Educational Administration And Supervision 29 (January 1943): 1.

values, and discrepancy measures "are highly significant in predicting evaluator ratings of teachers".[1] In a survey of classroom instruments used for evaluating teaching performance, Rosenshine found that it was very difficult to determine whether a rating scale was being used to judge the value or to estimate the frequency of a behavior.[2] The Committee on IOTA established four guidelines in the selection of an instrument for rating teaching effectiveness: (1) The instrument must be valid. It must sample the areas of competence defined by the criteria. (2) Devising a rating instrument calls for selection of what is most important, and requires explicit and considered judgment. (3) The instrument must be adapted for local use, and therefore must have local validity as well as general validity. (4) The instrument must be concise, have discriminative ability, and be free from personal bias.[3]

Performance Tests

Although no studies were found which have assessed the use of a performance test for measuring teaching effectiveness, two authors advocate the use of such a test as an alternative to other forms of teacher evaluation. McNeil described a situation wherein a number of teachers are given identical instructional tasks or objectives and a sample of a measure to be administered to pupils after the teaching has occurred.

---

[1]Lenora Perry Fagan, "Teacher Evaluation: Factors Influencing The Perceptions Of Evaluators" (Ph.D. dissertation, University of Toronto, 1983), abstract.

[2]Barak Rosenshine, "Evaluation Of Classroom Instruction," Review of Educational Research 40 (April 1970).

[3]The Committee On IOTA, p. 29.

After a specific period of time to plan the lesson and instructional plan, randomly assigned students are brought to the teacher. Following the instructional period, pupils complete a test which measures pupil attainment of the instructional objectives. The mean of the test scores determines the teacher's ability to teach the predetermined skill.[1] Popham advocates a similar kind of teaching performance test as a way to measure teaching performance without the "contaminating" factors which exist in a normal classroom environment.[2]

Wanat, however, questions the use of such performance tests. He sees a number of problems: (1) How representative are the tasks and conditions of the test? (2) The performance test does not recognize individual differences of teachers. (3) How real are the conditions of the test, the tasks expected? (4) There are other factors which still contaminate the results. (5) Is the test to be used in a punitive manner or diagnostically.[3]

---

[1]John D. McNeil, "Performance Tests: Assessing Teachers Of Reading - A Proposal," The Reading Teacher 25 (April 1972): 622.

[2]W. Popham, "Found: A Practical Procedure to Appraise Teacher Achievement in the Classroom," in Accountability for Teachers and School Administrators, ed. Allan C. Ornstein. (Belmont, California: Fearon Publishers, 1973), pp. 25-26.

[3]Stanley F. Wanat, "Performance Tests: Assessing Teachers Of Reading - A Response," The Reading Teacher 25 (April 1972): 623.

Clinical Supervision

In 1969, Goldhammer's book on clinical supervision was published posthumously, and introduced a significant departure from traditional methods of teacher evaluation.[1] In 1973, his mentor, Cogan, also published a book about clinical supervision, reiterating the concepts of collaboration, formal teacher evaluation cycles, focusing on selected teacher practices, and teacher autonomy.[2] Goldhammer's cycle of supervision includes five steps: (1) Preobservation conference, (2) The observation, (3) Analysis and strategy, (4) Supervision conference, and (5) Postconference analysis. Cogan's cycle of evaluation includes eight steps: (1) Establishing a relationship, (2) Planning with the teacher, (3) Planning the strategy of observation, (4) Observing instruction, (5) Analyzing the teaching-learning process, (6) Planning the strategy of the conference, (7) The conference, (8) Renewed planning. Other practitioners have since refined the cycle of supervision and adapted it to their individual needs. The critical factor in clinical supervision, however, is a genuine feeling of colleagueship and mutuality in the relationship.[3] These concepts, of course, are quite attractive to professional educators. Those who have embraced the concept seem to do so with almost a religious fervor. Consider the following:

---

[1]Robert Goldhammer, Clinical Supervision (New York: Holt, Rinehart and Winston, 1969).

[2]Morris L. Cogan, Clinical Supervision (Boston: Houghton Mifflin, 1973).

[3]Charles A. Reavis, "Clinical Supervision: A Review Of The Research," Educational Leadership 35 (April 1978): 580.

> The development of colleagueship between teachers and supervisors and among teachers seems to offer three major benefits: (1) Mobilization of the human resources of the school for the formidable task of instructional improvement; (2) increased intrinsic rewards to enhance job satisfaction for teachers; and (3) increased likelihood of successful implementation of instructional innovations.[1]

Although relatively few researchers have studied clinical supervision, those who have discovered rather favorable results. Reavis looked at three studies of teacher attitudes toward clinical supervision, and found that teachers tended to favor it as a process.[2] Reavis also states that no study has found traditional supervision effective in changing teaching behaviors when compared to clinical supervision.[3] In a study by Tomblin, however, a participatory style of teacher evaluation very similar to clinical supervision was compared with a more traditional observation style of evaluation, and no significant differences were found between the groups on (1) attitudes toward evaluation, (2) self-improvement and growth, and (3) attitudes toward school problems and school administrators.[4] In spite of these findings, both teachers and principals who were involved in the study indicated that they preferred the participatory method, but that because of the amount of time required, full implementation of the model possibly did not occur. Another study, which explored the use of the clinical supervision model in the context of peer supervision, found that

---

[1]Robert J. Alfonso and Lee Goldsberry, "Colleagueship In Supervision," In Supervision Of Teaching, ed. Thomas J. Sergiovanni (Alexandria, Va.: Association For Supervision and Curriculum Development, 1982) p. 106.

[2]Reavis, p. 584.

[3]Ibid.

[4]Elizabeth Alene Risinger Tomblin, "Effects Of Participatory And Nonparticipatory Methods Of Teacher evaluation On Selected Teacher Variables" (Ph.D. dissertation, University of Colorado at Boulder, 1976), abstract.

the underlying assumptions of the model are incongruent within a school setting.[1]

In spite of its initial attractiveness, therefore, clinical supervision may suffer from some significant limitations. The problem of time has been mentioned. In addition, Harris notes three further problems: Clinical supervision does not call for clearly specified public criteria of performance expectations. The model tends to ignore data gathering and analysis procedures. And even though it emphasizes goal setting and improvement as a process, it disregards the details of a systematic objective evaluation process.[2]

Supervision By Objectives

Another model for supervising and evaluating classroom teachers which departs significantly from traditional methods is a system derived from commerce and industry called Management By Objectives, or MBO. In many ways similar to clinical supervision, MBO differs in that, while clinical supervision deals with a specific teaching episode, MBO deals with a specific identified aspect of teaching which needs improvement. One of the strongest supporters of MBO and its application to education, Redfern established six components in the evaluation process when using a supervision by objectives model: (1) Responsibility criteria, (2) Identify needs, (3) Determine objectives and Action plans, (4) Carry out Action

---

[1]Shirley A. McFaul and James Cooper, "Peer Clinical Supervision: Theory vs. Reality," Educational Leadership 41 (April 1984).

[2]Ben M. Harris, "Teacher Evaluation As A Developmental System," ERIC # ED224800, 1983.

plans, (5) Assess results, (6) Discuss results.[1] According to Redfern, the focus of any evaluation system must be on the effectiveness of the results, and not on the individual effort of the teacher. Such a statement is reminiscent of those supporting student achievement as a means of evaluating teacher performance, but leads to a more practical and realistic process for measuring outcomes. Crenshaw and Hoyle, for example, point out that although outcomes are important in the measurement of performance, the difference between school and industry has primarily been a confusion over results. Test scores, they say, are only one part of the desired outcome of the teaching experience.[2] Using the results of a number of studies which have looked at an objectives-based approach to evaluation, McNeil concluded that such a model is generally more effective than traditional evaluation methods.[3] According to McNeil, not only do teachers prefer this method of evaluation, but a study of elementary students indicates that achievement is greater when specific goals are set.[4] Eads found that teachers who are evaluated under a supervision by objectives model tend to feel more positive toward teacher evaluation than when evaluated by other methods.[5]

---

[1]George B. Redfern, Evaluating Teachers and Administrators: A Performance Objectives Approach (Boulder, Colorado: Westview Press, 1980), p. 13.

[2]Crenshaw and Hoyle, p. 43.

[3]John D. McNeil, "Concomitants Of Using Behavioral Objectives In The Assessment Of Teacher Effectiveness," The Journal Of Experimental Education 36 (Fall 1967): 71.

[4]Ibid.

[5]Albert Edward Eads, Jr., "A Study Of The Attitudes Of Teachers Toward A 'Supervision By Objectives' Teacher Evaluation Model" (Ph.D. dissertation, University of South Carolina, 1974), abstract.

Although most districts who have adopted an objectives based evaluation program seem to be satisfied with the results, the NSPRA has discovered some problems with the model.[1] One of the problems seems to be that teachers may write objectives which do not challenge them. Moreover, objectives or goals may be too vague, too global, or too ambitious. Thus, a great deal of inservice training is required for this system to be effective. The premier problem, however, as with clinical supervision, seems to be time. An MBO system demands a great deal of time from the teacher and the supervisor.[2]

---

[1]National School Public Relations Association, p. 15.
[2]Ibid.

Evaluation Characteristics Preferred By Teachers

The first section of this chapter examined teachers' attitudes toward evaluation, and found that most teachers seem to acknowledge the need for evaluation, but mistrust the manner in which it is implemented. A review of several evaluation models in the second section uncovered a number of specific characteristics of evaluation which teachers seem to like, such as collegiality, specific objectives, and peer evaluation. This final section summarizes a group of such characteristics, supported by research, which contribute to improved teacher attitudes toward evaluation, and have been shown to be preferred by teachers.

Teacher Participation

> The growing practice--and most noteworthy new trend--is to involve teachers in the establishment of evaluation programs. The unilateral imposition of the administrator is going the way of the dinosaur.[1]

Although the NSPRA reports a growing trend of teacher involvement in the establishment of evaluation programs, it is not clear whether teacher involvement is as wide spread as the NSPRA would like us to believe. Tobia, for example, who determined that involving Pennsylvania teachers in developing a teacher evaluation process was positively related to their attitudes toward evaluation, found little or no teacher involvement in evaluation across the commonwealth of Pennsylvania.[2] Participation in developing the evaluation process has been shown to lead to better teacher

---

[1]Ibid.

[2]Edward Francis Tobia, "The Relationship Between Teacher Participation in the Development of a Teacher Evaluation Process and Teacher Attitude Toward Evaluation" (Ed.D. dissertation, Temple University, 1984), abstract.

attitudes toward evaluation, which in turn lead to higher benefits from the evaluation process.[1] Henderson found that teachers who felt that had high participation in school decision making had a higher morale than teachers who thought they did not participate.[2] Likewise, Richards found that participatory decision making was an important factor in the attitudes of teachers toward evaluation.[3] Paulin doscovered that secondary teachers in a district containing 7 high schools were more receptive to evaluation when they had greater input into the evaluation process. Her findings indicate that evaluation will be most effective when teachers are equitably represented in evaluation design and implementation.[4]

Young and Heichberger's survey of elementary teachers in Western New York supports these findings. The teachers in their study felt strongly that teachers should play a role in the development of the teacher evaluation program.[5] In a dissertation study, Miller discovered that, while most of the middle school teachers in her survey reported that evaluation

---

[1] Wolf, p. 161.

[2] Lester F. Henderson, "Elementary Teachers' Satisfaction and Morale and Perceived Participation in Decision-Making" (Ed.D. dissertation, University of Arkansas, 1967), abstract.

[3] William Howard Richards III, "The Effect of a Professional Growth and Evaluation Cycle Upon Experienced Teacher Attitudes" (Ed.D. dissertation, Boston University, 1983), abstract.

[4] Pauline Paulin, "The Politics of Evaluation at the Local Level: A View Through Teachers' Perspectives," paper presented at the Annual Meeting of the American Educational Research Association, Los Angeles, 13-17 April, 1981.

[5] Young and Heichberger, p. 11.

was not a supportive experience, having more meaningful input into the process would help alleviate those feelings.[1]

The common conclusion of these 7 separate research studies is that teacher participation in evaluation decisions leads to better attitudes and morale, and consequently to more effective use of the evaluation. Specifically, teachers prefer to be involved in 4 distinct parts of the evaluation process:

(1) Teachers prefer to participate in designing the evaluation form. In Houston's study of Tennessee teachers and administrators, he determined a more positive perception by teachers of the overall evaluation process when they participated in the development of evaluation procedures and instruments. One hundred percent of the suburban and rural elementary teachers surveyed by Young and Heichberger in Western New York State indicated that teachers should take part in developing or selecting evaluation instruments.[3]

(2) Teachers prefer to participate in the design of the evaluation process. Tobia found evidence that involving Pennsylvania teachers in developing a teacher evaluation process was positively related to teacher attitude toward evaluation.[4]

(3) Teachers prefer to participate in developing the policies which govern evaluation.

(4) Teachers prefer to participate in establishing the goals and purposes of evaluation. Bolton indicates that goals are more likely to

---

[1]Mary Michaelle Miller, "The Evaluation Style, Methods, and Modes Preferred By Middle School Teachers and Administrators in an Urban Setting" (Ed.D. dissertation, University of San Francisco, 1981), abstract.

[2]Houston, abstract.

[3]Young and Heichberger, p. 15.

[4]Tobia, abstract.

be understood and attained when they are cooperatively developed by the teacher and principal.[1]

In a 1974 NSPRA summary and compilation of information from teachers' associations across the country, teachers' groups assert that teachers should be included in shaping the policies, setting the goals, designing the instruments, and carrying out the procedures of teacher evaluation.[2]

Activities Prior to Classroom Observation

The preobservation and postobservation conferences were discussed previously in the context of clinical supervision and management by objectives. Although a number of writers were shown to support the concept of pre- and postconferences, the number of studies indicating that teachers prefer to be involved in preconferences is somewhat limited. In the only study which specifically assessed teachers' feelings regarding a preconference, S. D. Jones determined that the preconference and postconference were viewed positively by teachers in the study group. Specifically, he discovered what seemed to be a better attitude toward evaluation in school districts using the clinical supervision model.[3] Several other studies, however, drew conclusions which strongly support the concept of a pre-observation conference, even though only one specifically mentions it by name. In an examination of the perceptions of selected Idaho principals and teachers, Bauer discovered that the teachers in the study con-

---

[1] Bolton, p. 16.

[2] National School Public Relations Association, p. 56.

[3] Sheridan Davis Jones, "A Model For Identifying Evaluation Procedures That Have A Positive Influence On Teacher Attitude" (Ph.D. dissertation, Oregon State University, 1981), abstract.

sidered the activities which occur prior to classroom observation as more important than did the principals, supporting the notion that a preconference might improve teachers' feelings about evaluation.[1] In a study of educators in Christian schools, Farris found that the principals and experienced teachers in the study group agreed on the importance and the content of a preconference, although they agreed on little else.[2] In Young and Heichbergers' study, 70 percent of the teachers expressed a desire to meet with supervisors and discuss objectives and plans together.[3]

In a dissertation study involving Montana teachers, Fraser discovered that 64 percent of the teachers in the study group wanted to use a preobservation conference to reach agreement with the principal on lesson objectives, while 67 percent wanted a preconference to discuss the data to be collected and how it was to be gathered.[4]

After collecting information from teachers' groups across the country, the NSPRA stated in its 1974 report that, according to the teachers' organizations surveyed, the time, place, and conditions of any visitation must be agreed upon in advance by teacher and evaluator.[5] The report goes on to assert that the criteria for evaluation and the traits to be judged should be agreed to and clearly understood by all parties before the process begins.[6]

---

[1]Shirley S. Bauer, "Perceptions of Selected Idaho Educators Regarding Supervision" (Ph.D. dissertation, University of Idaho, 1975), abstract.

[2]Farris, abstract.

[3]Young and Heichberger, p. 13.

[4]Ken P. Fraser, "Supervisory Behavior and Teacher Satisfaction" (Ed.D. dissertation, Montana State University, 1979), abstract.

[5]National School Public Relations Association, p. 56.

[6]Ibid.

In related studies, McNeil and Bolton provide results which support the NSPRA assertion. McNeil used 77 university students in a student teaching role for two days. He discovered that when a supervisor and a teacher agree ahead of time on what constitute evidence of success, that the supervisor views the teacher as having achieved greater success.[1] Similarly, from a review of research conclusions, Bolton is able to state that involving teachers in the development of evaluation criteria may improve the morale of the staff.[2]

Summarizing the studies related to activities prior to the classroom observation, the following three conclusions may be drawn:

(1) Teachers prefer to meet with the principal for a conference prior to a classroom observation.

(2) Teachers prefer to reach agreement with the principal on the time and place of the observation, prior to the observation.

(3) Teachers prefer to reach agreement with the principal on the criteria for the evaluation prior to the observation.

Activities Following Classroom Observation

In an unpublished "Occasional Paper" from Iowa State University College of Education, which summarized the application of research findings in the area of evaluation of teacher performance, Manatt in 1982 asserted that most writers in the area of teacher evaluation agree that the post-observation conference is the most important for changing teachers' behavior. At the same time, he cited major disagreements among

---

[1]McNeil, "Concomitants", p. 70.
[2]Bolton, p. 16.

those writers over the viability of a pre-observation conference.[1] This supervisory conference, borrowed from the appraisal conference in private business and industrial settings, has been the subject of numerous investigations.[2] A 1972 HEW report written by Bolton focused on the post-observation conference as an essential component of the teacher evaluation process, drawing several conclusions from research findings.[3] Among the important insights were the following: (1) Criticism builds defensiveness; (2) Praise has little effect on future productivity; (3) Mutual goal-setting improves performance; (4) Teachers accept decisions more readily if the focus is on improving performance and the situation.

A few studies have indicated that teachers prefer that the post-conference occur as part of the evaluation process. In her dissertation study, Riddile surveyed suburban teachers' perceptions of an evaluation by an objectives process in a large school district, and found that a postconference held soon after the observation contributed to teachers' perceptions of the effectiveness of the evaluation process.[4] Fraser's survey of Montana teachers discovered that 96 percent of the teachers in the study wanted supportive feedback from their supervisor after each observation visit.[5] Jones' dissertation study affirmed that the teachers who participated viewed the preconference and postconference in a

---

[1]Mannatt, pp. 3 and 5.
[2]Ibid. p. 5.
[3]Bolton, p. 31.
[4]Marianne Margrave Webb Riddile, "Teachers' Perceptions Of An Evaluation By Objectives Process in a Large Suburban School District" (Ed.D. dissertation, George Peabody College for Teachers, 1982), abstract.
[5]Fraser, abstract.

positive way.[1] When Jensen interviewed 46 primary and intermediate teachers, more than half of them recommended that teacher evaluation practices include an opportunity for the teacher to talk with the evaluator.[2]

The NSPRA summary of salient points from teachers' organizations across the country asserts that an opportunity must be provided for the teacher to consult with the evaluator before the evaluation report goes into the permanent file.[3] This, coupled with the above studies, leads to the conclusion that teachers prefer a postobservation conference.

Many of these same studies conclude that supportive comments following the observation are preferred by teachers, and that these comments should be based on observed strengths and weaknesses. The teachers in Riddile's study perceived the evaluation process to be more effective when such comments were made by the principal.[4] Fraser's study specifies that feedback should be supportive,[5] and Bolton concluded that teachers accept decisions more readily if the focus is on improving performance.[6] Statements by teachers' organizations in the NSPRA survey indicate that evaluation must take place in a constructive and nonthreatening atmosphere.[7] Bolton reminds us, however, that praise alone has very little effect on future productivity.[8] Nevertheless, the research studies indicate that teachers prefer that the principal provide supportive comments based on observed strengths and weaknesses following an observation.

---

[1] Sheridan Davis Jones, abstract.
[2] Jensen, p. 136.
[3] National School Public Relations Association, p. 57.
[4] Riddile, abstract.
[5] Fraser, abstract.
[6] Bolton, p. 31.
[7] National School Public Relations Association, p. 57
[8] Bolton, p. 31.

In addition to the postconference and the supportive comments, and most likely during the postconference, teachers prefer that the principal give them a copy of the evaluation report, and offer them an opportunity to reply to the report if they wish. In its summary of salient points from teachers' organizations, the NSPRA states that "opportunity must be provided for the teacher to see the evaluator's report... and to write a reply to the report if he wishes, that will be attached to it in his files." Teachers in Riddile's study indicated that the evaluation process was more effective if they received a written summary of the evaluation feedback.[1] Bolton maintained that open files of formal written evaluations are essential to an effective evaluation system, and that school districts should provide copies of all evaluation reports to the teacher.[2] In addition, as early as 1967, nearly 20 years ago, 63 percent of the 603 comprehensive agreements on file with the NEA Research Division, which covered teaching staffs in school systems enrolling 1,000 or more pupils, contained clauses on the evaluation of teachers, many of which included a requirement for the teacher's review of the written evaluation report and an opportunity to respond to any adverse comments.[3]

In summary, four conclusions have been drawn from the studies reviewed in this section:

(1) Teachers prefer to meet with the principal for a conference following the observation.

---

[1]Riddile, abstract.
[2]Bolton, p. 32.
[3]National Education Association, p. 73.

(2) Teachers prefer that during the postconference the principal provide supportive comments based on observed strengths and weaknesses.

(3) Following the conference, teachers prefer that the principal give them a copy of the evaluation report.

(4) Teachers prefer that the principal allow them to attach a reply to the evaluation report if they wish.

Competency of the Principal

Even though most teachers do not question that the principal should be responsible for their evaluation, a concern of many teachers is whether the principal is qualified to evaluate them.[1] The 1969 NEA survey of school systems discovered that the principal was the sole person to evaluate teachers in over half of the school systems responding.[2] The 1974 NSPRA survey reported similar findings.[3] In a dissertation study by Hickman, who surveyed teacher evaluation systems in Newfoundland and Labrador, it was determined that principals in the study were primarily responsible for the evaluation of teachers. But Hickman's study also discovered two major problems with this model: principals seem to lack sufficient time to devote to the evaluation process, coupled with a need for additional inservice.[4]

The importance of principal competence in the area of teacher evaluation has been noted by several studies. In an unpublished paper seeking to apply the findings of significant research studies to the process of teacher evaluation, Manatt has determined that one of the two major predictors of a successful evaluation conference is the supervisor's superior knowledge of teaching techniques based on an undergirding learning theory. The other is the fundamental interpersonal relationship

---

[1]Ibid, p. 71.
[2]National Education Association, p. 67.
[3]National School Public Relations Association, p. 10.
[4]George Augustus Hickman, "A Study of Teacher Evaluation Systems In The Province of Newfoundland and Labrador" (Ed.D. dissertation, University of Toronto, 1983), abstract.

between the teacher and the supervisor.[1] In a dissertation study by Barnes which surveyed principals and teachers in the state of Florida, the teachers who were surveyed thought that principals needed training in areas dealing with their interactions with teachers.[2] A case study by Fouke of one particular school district demonstrated that specific administrator training improved the quality of evaluation in the district.[3] A dissertation study by Ferguson discovered that principals in the state of Washington, while claiming to have adequate preparation for evaluating teachers, nevertheless felt that more training to strengthen teacher evaluation skills would be useful.[4] Furthermore, after surveying public and private schools in the state of Illinois, Meyer concluded that administrators who carry out evaluations should be better trained.[5]

Regarding the competency of the principal to evaluate teachers, four specific teacher concerns are apparent: (1) The principal should also be evaluated. The NSPRA, from its review of statements by teachers' organizations, indicates that those who do the evaluating should be evaluated reguarly, and that, "if teachers are to be evaluated, then all other

---

[1]Manatt, p. 14.

[2]Barbara Kininessi Barnes, "An Identification Of Perceived In-service Training Needs Of Florida Public School Principals Relative To Teacher Evaluation" (Ph.D. dissertation, Florida State University, 1983), abstract.

[3]Linda Graves Fouke, "A Comparative Study Of Teacher Evaluations Written By Administrators" (Ed.D. dissertation, Columbia University Teachers College, 1983), abstract.

[4]Victor Simon Ferguson, "Perceptions Concerning Teacher Evaluation In The State of Washington" (Ed.D. dissertation, Seattle University, 1981), abstract.

[5]D. Eugene Meyer, "A Survey Of Public And Private Schools In Illinois Regarding Faculty, Administration and Supportive Personnel Evaluations, Procedures, and Instruments Used," Illinois School Research and Development 18 (Spring 1982).

educational personnel should be evaluated, too, up to the highest levels of administration."[1] Bolton writes that "... only when systematic evaluation of evaluators (e.g., principals, supervisors) occurs will teachers more readily accept accountability functions."[2]

(2) The principal should be trained in evaluating teachers. The NSPRA reports this as a "must" from teachers' organizations.[3] Barnes' study of Florida principals' training needs concluded that Florida teachers thought their principals should be trained in their interactions with teachers during the evaluation process.[4] Bolton found that evaluators need to be trained to avoid allowing their personal biases and prejudices to effect the accuracy of observations.[5] In a survey of teachers and administrators in Western New York State and Ontario, Canada, Sapone discovered a real need for administrators responsible for appraisal and evaluation to have complete in-service training in appraisal and evaluation methods.[6]

(3) Teachers must have confidence in the principal's expertise in the areas which are to be evaluated. Paulin's survey of high school teachers in 7 schools revealed that the teachers in her study were more willing to be evaluated when they had confidence in the expertise of the evaluator.[7] In a 1975 article addressed to Canadian administrators, Pederson summarized several teacher concerns, among them the agreement that"the majority of those who do the observing and evaluating are

---

[1]National School Public Relations Association, p. 57.
[2]Bolton, p. 35.
[3]National School Public Relations Association, p. 57.
[4]Barnes, abstract.
[5]Bolton, p. 35.
[6]Sapone, p. 30.
[7]Paulin, p. 7.

academically and pedagogically unqualified to do so."[1] In Jensen's study, in which 46 primary and intermediate teachers were interviewed, 41 percent thought that the evaluator's teaching experience and "view of teaching" were influential factors in the effectiveness of teacher evaluation. The concern expressed was that perhaps the evaluators were too far and too long removed from the classroom.[2]

(4) Teachers prefer that the principal include another teacher on an evaluation team. Fraser, in his survey of Montana teachers, discovered that 54 percent of the teachers in the study would like for their supervisor to use group methods of supervision, including one or more colleagues skilled in the teacher's speciality.[3] Bolton suggests that various persons can help collect information for teacher evaluation, including a teacher's peers.[4] The NSPRA states that "Evaluation must not be done by just one person, but by a team, including at least one peer skilled in the teacher's specialty."[5]

---

[1]K. George Pederson, "Improving Teacher Effectiveness," Education Canada 15 (Fall 1975): 17.
[2]Jensen, p. 134 and 136.
[3]Fraser, abstract.
[4]Bolton,p. 25.
[5]National School Public Relations Association, p. 57.

Preferred Conditions of the Evaluation

Many authors have written about the problems which result from traditional teacher evaluation procedures. Most have identified the process as a threatening, negative, one-sided experience. Most of these authors were, at the same time, advocating an alternative model of teacher evaluation. In the wake of such advocacy, a few studies have attempted to determine whether, from the teacher's point of view, such methods are in fact superior, or at least preferred. From these studies, a list of nine preferred conditions of evaluation has been identified.

(1) Teacher evaluation should be formative, not summative. A dissertation study by Scandrett, in which teachers from a rural Missouri high school developed objectives which formed the basis of their evaluation, found that teachers preferred formative evaluation as opposed to summative evaluation.[1] The NSPRA, in its review of statements by many teachers' organizations across the country, listed several salient points, two of which are supportive of the idea of formative teacher evaluation:

> The purpose of teacher evaluation must be clearly understood to be improvement of instruction.
>
> Evaluation should be an on-going long-term process, that takes note of a teacher's overall performance and of progress between periods of evaluation.[2]

Bolton cited evidence in his report that teachers welcome evaluation if the major focus is on improving rather than fault-finding.[3]

---

[1]Terrence Glen Scandrett, "A Comparison of Summative and Formative Teacher Evaluation at a Rural Missouri High School" (Ed.D. dissertation, Saint Louis University, 1983), abstract.

[2]National School Public Relations Association, p. 57.

[3] Bolton, p. 3.

Ninety-three percent of the teachers in Young and Heichberger's survey of Western New York State elementary teachers responded that evaluation should focus attention on improving the performance of the teacher, rather than on tenure and termination decisions.[1]

(2) Teacher evaluation should be used to diagnose teacher performance. Eighty-seven percent of the teachers in Young and Heichberger's study supported this idea.[2] Pederson's report to Canadian administrators made the point that evaluation should try to become diagnostic rather than judgmental.[3]

(3) Teacher evaluation should be non-punitive, constructive, and non-threatening. Although these conditions are somewhat more esoteric than the previous two, they are supported by a number of documents and studies. During Jensen's interviews of elementary and intermediate teachers, more than half commented on the threatening aspects of evaluation. Over 70 percent wanted the evaluator to have a pro-attitude when he/she walked into the classroom. Many of the teachers discussed instances where teacher evaluation was threatening, and therefore less effective.[4]

The idea that teachers view evaluation as punitive, non-constructive, and threatening, and therefore ineffective, appears in numerous other articles and studies. The NSPRA survey of teachers' organizations asserts that when teachers view the evaluation process as punitive,

---

[1]Young and Heichberger, p. 16.
[2]Young and Heichberger, p. 11.
[3]Pederson, p. 18.
[4]Jensen, p. 137.

teacher effectiveness is lowered because teacher morale is lowered. Therefore, the report states that evaluation must take place in a constructive and non-threatening atmosphere.[1] Pederson states that teachers view evaluation as a unilateral exercise of power, and that the punitive nature of evaluation is seen as a threat to professional status and personal freedom.[2] Seventy percent of the teachers in Young and Heichberger's study reported that the supervisor is quite often seen as potentially dangerous to a teacher.[3]

(4) Teacher evaluation should be a cooperative, collegial effort. This concept is strongly supported by information in the NSPRA report, which states, "the teacher must feel that improvement of his performance is a cooperative effort involving him, his evaluators, and others on the school staff."[4] Fraser's survey of Montana teachers reports that 99 percent of the teachers in the study wanted such a collegial relationship.[5] Sixty-two percent of the teachers in Young and Heichberger's study wanted a helping relationship, while an additional 36 percent wanted colleagueship.[6] Several studies by Blumberg discovered that teachers feel supervisors do not treat them as collegial equals, as they would like to be treated.[7] Cole compared the use of collegial techniques for teacher evaluation with administrator-dominated techniques by pairing 32 schools. Although the principals in the study revealed no significant

---

[1]National School Public Relations Association, p. 57.
[2]Pederson, p. 17.
[3]Young and Heichberger, p. 14.
[4]National School Public Relations Association, p. 57.
[5]Fraser.
[6]Young and Heichberger, p. 16.
[7]Blumberg, p. 53.

difference in attitude toward teacher evaluation, teachers who were evaluated by collegial techniques had a much more positive attitude toward evaluation, and were more likely to feel that the evaluation would be used for the improvement of instruction.[1]

(5) Teacher evaluations should be based on mutual trust. In a study of the relationships between supervisory leadership and situational factors in which supervision takes place, Campbell found that teachers valued warmth, mutual trust, friendship, and respect.[2] Goldstein identified that a lack of mutual trust contributed to teachers' views of supervision as being too evaluative and not of a helping order.[3] Paulin's survey of high school teachers discovered an unwillingness to participate in the evaluation process when trust in the evaluator was lacking.[4] One of the qualities of a supervisor most often mentioned by the teachers in Young and Heichberger's study was honesty.[5]

(6) The evaluator should show concern for the teacher as a person. Although such a statement should be self-evident, the fact is that in many circumstances teachers do not feel so treated. Blumberg's study, which included 166 non-randomly chosen teachers, who reportedly represented a cross-section, reported that teachers see supervisors as not really understanding them and their problems.[6] Ninety-nine percent of the Montana

---

[1]Charles C. Cole, "A Comparison of Two Methods of Teacher Evaluation" (Ed.D. dissertation, North Texas State University, 1977), abstract.

[2]Ona L. Campbell, "The Relationship Between Eight Situational Factors and Low Scores on the Leadership Dimensions of Instructional Supervisors" (Doctoral dissertation, North Texas State College, 1961).

[3]William Goldstein, "An Enlightened Approach to Supervising Teachers," The Clearinghouse 46 (March 1972): 392.

[4]Paulin, p. 7.

[5]Young and Heichberger, p. 18.

[6]Blumberg, p. 53.

teachers in Fraser's study wanted the supervisor to show real concern for the teacher as a person.[1]

(7) Teacher evaluation should be an on-going process, not a "one-shot" procedure. The primary support for this statement comes from the summary of documents by the NSPRA, which asserts "Evaluation should be an on-going, long term process ... not a one-shot, stand-or-fall rating."[2] In his article "An Artistic Approach to Supervision", Eisner states, "the one-shot, 40 minute visit severely constrains what a supervisor is able to do, if for no other reason than the problem of establishing rapport." [3]

(8) Observations for teacher evaluation should be of adequate length and frequency. Riddile studied the effectiveness of evaluation by objectives, and found several factors which contributed to teachers' perceptions of effectiveness, which have been mentioned previously. But, in addition to those factors, certain behaviors of the principals seemed also to contribute to a perception of effectiveness. Primary among those was the amount of time spent observing the teacher.[4] Houston's study in Tennessee discovered a discrepency between teachers and administrators regarding preferred length of observations and frequency of observations, with teachers reporting a higher level of satisfaction with the evaluation process as classroom observations occurred with more frequency and length.[5]

---

[1]Fraser, abstract.
[2]National School Public Relations Association, p. 57.
[3]Elliot W. Eisner, "An Artistic Approach to Supervision," in Supervision of Teaching, ed. Thomas J. Sergiovanni (Alexandria, Va.: Association for Supervision and Curriculum Development, 1982), p. 61.
[4]Riddile, abstract.
[5]Houston, abstract.

The 46 teachers interviewed by Jensen mention time spent on evaluation more often (85 percent) than any other factor which can facilitate or hinder the evaluation process.[1]

(9) Teacher evaluation should be part of a comprehensive plan of career development and improvement of teaching performance. Several of the studies already mentioned are strongly supportive of this idea. Specifically, the summary of teachers' association statements by the NSPRA asserts that "Evaluation ... must go hand in hand with a comprehensive plan of career development and improving total teacher performance ..."[2] In Frasers' study of Montana teachers, 96 percent wanted to participate in in-service education designed specifically to help them teach more effectively.[3]

Purposes and Criteria of Evaluation

Much has been said about process, and successful evaluation certainly depends upon an appropriate process. But valid criteria upon which to evaluate performance is also important. In an article discussing how to improve teacher evaluation, Musella has concluded that the development of criteria is a prerequisite to any assessment. The criteria need not be universally accepted; it is enough that both parties are aware of the criteria and accept the criteria.[4] Bolton says: "When decisions regarding the development of criteria are based upon empirically supported and rational considerations, relevance and useability are more likely to be

---

[1]Jensen, p. 133.
[2]National School Public Relations Association, p. 57.
[3]Fraser, abstract.
[4]Musella, p. 19.

ensured."[1] In an article directed at principals, Crenshaw and Hoyle state that evaluation based on predetermined performance criteria and expected levels of performance are valid and should be one of the principal methods of teacher evaluation."[2]

In an attempt to organize the criteria by which teacher effectiveness may be judged, Mitzel divides them into 3 classes: (1) pre-sage (I.Q., NTE scores, degree status, hours in education, marital status, etc.); (2) process measures (what happens in the classroom); (3) product measures (change that occurs in students as a result of what happens in the classroom).[3] The NSPRA lists 6 attributes of the successful teacher that most districts presently attempt to assess: teacher-pupil relationships, classroom management and procedure, staff relationships, community relationships, professional attributes, and professional growth.[4] The Michigan Education Association comments on 3 environments in which teachers are often evaluated: community, institution, and classroom. Their position is that behavior in the community is not an appropriate criterion for formal evaluation of teacher performance. The Association recommends that a teacher's technical competence in the classroom be assessed in terms of (1) planning and organizing in relation to stated goals, (2) knowledge of the subject matter, (3) methodology, (4) classroom control, (5) client relationships, and (6) the management and condition of the milieu.[5]

---

[1]Bolton, p. 17.
[2]Crenshaw and Hoyle, p. 42
[3]H. E. Mitzel, "Teacher Effectiveness," in the Encyclopedia of Educational Research, 3rd., ed. Chester W. Harris (New York: Macmillan, 1960), pp. 1483-1484.
[4]National School Public Relations Association, p. 11.
[5]Ibid., p. 12.

In an article which purports to describe teachers' feelings about evaluation, Wolf asserts that teachers believe that the criteria for evaluating effective teaching are too vague and ambigious to be of worth. Current appraisal techniques, in their opinion, do not collect information which accurately characterizes performance.[1] In addition, according to Wolf, teachers see the idiosyncrasies of the rater interfering with the evaluation. In that same direction, Morrison notes that evaluation implies a judgment of worth, which derives from the values of the judge, and that much of the controversy surrounding observation is the result of differences in educational value judgments.[2] Of course, such value judgments lead to bias, which is unquestionably a problem when developing criteria for evaluation.

Musella asserts that bias introduced by the perceptual-cognitive view of the rater is a problem in evaluation which limits the usefulness of formal evaluative criteria.[3] Christner found bias in administrators, in spite of many hours of workshop training on observation techniques, examining lesson plans, and other objective forms of data gathering.[4] He also discovered problems with stereotypes.[5] Holley and Hickman note that these kinds of evaluator unreliability lead to evaluation documents which provide false or dubious information concerning practices which they purport to assess.[6] A study by Start, which rated supervisor and teacher

---

[1]Wolf, p. 160.

[2]Edward J. Morrison, "Performance Observation as an Approach to Teacher Evaluation," Educational Horizons 52 (Summer 1974): 170.

[3]Musella, pp. 18-19.

[4]A. Christner, "The J.R. Syndrome: Administrator Bias in Teacher Evaluation," presented at the Annual Meeting of the American Educational Research Association, April 1981, p. 9.

[5]Ibid., p. 8.

[6]Holley and Hickman, p. 5.

personalities on sixteen factors, found that the highest ratings of teaching ability were granted to those teachers whose profiles were either similar to or very different from that of the head teachers.[1]

In developing specific criteria for teacher evaluation, the literature is not clear as to which criteria should be used. Burnett could not find reliability regarding any of the following five presage criteria: (1) academic success in college, (2) professional references, (3) aggregate of professional education classes, (4) college achievement and professional education courses, and (5) student teaching courses.[2] Jensen's study, which had 46 teachers listing criteria for teacher evaluation, found the most common criterion to be classroom atmosphere, tone or climate.[3] Other important aspects of evaluation included: (1) rapport with children, (2) planning, (3) involvement in the faculty, (4) communication with parents, (5) knowledge of the subject, (6) physical organization and appearance of the classroom.[4]

Since many teacher evaluation forms include personality traits, Crenshaw and Hoyle looked at research on the topic, and found nothing to indicate a strong cause and effect relationship between personality traits and teaching effectiveness.[5] Musella surveyed several studies and also found little evidence that certain personality characteristics are more desireable than others for teaching in general.[6] Stephens cites a

---

[1]Start, abstract.

[2]Richard Mulkey Burnett, "An Evaluation of Certain Criteria Used in Selecting Teachers for Public Schools" (Ed.D. dissertation, University of Texas, 1966), abstract.

[3]Jensen, p. 135.

[4]Ibid.

[5]Crenshaw and Hoyle, p. 41.

[6]Musella, p. 17.

monograph by Tarveggia and Dubin called The Teaching Learning Paradox, which analyzed the data of almost one hundred comparison studies of different teaching methods, and found "no shred of evidence to indicate any basis for preferring one teaching method over another as measured by the performance of students on course examinations."[1] A more promising report from the University of North Carolina at Chapel Hill, however, offers 26 specific criteria of effective teaching, each of which is supported by several quality research studies as being associated with measurable student performance gains.[2]

For the purposes of this study, four conclusions regarding the basis for teacher evaluation have been drawn: (1) Teachers prefer that evaluation results should not be used for tenure and termination decisions, but be limited to the improvement of instruction. The NSPRA collection of documents from teachers' organization states that "The purpose of teacher evaluation must be clearly understood to be improvement of instruction, not for formal, legalistic purposes of firing, determination of tenure, salary, and promotion."[3] In the survey by Young and Heichberger, 93 percent of the responding teachers agreed that evaluation should focus on improvement rather than tenure and termination decisions.[4]

Teachers prefer that evaluation include a goal-setting process. Sapone's survey of 70 administrators and their teachers determined that teachers in the survey ranked the goal-setting process as their most important priority.[5] Seventy percent of the teachers surveyed by Young

---

[1]J. M. Stephens, The Process of Schooling, (New York: Holt, Rinehart, 1967), p. 72.
[2]Group for the Study of Effective Teaching, p. 122.
[3]National School Public Relations Association, p. 57.
[4]Young and Heichberger, p. 16.
[5]Sapone, p. 29.

and Heichberger indicated that the process of establishing goals and working together to evaluate those goals was the most desired form of supervision.[1] A dissertation study by Carpenter surveyed elementary teachers in an elementary district comprised of 17 schools. The teachers in the study preferred a goal-setting evaluation process to the descriptive process then in use.[2]

(3) Teachers prefer that the principals not use a subjective survey sheet to evaluate performance. The teachers in Carpenter's study preferred not to be evaluated by a descriptive method.[3] The American Federation of Teachers passed a resolution at its 1973 convention stating that "any scaled rating of teachers nurtures the exercise of political pressure and creates disharmony among the members of a school's staff.[4] Pederson states that "at best, (survey sheets are) only highly subjective appraisals of skills considered relevant by the observer.[5]

(4) Teachers prefer that evaluation be based on objective, substantive criteria. Seventy percent of the teachers interviewed by Jensen mentioned criteria and evidence for evaluation as significant factors in teacher evaluation.[6] Bolton cites research evidence that criteria should be based on empirically supported and rational considerations.[7] Wolf indicates that teachers believe that the standards for evaluating what is effective teaching are too vague and ambitous to be worth anything.[1]

---

[1]Young and Heichberger, p. 16.
[2]Dennis Arthur Carpenter, "A study of Teachers' Perceptions of Two Teacher Evaluation Processes" (Ed.D. dissertation, George Peabody College, 1981), abstract.
[3]Ibid.
[4]National School Public Relation Association, p. 16.
[5]Pederson, p. 17.
[6]Jensen, p. 134.
[7]Bolton, p. 17.
[8]Wolf, p. 160.

CHAPTER III

PRESENTATION AND ANALYSIS OF THE DATA

The research questions to be answered by this study were:

1. Do principals, in their evaluation of teachers, include procedures that are congruent with practices that have been identified in the literature as being associated with increased teacher satisfaction with evaluation?

2. Is there a difference between what a principal would ideally do and what he/she actually does in the process of evaluating teachers?

3. What effect, if any, does each of the following have on research questions 1 and 2?

   A. Number of years as a teacher.
   B. Number of years as a principal.
   C. Number of years in education.
   D. Highest degree earned.
   E. Staff size.
   F. Average income of families in the community.

4. Are district-approved teacher evaluation policies and procedures congruent with practices that have been identified in the literature as being associated with increased teacher satisfaction with evaluation?

To answer these questions, the information from Chapter II regarding teacher-preferred characteristics of evaluation was compiled and organized. In Chapter II, a review of 30 articles revealed 28 characteristics of evaluation which result in increased teacher satisfaction with evaluation of their performance. These characteristics were organized into 6 clusters, as follows:

I Teacher participation in evaluation decisions.

- I-1 Teachers prefer to participate in designing the evaluation instrument.
- I-2 Teachers prefer to participate in designing the evaluation process.
- I-3 Teachers prefer to participate in shaping the policies which govern evaluation.
- I-4 Teachers prefer to participate in establishing the goals of evaluation.

II Activities prior to classroom observation.

- II-1 Teachers prefer to meet with the principal for a conference prior to an observation.
- II-2 Teachers prefer to establish the time and place of the observation prior to its occurrence.
- II-3 Teachers prefer to establish the criteria for the evaluation with the principal prior to its occurrence.

III Activities following the classroom observation.

- III-1 Teachers prefer to meet with the principal for a conference following the observation.
- III-2 Teachers prefer that during the postconference the principal provide supportive comments based on observed strengths and weaknesses.
- III-3 Following the conference, teachers prefer that the principal give them a copy of the evaluation report.
- III-4 Teachers prefer that the principal allow them to attach a reply to the evaluation report if they wish.

IV Training and competence of the evaluator.

- IV-1 Teachers prefer that the principal also be evaluated by his or her supervisor.
- IV-2 Teachers prefer that the principal receive in-service training in the evaluation procedures to be used.
- IV-3 Teachers prefer that the principal have expertise in the areas to be evaluated.
- IV-4 Teachers prefer that the principal include another teacher on an evaluation "team".

V Conditions of the evaluation.

- V-1 Teachers prefer that evaluation be more formative than summative.
- V-2 Teachers prefer that evaluation be diagnostic.
- V-3 Teachers prefer that evaluation be non-threatening.

V-4 Teachers prefer that evaluation be a cooperative effort with the principal.
V-5 Teachers prefer that there be mutual trust between the principal and the teacher.
V-6 Teachers prefer to be viewed more as a person than as an employee during the evaluation process.
V-7 Teachers prefer that evaluation be an on-going process.
V-8 Teachers prefer that evaluation be based on adequate frequency of observations.
V-9 Teachers prefer that evaluation be part of a plan for the improvement of total teaching performance.

VI Purposes and Criteria of the evaluation.

VI-1 Teachers prefer that the evaluation results not be used for tenure and termination decisions, but be limited to improving instruction.
VI-2 Teachers prefer that evaluation include a goal-setting process.
VI-3 Teachers prefer that the principal not use a subjective survey sheet to evaluate performance.
VI-4 Teachers prefer that evaluation be based on objective, substantive criteria.

The Questionnaire

The questionnaire which was developed from the 28 items in the 6 clusters included, in addition to 7 demographic and personal information queries, a total of 92 responses, divided into 2 sections. The first section asked the principal to respond according to present practice; the second section asked the principal to respond to an identical question according to his/her opinion of "ideal" practice. Of the 92 responses, 54 specifically referred to one of the identified preferred characteristics of evaluation (Table 3.1). Twenty-two additional items were included on the questionnaire to provide alternative responses, while 16 provided additional information of interest to the study.[1]

[1] A copy of the survey instrument appears in Appendix A.

Response

A total of 138 questionnaires were mailed to elementary principals, of which 81 were returned, yielding a response rate of 59 percent. Identical instruments had been mailed to a pilot group of 13, but the data from that group was not included in the results. An additional group of 22 principals had been interviewed face-to-face, but the results, though of interest, were considered incomparable to the mailed questionnaire, and were also not included in the analysis.

On most of the 28 scaled items, the interviews produced a higher percentage of "1" responses ("always" for present practice; "very important" for ideal practice) then did the mailed survey. Similarly, the interviews produced a higher percentage of responses on most of the alternatives provided in questions #15 to #17 and #33 to #35, and tended more toward the extremes in each of the 11 characteristics in questions #18 and #36. Because of this tendency, the interview responses were separated from the survey responses for purposes of analyzing the data. The presentation of the questionnaire responses which follows is drawn from the 81 principals who responded to the mailed questionnaire.

Presentation of Questionnaire Responses

The questionnaire presented 47 pairs of identical questions, one for present practice, an identical one for the respondent's opinion of "ideal" practice.

In this section, the responses to each pair of identical questions are presented, along with an analysis of those responses. For each pair of responses, a t -test was run on the difference between the ideal and

the present practice responses, to determine whether the difference was statistically significant. The analysis following each identical pair of questions is based on the sum of the frequencies of "usually" and "always" for present practice, and the sum of the frequencies of "moderately important" and "very important" for ideal practice, unless indicated otherwise.

Demographic and Personal Information

The first 7 questions (A through G) relate to personal and demographic information.

Question A:	How many years were you a teacher?	
Response:	1 through 5 years	22.2%
	6 through 10 years	50.6%
	over 10 years	24.7%
	no response	2.5%
		100.0%

Over 75 percent of the responding principals were teachers for more than 5 years before they became principals. Twenty-five percent were teachers for more than 10 years before assuming a principal's post. Thus, a large number of responding principals were on the other side of the evaluation table for a number of years before becoming a principal.

Question B:	How many years have you been a principal?	
Response:	1 through 5 years	22.2%
	6 through 10 years	27.2%
	over 10 years	48.1%
	no response	2.5%
		100.0%

Almost 80 percent of the responding principals have been a principal for longer than 5 years, while nearly 50 percent have held such posts for more than 10 years, revealing a high level of administrative experience among the respondents, in addition to several years of teaching experience.

Question C: How many total years have you worked in education?

Response:		
	9 through 15 years	21.0%
	16 through 25 years	46.9%
	over 25 years	29.6%
	no response	2.5%
		100.0%

All of the responding principals in this study have been in education for at least 9 years. Nearly 80 percent have been educators for more than 15 years, indicating that the responses to this questionnaire have come from a highly experienced group of educators.

Question D: What is the highest degree you have attained?

Response:		
	M.A.	34.6%
	M.A. + 30 hours	28.4%
	M.A. + 45 hours	22.2%
	Doctorate	12.3%
	no response	2.5%
		100.0%

Of the 81 principals in the study, 28 hold a Master's Degree, 23 have 30 hours beyond a Master's, 18 have 45 hours beyond a Master's, and 10 have earned a Doctorate. Thus, approximately two-thirds of the responding principals have pursued education beyond a Master's Degree.

In summary, the overwhelming majority of the participants in this study were both highly educated and experienced educators.

Question E: Please estimate the average household income of the community served by your school.

Response:		
	Under $15,000	16.0%
	$15,000 - $30,000	54.3%
	$30,000 - $45,000	17.3%
	over $45,000	7.4%
	no response	4.9%
		100.0%

The districts included in the study form a fairly standard distribution by family income. Over half are in the range of $15,000 to $30,000. Only 16 percent are below $15,000, while approximately 25

percent are over $30,000. A review of the respondents has indicated that all geographic sections in the study area were represented in the survey responses, thus validating the reported distribution of income.

Question F: Of the families in the community served by your school. approximately what percentage falls in each of the above income categories?

Response:		
	Under $15,000	mean 27.8%
	$15,000 to $30,000	mean 36.1%
	$30,000 to $45,000	mean 19.7%
	over $45,000	mean 14.7%
	no response	15

Throughout the study area, the distribution of incomes is substantially similar to the results obtained in question "E". Given the large number who did not respond to this question, however, the results were not used for further analysis.

Question G: What is the size of your teaching staff?

Response:		
	Under 16	14.8%
	16 to 25	49.4%
	over 25	33.3%
	no response	2.5%
		100.0%

Over 80 percent of the responding principals work in schools with more than 16 staff members, while nearly half serve school with a staff size of 16 to 25. Only 15 percent work in schools with less than 16 staff members.

Summary

Over 75 percent of the responding principals were teachers for at least 5 years, while nearly 80 percent have been a principal for more than 5 years, and in the field of education for more than 15 years. Appraximately 65 percent have pursued education beyond a Master's Degree, and over 80

percent work in schools with more than 15 staff members. The schools represented in the study form a well-balanced economic cross-section.

Survey Responses

Question H: Have you received training in evaluating teachers?

This question, and the one which follows, were written in a unique format, in which, instead of a pair of identical questions, the answer to one question implies either present or ideal practice. The format for the answer and the original response frequency was:

79.0%	yes
4.9%	no
13.6%	no, but I would like to

To tally the results, all "yes" answers were considered present practice. All "no" answers and "no, but I would like to" answers were considered "no" for present practice. The "no, but I would like to" answers were considered "yes" for ideal practice, and when added to the "yes" answers, the sum became the totals of "yes" for ideal practice.

This procedure is admittedly less than ideal, because respondents who answered "yes" were not given the opportunity to indicate whether, in their opinion, such training is important to ideal practice, as they were on other questions. After some consideration, however, it was decided to include the information in the study results and, for ease of analysis, to change the format of the answer as follows:

Response:		Present	Ideal
1	Yes	79.0%	92.6%
2	No	18.5%	4.9%
	no response	2.5%	2.5%
		100.0%	100.0%
		$\bar{X}$=1.19;s.d.=40	$\bar{X}$=1.05;s.d.=0.22

Nearly 80 percent of the principals responding indicated that they have received training in teacher evaluation. An additional 11 principals who have not received training would like to, leaving only 4 principals who have not received training and who see no need for it. Applying the _t_ -test, the difference between ideal and present practice is significant ($t=3.55$; $p < .001$).

Teachers have indicated that they prefer the principal to have in-service training in the evaluation procedures to be used. Principals here have indicated that they are trained in evaluation, though no specific evaluation procedure was mentioned. Nevertheless, the high response rate for this question indicates high congruence.

Question I: Are you usually evaluated by your supervisor?

The format of this question is identical to the previous question. The format of the answer and the original response frequency was:

87.7% yes
2.5% no
4.9% no, but I would like to

Combining the answers as in Question "H", the reformatted response is:

Response:		Present	Ideal
1	Yes	87.7%	92.6%
2	No	7.4%	2.5%
no response		4.9%	4.9%
		100.0%	100.0%
		$\bar{X}=1.08$; s.d.=0.27	$\bar{X}=1.03$; s.d.=0.16

Almost 88 percent of the principals in the study reported that they are evaluated by their supervisor. Only 2 principals are not evaluated and do not wish to be, placing this item in high congruence with teachers' preferences, both in present practice and in ideal practice. The difference

between present and ideal practice, applying the t-test, is significant ($t=2.04; p<0.5$).

NOTE: All further questionnaire items are presented in pairs, the first question representing present practice, the second representing ideal practice. In those cases where 2 or more pairs of questions refer to a common evaluation characteristic, they are presented in a group.

Question J: How many observations do you usually make when you evaluate a teacher?
Question #37: In your professional opinion, how many formal observations is ideal when evaluating a teacher?

Response:	Present	Ideal
1	19.8%	4.9%
2	29.6%	29.6%
3 or more	48.1%	59.3%
no response	2.5%	6.2%
	100.0%	100.0%
	$\bar{X}=2.30$; s.d.=0.77	$\bar{X}=2.57$; s.d.=0.60

Question K: What is the average length of a classroom observation when you evaluate a teacher?
Question #38: In your professional opinion, what is the ideal length of a classroom observation?

Response:	Present	Ideal
1 10 minutes	4.9%	1.2%
2 20 minutes	8.6%	7.4%
3 30 minutes	27.2%	18.5%
4 40 minutes	30.9%	39.5%
5 50 minutes	7.4%	9.9%
6 60 minutes	16.0%	14.8%
7 over 60 minutes	2.5%	3.7%
no response	2.5%	4.9%
	100.0%	100.0%
	$\bar{X}=3.92$; s.d.=1.43	$\bar{X}=4.15$; s.d.=1.32

Although no standard has been established, for the purposes of this study an observation of at least 40 minutes, repeated at least 3 times, is considered adequate, although local conditions might vary. Nearly half of the principals in the study presently include 3 or more observations in their evaluation of a teacher, and almost 60 percent observe an average of 40 minutes or more.

In ideal practice, nearly 60 percent of the principals would make at least 3 observations, a difference of more than 11 percentage points. Almost 70 percent would observe at least 40 minutes, also a difference of more than 11 percentage points. Applying the t -test to the number of observations, the difference between ideal and present practice is significant ($t=1.85$; $p < .07$)

Question #1 and #19: Classroom observation is a significant part of the process of evaluating teachers.

Response:		Present			Ideal
1	Always	72.8%	1	Very Important	72.8%
2	Usually	18.5%	2	Moderately important	24.7%
3	Sometimes	7.4%	3	Midly important	2.5%
4	Seldom	1.2%	4	Mildly unimportant	0.0%
5	Rarely	0.0%	5	Moderately unimportant	0.0%
6	Never	0.0%			0.0%
		100.0%			100.0%
		$\bar{X}=1.37$; s.d.=0.68			$\bar{X}=1.30$; s.d.=0.51

Over 90 percent of the respondents indicated that classroom observation is usually or always part of the evaluation process in their schools. Nearly 98 percent of the principals in the survey indicated that classroom observation is moderately or very important in an ideal evaluation system. Applying a t -test, the difference between ideal and present practice is not significant, due in part to the ceiling effect created by the high responses to present practice, as well as the small difference between the means.

Question #2 and #20. Before a classroom observation, the teacher and supervisor agree on the time and place of the observation.

Response:	Present		Ideal
1 Always	43.2%	1 Very Important	34.6%
2 Usually	17.3%	2 Moderately Important	21.0%
3 Sometimes	18.5%	3 Mildly Important	17.3%
4 Seldon	1.2%	4 Mildly Unimportant	6.2%
5 Rarely	13.6%	5 Moderately unimportant	12.3%
6 Never	6.2%	6 Very unimportant	8.6%
	100.0%		100.0%
	$\bar{X}$=2.43;s.d.=1.65		$\bar{X}$=2.67;s.d.=1.69

Over 60 percent of the principals surveyed usually or always established the time and place of the observation with the teacher, prior to classroom observation. Only 55 percent of the principals, however, indicated that this practice is moderately or very important in an ideal evaluation model, making this item one of the few which actually showed a negative difference from present practice to ideal practice. Applying the t -test, the difference only approaches significant (t=1.92;p <.60).

Question #3 and #21. Before a classroom observation, the teacher and and supervisor agree on the criteria for the observation.

Response:	Present		Ideal
1 Always	48.1%	1 Very Important	50.6%
2 Usually	14.8%	2 Moderately Important	19.8%
3 Sometimes	14.8%	3 Mildly Important	14.8%
4 Seldom	7.4%	4 Mildly unimportant	2.5%
5 Rarely	9.9%	5 Moderately Unimportant	6.2%
6 Never	3.7%	6 Very Unimportant	2.5%
	100.0%		100.0%
	$\bar{X}$=2.23;s.d.=1.53		$\bar{X}$=2.05;s.d.=1.47

Question #4 and #22. Before the observation, the teacher and supervisor agree on the instrument to be used in the evaluation.

Response:	Present		Ideal
1 Always	85.2%	1 Very Important	79.0%
2 Usually	0.0%	2 Moderately Important	0.0%
3 Sometimes	4.9%	3 Mildly Important	4.9%
4 Seldom	0.0%	4 Mildly Unimportant	3.7%
5 Rarely	1.2%	5 Moderately Unimportant	7.4%
6 Never	8.6%	6 Very unimportant	1.2%
no response	0.0%	no response	1.2%
	100.0%		100.0%
	$\bar{X}$=1.65;s.d.=1.5		$\bar{X}$=1.9;s.d.=1.54

These two pairs of questions requested very similar information. The specific preference of teachers, according to the literature, is to

agree on the criteria of the evaluation prior to an observation. Although it could be argued that agreeing on the instrument establishes the criteria, some instruments are simply narrative forms, and include very little criteria. Therefore, the more important response for the purpose of this study is the first one. Over 60 percent of the principals in the survey usually or always establish the criteria for the evaluation prior to observing in the calssroom. Over 70 percent consider this step moderately or very important to an ideal evaluation system. Applying the t -test, this difference is not significant. On the other hand, over 85 percent usually or always establish the instrument to be used, while only 79 percent indicate that such action is moderately or very important in an ideal evaluation situation. This difference approaches significance ($t=1.71; p < .09$).

Question #5 and #23: Before the observation, the supervisor meets with the teacher for a pre-observation conference.

Response:	Present		Ideal
1 Always	29.6%	1 Very Important	33.3%
2 Usually	11.1%	2 Moderately Important	21.0%
3 Sometimes	25.9%	3 Mildly Important	17.3%
4 Seldom	9.9%	4 Mildly Unimportant	6.2%
5 Rarely	11.1%	5 Moderately Unimportant	11.1%
6 Never	11.1%	6 Very Unimportant	9.9%
no response	1.2%	no response	1.2%
	100.0%		100.0%
	$\bar{X}=2.95$; s.d.=1.71		$\bar{X}=2.70$; s.d.=1.71

The pre-observation conference is an integral part of the clinical supervision model and the MBO model. It is advocated by a number of writers, and research indicates that teachers prefer that it be part of the evaluation process. Nevertheless, only about 41 percent of principals surveyed usually or always hold such conferences. In ideal practice, over 54 percent indicate the preconference is moderately or very important, a difference

of nearly 14 percentage points. Applying a *t*-test, however, the difference between ideal and present practice is not significant.

Question #6 and #24. Following the observation, the supervisor provides the teacher with a copy of the evaluation report.

Response:	Present		Ideal
1 Always	92.6%	1 Very Important	93.8%
2 Usually	4.9%	2 Moderately Important	3.7%
3 Sometimes	1.2%	3 Mildly Important	0.0%
4 Seldom	0.0%	4 Mildly Unimportant	1.2%
5 Rarely	0.0%	5 Moderately Unimportant	0.0%
6 Never	1.2%	6 Very Unimportant	1.2%
	$\bar{X}$=1.14;s.d.=0.63		$\bar{X}$=1.14;s.d.=0.67

Nearly 100 percent of the principals in the survey reported that they usually or always provide the teacher with a copy of the evaluation report following an observation. An almost identical number indicated that doing so is moderately or very important to an ideal evaluation system. As a consequence, congruence between principals' practice and teachers' preference is very high on this item, both in present practice and in ideal practice. Because of the ceiling effect, the difference between ideal and present practice is not significant.

Question #7 and #25. Following the observation, the teacher and supervisor hold an evaluation conference.

Response:	Present		Ideal
1 Always	81.5%	1 Very Important	88.9%
2 Usually	14.8%	2 Moderately Important	9.9%
3 Sometimes	2.5%	3 Mildly Important	0.0%
4 Seldom	0.0%	4 Mildly Unimportant	0.0%
5 Rarely	0.0%	5 Moderately Unimportant	0.0%
6 Never	1.2%	6 Very Unimportant	1.2%
	100.0%		100.0%
	$\bar{X}$=1.26;s.d.=0.70		$\bar{X}$=1.16;s.d.=0.62

Ninety-six percent of the principals in the survey usually or always hold a post-observation conference. Ninety-nine percent consider a post-observation conference to be moderately or very important in an ideal evaluation system. This item produces high congruence with teachers' pre-

ferences, both in actual practice and in ideal practice. Because of the ceiling effect, the difference between ideal and present practice is not significant.

Question #8 and #26: The teacher is allowed to attach a reply to the evaluator's report.

Response:		Present			Ideal
1	Always	91.4%	1	Very Important	87.7%
2	Usually	4.9%	2	Moderately Important	8.6%
3	Sometimes	1.2%	3	Mildly Important	2.5%
4	Seldom	0.0%	4	Mildly Unimportant	0.0%
5	Rarely	2.5%	5	Moderately Unimportant	0.0%
6	Never	0.0%	6	Very Unimportant	1.2%
		100.0%			100.0%
		$\bar{X}$=1.17;s.d.=0.69			$\bar{X}$=1.20;s.d.=0.68

Ninety-six percent of the principals who responded usually or always allow the attachment of a reply to the evaluation report, and 96 percent consider it moderately or very important. This item results in high congruence between principals and teachers, both in present practice and in ideal practice. Because of the ceiling effect, the difference between ideal and present practice is not significant.

Question #9 and #27. During the evaluation process, the evaluator provides supportive comments based on observed strengths and weaknesses.

Response:		Present			Ideal
1	Always	82.7%	1	Very Important	90.1%
2	Usually	14.8%	2	Moderately Important	7.4%
3	Sometimes	1.2%	3	Mildly Important	0.0%
4	Seldom	0.0%	4	Mildly Unimportant	0.0%
5	Rarely	0.0%	5	Moderately Unimportant	0.0%
6	Never	0.0%	6	Very Unimportant	1.2%
no response		1.2%	no response		1.2%
		100.0%			100.0%
		$\bar{X}$=1.18;s.d.=0.41			$\bar{X}$=1.14;s.d.=0.61

Ninety-eight percent of principals in the survey usually or always provide supportive comments to teachers based on observed strengths and weaknesses during the evaluation process. Ninety-eight percent indicated that such comments are at least moderately important to ideal

evaluation situations, with over 90 percent responding that supportive comments are very important. Principals and teachers are highly congruent on this item, both in ideal and present practice. Because of the ceiling effect the difference between ideal and present practice is not significant.

Question #10 and #28. The overall results of the teacher evaluation process lead to specific in-service training designed to improve classroom teaching.

Response:		Present			Ideal
1	Always	8.6%	1	Very Important	44.4%
2	Usually	28.4%	2	Moderately Important	37.0%
3	Sometimes	40.7%	3	Mildly Important	13.6%
4	Seldom	9.9%	4	Mildly Unimportant	1.2%
5	Rarely	9.9%	5	Moderately Unimportant	1.2%
6	Never	1.2%	6	Very Unimportant	1.2%
no response		1.2%	no response		1.2%
		100.0%			100.0%
		$\bar{X}$=2.87;s.d.=1.13			$\bar{X}$=1.80;s.d.=0.97

The difference between ideal and present practice is quite dramatic on this item. In actual practice, only 37 percent of the respondents usually or always provide in-service training based on the evaluation. In ideal practice, 82 percent of the respondents indicated that specific in-service training which follows the evaluation is moderately or very important. Applying the t-test, the difference between present practice and ideal practice is significant (t=8.55;p <.0001).

Question #11 and #29: The Supervisor is knowledgeable in the academic area he/she is expected to evaluate.

Response:		Present			Ideal
1	Always	23.5%	1	Very Important	38.3%
2	Usually	60.5%	2	Moderately Important	32.1%
3	Sometimes	12.3%	3	Mildly Important	19.8%
4	Seldom	1.2%	4	Mildly Unimportant	4.9%
5	Rarely	0.0%	5	Moderately Unimportant	2.5%
6	Never	0.0%	6	Very Unimportant	1.2%
no response		2.5%	no response		1.2%
		100.0%			100.0%
		$\bar{X}$=1.92;s.d.=0.64			$\bar{X}$=2.05;s.d.=1.16

While 84 percent of the principals in the survey indicated that they usually or always feel confident about their knowledge of the academic subject being evaluated, only about 70 percent indicated that such expertise is moderately or very important, although nearly 40 percent rated it as very important. Applying the *t* -test, the difference between present and ideal practice is not significant.

Question #12 and #30: Teacher evaluation is done by a "team", including one person skilled in the teacher's specialty.

Response:	Present		Ideal
1 Always	0.0%	1 Very Important	7.4%
2 Usually	2.5%	2 Moderately Important	22.2%
3 Sometimes	9.9%	3 Mildly Important	22.2%
4 Seldom	9.9%	4 Mildly Unimportant	9.9%
5 Rarely	17.3%	5 Moderately Unimportant	17.3%
6 Never	59.3%	6 Very unimportant	18.5%
no response	1.2%	no response	2.5%
	100.0%		100.0%
	$\bar{X}$=5.26;s.d.=1.12		$\bar{X}$=3.67;s.d.=1.63

Although surveys of teachers have been somewhat inconclusive regarding teachers' preferences as to who should evaluate them, several studies have found that teachers prefer to be evaluated when a peer who is skilled in their specialty is included. Only 2.5 percent of principals in the present study indicated that this usually or always happens. In ideal practice, nearly 30 percent regard team evaluation as moderately or very important. Nevertheless, principals and teachers are not congruent on this characteristic of evaluation, due probably to the fact that only elementary principals were surveyed. There are, after all, only a few "specialities" in the elementary schools. Moreover, elementary principals were probably originally elementary teachers, giving them confidence in their own expertise to evaluate in all or most areas. (Remember that 84 percent indicated they usually were knowledgeable in the

academic area to be evaluated.) The difference between present practice and ideal practice on this item, when the *t* -test is applied, is significant ($t=9.92$; $p < .0001$).

Question #13 and #31. Teacher evaluation is based on objective, substantive criteria.

Response:		Present			Ideal
1	Always	43.2%	1	Very Important	54.3%
2	Usually	42.0%	2	Moderately Important	32.1%
3	Sometimes	6.2%	3	Mildly Important	7.4%
4	Seldom	2.5%	4	Mildly Unimportant	2.5%
5	Rarely	2.5%	5	Moderately Unimportant	0.0%
6	Never	1.2%	6	Very Unimportant	2.5%
no response		2.5%	no response		1.2%
		100.0%			100.0%
		$\bar{X}=1.79$; s.d.=1.01			$\bar{X}=1.68$; s.d.=1.03

Principals' practice and teachers' preferences are highly congruent regarding objective, substantive criteria for evaluation. Over 85 percent of the principals reported that evaluation is usually or always based on such criteria, and more than 86 percent indicated that substantive criteria are moderately or very important in an ideal evaluation system, with over 54 percent rating it as very important. Only 2 principals rated it as very unimportant. Applying the *t* -test, the difference between present practice and ideal practice is not significant:

Question #14 and #32. Teacher evaluation includes a process for establishing specific goals and objectives for the teacher.

Response:		Present			Ideal
1	Always	43.2%	1	Very Important	63.0%
2	Usually	42.0%	2	Moderately Important	23.5%
3	Sometimes	6.2%	3	Mildly Important	8.6%
4	Seldom	2.5%	4	Mildly Unimportant	0.0%
5	Rarely	2.5%	5	Moderately Unimportant	0.0%
6	Never	1.2%	6	Very Unimportant	1.2%
no response		2.5%	no response		3.7%
		100.0%			100.0%
		$\bar{X}=1.79$; s.d.=1.01			$\bar{X}=1.68$; s.d.=1.03

Over 85 percent of the principals reported that evaluation in their schools usually or always includes a process for establishing

specific goals and objectives for the teacher. Over 87 percent indicated that such a process is moderately or very important in an ideal evaluation system, with 63 percent rating it as very important. Such responses translate into high congruence with teachers' preferences. The difference between ideal practice and present practice on this characteristic, when the t -test is applied, is significant ($t=3.73; p < .001$).

Question #15: Which of the following are purposes of teacher evaluation in your school?
Question #33: In your professional opinion, which of the following should be purposes of teacher evaluation?

Response:	Present Practice			Ideal Practice		
		$\bar{X}$			$\bar{X}$	
Professional Growth	60.5%	4.16	3.94	70.4%	3.37	3.68
Program Evaluation	30.9%	6.53	3.72	43.2%	5.54	3.99
Improvement of Instruct.	83.8%	1.49	1.94	100.0%	1.00	0.00
Reward of Performance	19.8%	7.42	3.21	42.0%	5.64	3.97
Tenure and Termination	76.5%	2.88	3.41	63.0%	3.96	3.89
Teacher Accountability	48.1%	5.15	4.02	59.3%	4.26	3.96
Modification of Assign.	22.2%	7.22	3.35	35.8%	6.14	3.86
Other	8.6%			3.7%		

Improvement of instruction is the most frequently indicated purpose for evaluation. Almost 84 percent of the principals in the study indicated that improvement of instruction is a purpose for evaluation in their schools, while 100 percent rated it as a purpose in an ideal evaluation. Applying the t -test, the difference between ideal and present practice for this characteristic is significant ($t=3.73; p < .0001$).

Professional growth, although related to improvement of instruction, is rated substantially lower. Only about 60 percent of the principals indicated that the teacher's professional growth is a purpose for evaluation in present practice, while over 70 percent reported that it should be a purpose. The difference in the responses between improvement of instruction and professional growth is curious, since one wonders how

the first can occur without the second. The difference between ideal practice and present practice for professional growth only approaches significance ($t=1.92; p < .06$).

Tenure and termination decisions is the second highest indicated purpose for evaluation in present practice. Over 76 percent of the principals listed it as a purpose for evaluation. Teachers have indicated, however, that such decisions should not be a reason for evaluation. On this item, then, low congruence exists between teachers' preferences and principals' actual practice. In ideal practice, however, only 63 percent of the principals listed tenure and termination as a purpose, resulting in better congruence, but still not high. The difference between ideal and present practice, applying the t-test, is significant ($t=2.78; p < .01$).

Almost 20 percent of the principals in the study presently use evaluation results to reward superior performance, while 42 percent indicated that rewarding performance should be a purpose of evaluation. Applying a t-test, the difference between ideal and present practice on this item is significant ($t=4.47; p < .0001$).

Over 48 percent of the principals indicated that teacher accountability is presently a purpose of evaluation. Almost 60 percent indicated that accountability should be a purpose of evaluation. The difference between the two only approaches significance ($t=1.91; p < .06$).

Just over 20 percent of the principals in the study listed modification of teacher assignment as a purpose of evaluation, while nearly 36 percent indicated that it should be a purpose. The difference between the two is significant ($t=2.48; p < .05$).

Seven of the principals in the study listed other purposes for teacher evaluation. These included establishing specific goals for the following year; developing annual inservice themes; promoting success (which is closely related to professional growth); positive reinforcement of teachers; meeting the contractual agreement; and fulfilling a district requirement.

Question #16: Which of the following form the basis of teacher evaluation in your school?
Question #34: In your professional opinion, which of the following should form the basis of teacher evaluation?

Response:	Present Practice			Ideal Practice		
		$\bar{X}$	s.d.		$\bar{X}$	s.d.
Narrative description	56.8%	4.46	3.99	71.6%	3.27	3.63
Rating scale of compet.	37.0%	6.04	3.89	37.0%	6.04	3.89
Specific teacher object.	43.2%	5.54	3.99	65.4%	3.77	3.83
Rating scale of charact.	42.0%	5.64	3.97	32.1%	6.43	3.76
The "clinical cycle"	32.1%	6.43	3.76	45.7%	5.35	4.01
Other	3.7%			2.5%		

In present practice, the most frequently used format for teacher evaluation among the principals surveyed is the narrative description of performance. Nearly 57 percent of the principals who responded indicated that they use such a narrative, with 52 percent employing the narrative in combination with other procedures. Almost 72 percent indicated that a narrative should be part of the process, with 66 percent reporting it should be used in combination with other procedures. The difference between ideal practice and present practice is significant ($t=2.24; p < .05$).

Thirty-seven percent of the principals in the study employ a rating scale of teacher competencies, 3 percent using it along, and 34 percent using it in combination with other procedures. The same number

indicated that such a scale should be part of an ideal evaluation system, but all indicated this scale should be used in combination with other procedures. Because the responses for ideal and present practice are identical, there is no measurable difference for this item.

Forty-two percent of the principals use a rating scale of professional and personal characteristics, 36 percent in combination with other procedures, and 6 percent using it alone. Only 32 percent believe that this kind of scale should be part of an evaluation system. All but one of those believe it should be used in combination with other procedures. The difference between ideal and present practice is not significant.

Forty-three percent of the principals in the study include specific teacher objectives as a basis for evaluating teachers, 10 percent using them along, 33 percent in combination. Over 65 percent believe that objectives should be a basis for evaluation, 6 percent indicating they should be used alone, 59 percent in combination. According to several studies, teachers prefer evaluation to be based on specific objectives. Congruence between that preference and principals' practice is not high, although ideal practice results in higher congruence. The difference between ideal and present practice is significant ($t=3.51; p < .001$).

Approximately 32 percent of the principals in the study use the clinical cycle of supervision (pre-conference, observation, post-conference) as a basis for evaluation, 5 percent using it alone, 27 percent in combination. Over 45 percent indicated that the cycle should be used as a basis for evaluation, 7 percent indicating it should be the sole

basis, 38 percent reporting it should be used in combination. The difference between present and actual practice is significant ($t=2.25; p < .05$).

Three principals in the study indicated other bases for evaluation: a written summary of objectives achieved and the degree of achievement; student performance; and teacher self-evaluation.

Question #17: Teachers in your school participate in which of the following, if any?
Question #35: In your professional opinion, which of the following (if any) should teachers participate in?

Response:	Present Practice			Ideal Practice		
		$\bar{X}$	s.d.		$\bar{X}$	s.d.
Develop eval. policies	33.3%	6.33	3.80	54.3%	4.65	4.01
Designing eval. form	54.3%	4.65	4.01	64.2%	3.86	3.86
Designing theprocess	38.3%	5.94	3.91	59.3%	4.26	3.96
Establishing the goals	38.3%	5.94	3.91	74.1%	3.07	3.53

Thirty-three percent of the principals in the study reported that teachers are involved in developing the policies which govern teacher evaluation. Just over 54 percent indicated that teachers should be involved in developing policies. The difference between present and ideal practice is significant ($t=3.94; p < .001$).

Fifty-four percent of the principals in the study reported that teachers help design the evaluation form, while 64 percent indicated that teachers should participate in designing the form. This difference is not significant.

Thirty-eight percent of the principals in the study indicated that teachers participate in designing the evaluation process, and in establishing goals and purposes of evaluation. Fifty-nine percent indicated that teachers should help design the evaluation process, a difference which is significant ($t=3.23; p < .01$). Seventy-four

percent believe that teachers should help establish the goals, a difference which is also significant ($t=5.21; p < .0001$).

The following pair of questions requested 11 different responses. For each response, the respondent was asked to circle a number on a 6-point scale between two opposite characteristics of evaluation.

Question #18: How would you characterize the evaluation process in your school?
Question #36: In your professional opinion, how would you ideally characterize the process of teacher evaluation?

Response:		Present	Ideal
Objective	1	16.0%	34.6%
	2	25.9%	23.5%
	3	21.0%	12.3%
	4	18.5%	13.6%
	5	9.9%	7.4%
Subjective	6	3.7%	4.9%
No response		4.9%	3.7%
		100.0%	100.0%
		$\bar{X}=2.91$; s.d.=1.40	$\bar{X}=2.56$; s.d.=1.49

Only about 19 percent of the principals in this study rated present evaluation systems as more formative than summative, by circling a "1" or "2". Conversely, 37 percent indicated that their evaluation systems are more summative than formative, by circling a "5" or "6". Congruence, then, is quite low in present practice. In ideal practice, nearly 35 percent of the principals circled "1" or "2", characterizing the process as formative, while only 17 percent rated it more summative, by circling "5" or "6". The difference between ideal and present practice is significant. ($t=4.39; p < .0001$).

		Present	Ideal
Diagnostic	1	7.4%	16.0%
	2	11.1%	32.1%
	3	16.0%	25.9%
	4	25.9%	9.9%
	5	25.9%	6.2%
Descriptive	6	8.6%	4.9%
No response		4.9%	4.9%
		100.0%	100.0%
		$\bar{X}$=3.83;s.d.=1.41	$\bar{X}$=2.74;s.d.=1.33

Of the principals who responded to the survey, only about 19 percent rated their evaluation system as more diagnostic than descriptive, by circling a "1" or "2". Conversely, almost 35 percent rated their system as more descriptive than diagnostic, by circling a "5" or "6". In present practice, then, congruence is rather low between principals' practice and teachers' preference. In ideal practice, 48 percent of the principals characterized evaluation as more diagnostic, by circling a "1" or "2", while only 11 percent indicated it should be descriptive, by circling a "5" or "6", resulting in much higher congruence. The difference between ideal practice and present practice is significant. (t=4.54;p <.0001).

		Present	Ideal
Cooperative effort	1	9.9%	39.5%
	2	12.3%	21.0%
	3	13.6%	12.3%
	4	21.0%	11.1%
	5	18.5%	4.9%
Administrative task	6	19.8%	3.7%
No response		4.9%	7.4%
		100.0%	100.0%
		$\bar{X}$=3.92;s.d.=1.60	$\bar{X}$=2.28;s.d.=1.40

Twenty-two percent of the principals in the survey characterized their evaluation systems as a cooperative effort, by circling a "1" or "2". Thirty-eight percent indicated it is more an administrative task, by

circling a "5" or "6". Almost 35 percent rated their evaluation system midway on the scale, with 14 percent leaning toward a cooperative effort, and 21 percent leaning toward an administrative task. Teachers prefer that evaluation be a cooperative effort between the teacher and the principal. Principals' present practice is therefore not very congruent with that preference. In ideal practice, however, almost 60 percent characterized evaluation as a cooperative effort, by circling a "1" or "2", and only about 9 percent rated it an administrative task, by circling a "5" or "6", resulting in much higher congruence. The difference between present practice and ideal practice is significant ($t=7.90; p<.0001$).

		Present	Ideal
Teacher as colleague	1	11.1%	37.0%
	2	22.2%	23.5%
	3	24.7%	9.9%
	4	11.1%	13.6%
	5	21.0%	4.9%
Teacher as subordinate	6	3.7%	6.2%
no response		6.2%	4.9%
		100.0%	100.0%
		$\bar{X}=3.20$; s.d.=1.44	$\bar{X}=2.44$; s.d.=1.57

In present practice, 33 percent of the principals circled a "1" or "2", indicating that they consider the teacher to be a colleague rather than a subordinate. Conversely, almost 25 percent circled a "5" or "6", indicating that they consider the teacher a subordinate more than a colleague. Teachers indicate that they prefer collegiality in the evaluation process. Thus, congruence between principals' present practice and the reported preference of teachers is quite low. In ideal practice, however, almost 61 percent of the principals characterized the teacher as more a colleague, by circling a "1" or "2", while only 11 percent circled a "5" or "6", indicating that they consider the teacher a subordinate, resulting in much higher congruence. The difference between ideal practice

and present is significant ($t=3.70; p<.0001$.)

		Present	Ideal
On going	1	19.8%	49.4%
	2	18.5%	18.5%
	3	16.0%	9.9%
	4	16.0%	4.9%
	5	13.6%	6.2%
Specific Time Frame	6	8.6%	6.2%
		7.4%	4.9%
		100.0%	100.0%
		$\bar{X}=3.12$; s.d.=1.65	$\bar{X}=2.16$; s.d.=1.60

In present practice, the responses from principals in the study are rather evenly distributed between on-going evaluation and a specific time-frame for evaluation. Thirty-eight percent rated the process as on-going, by circling a "1" or "2", while 22 percent characterized it as a specific time-frame, by circling a "5" or "6". Thirty-two percent rated it midway on the scale. Congruence on this item between teachers' preference and principals' practice is not high.

In ideal practice, 61 percent of the principals in the study characterized evaluation as more on-going than time-limited, by circling a "1" or "2", while only 12 percent circled a "5" or "6" indicating that they would prefer a specific time frame, resulting in higher congruence. The difference between present practice and ideal practice is significant ($t=7.29; p<.001$).

		Present	Ideal
Non-punitive	1	37.0%	53.1%
	2	37.0%	19.8%
	3	11.1%	13.6%
	4	4.9%	4.9%
	5	1.2%	0.0%
Punitive	6	1.2%	1.2%
no response		7.4%	7.4%
		100.0%	100.0%
		$\bar{X}=1.93$; s.d.=1.04	$\bar{X}=1.72$; s.d.=1.04

Seventy-four percent of the principals in the study reported that their evaluation system is non-punitive, by circling a "1" or "2". Only

2 percent circled a "5" or "6", indicating that evaluation is punitive. In ideal practice, 73 percent characterized evaluation as non-punitive by circling a "5" or "6". However, a shift of 16 percentage points occurred from "moderately" to "very important". The difference between ideal practice and present practice approaches significance ($t=1.87; p < .07$).

		Present	Ideal
Mutual Trust	1	21.0%	64.2%
	2	32.1%	19.8%
	3	21.0%	8.6%
	4	18.5%	2.5%
	5	2.5%	0.0%
Suspicion	6	1.2%	0.0%
no response		3.7%	4.9%
		100.0%	100.0%
		$\bar{X}=2.51$; s.d.=1.19	$\bar{X}=1.47$; s.d.=0.77

Of the principals responding to the survey, 53 percent reported there is more mutual trust than suspicion in their evaluation systems, by circling a "1" or "2". Only about 4 percent reported more suspicion than trust by circling a "5" or "6". In ideal practice, 84 percent of the principals circled a "1" or "2", indicating that they would prefer mutual trust. None would choose suspicion. The difference between present and ideal practice is significant. ($t=7.29; p < .001$).

		Present	Ideal
Teacher as a person	1	18.5%	27.2%
	2	23.5%	24.7%
	3	19.8%	16.0%
	4	17.3%	7.4%
	5	11.1%	8.6%
Teacher as an employee	6	4.9%	11.1%
no response		4.9%	4.9%
		100.0%	100.0%
		$\bar{X}=4.076$; s.d.=1.47	$\bar{X}=4.2$; s.d.=1.70

Forty-two percent of the principals in the study circled a "1" or "2" for present practice, indicating that they view the teacher more as a

person than an employee, while 16 percent characterized the teacher as more employee than person by circling a "5" or "6". In ideal practice, 52 percent of the principals rated the teacher as more a person by circling a "1" or "2", but 20 percent circled a "5" or "6", indicating that they view the teacher more as an employee. The difference between ideal and present practice is not significant.

		Present	Ideal
Non-threatening	1	21.0%	44.4%
	2	37.0%	25.9%
	3	22.2%	16.0%
	4	11.1%	6.2%
	5	1.2%	0.0%
Threatening	6	0.0%	1.2%
No response		7.4%	6.2%
		100.0%	100.0%
		$\bar{X}$=2.30;s.d.=1.00	$\bar{X}$=1.84;s.d.=0.95

Fifty-eight percent of the principals in the study characterized evaluation in their schools as non-threatening by circling a "1" or "2". Only 1 percent indicated that it is threatening by circling a "5" or "6". However, 33 percent indicated some degree of threat by responding in the mid-range.

In ideal practice, 70 percent of the principals characterized evaluation as non-threatening by circling a "1" or "2", and only 1 percent circled a "5" or "6", indicating that they would have it be threatening. Twenty-two percent still responded in the mid-range, indicating that some degree of threat would be present in an ideal evaluation system. The difference between ideal and present practice is significant (t=3.40;p <.001).

		Present	Ideal
Satisfactory	1	13.6%	58.0%
	2	28.4%	21.0%
	3	22.2%	7.4%
	4	17.3%	3.7%
	5	4.9%	2.5%
Unsatisfactory	6	7.4%	1.2%
No response		6.2%	6.2%
		100.0%	100.0%
		$\bar{X}$=2.93;s.d.=1.44	$\bar{X}$=1.60;s.d.=1.11

Forty-two percent of the principals in the study indicated that their present evaluation systems are satisfactory, by circling a "1" or "2". Twelve percent circled a "5" or "6", indicating that they do not feel their systems are satisfactory. In ideal practice, 79 percent characterized evaluation as satisfactory by circling a "1" or "2", while 4 percent circled a "5" or "6", indicating that they still find it unsatisfactory. The difference between present practice and ideal practice is significant (t=6.98;p <.0001).

Organization of the Questionnaire Responses

In the previous sections, 99 separate questionnaire responses were presented as they appeared on the survey instrument. The first 7 were personal and demographic questions, leaving 92 items related directly to the study. The next 2 questions were designed to elicit information of use to both the present practice and ideal practice scales and hereafter appear as two pairs of responses. The other 90 questions were presented as 45 pairs--one question which asked for a response related to present practice; an identical question later in the questionnaire which requested a response related to ideal practice.

In Chapter II, the 28 teacher preferences regarding evaluation were grouped into 6 clusters:

I Teacher participation in evaluation decisions.
II Activities prior to classroom observation.
III Activities following classroom observation.
IV Training and competence of the evaluator.
V Conditions of the evaluation.
VI Purposes and criteria of the evaluation.

To organize the data for further analysis, the 92 questionnaire responses have been grouped according to those clusters. First, each of the present practice questions was paired with its corresponding ideal practice question, exactly as they were paired in the previous section. The result was 47 pairs of responses. Not every pair of items on the questionnaire, however, corresponded to one of the 28 identified teacher preferences. Conversely, some pairs either duplicated or closely resembled others, making it necessary to choose the pair which most closely represented the cluster item. In this manner, 28 pairs of questionnaire items were chosen for further analysis. Each pair corresponds to one of the 28 teacher preferences from one of the 6 clusters.

Table 3.1 presents the data in this manner. The first column indicates the percentage of principals who indicated a characteristic to be always or usually included in the evaluation system in their schools. The rate of inclusion ranges from 3 percent to 98 percent for the characteristics surveyed. The second column indicates the percentage of principals who reported each characteristic to be very important or moderately important in an ideal evaluation system. The rate of inclusion ranges from 30 percent to 99 percent for the characteristics surveyed.

TABLE 3.1 PERCENTAGE OF PRINCIPALS REPORTING A TEACHER PREFERRED CHARACTERISTIC TO BE USUALLY OR ALWAYS A PART OF PRESENT PRACTICE IN THEIR SCHOOLS AND THE PERCENTAGE WHO INDICATE A CHARACTERISTIC TO BE IMPORTANT IN IDEAL PRACTICE (N=81)

	PRESENT PRACTICE	IDEAL PRACTICE
Cluster I: Teacher Participation in Evaluation Decisions		
1. Teachers participate in designing the evaluation form.	54%	64%
2. Teachers participate in designing the process of evaluation.	38%	59%
3. Teachers participate in developing evaluation policies	33%	54%
4. Teachers participate in establishing goals and purposes.	38%	74%
Cluster II: Activities Prior to Classroom Observation		
1. Teacher and Principal meet for pre-conference	41%	54%
2. Teacher and Principal agree on time and place of observation	61%	56%
3. Teacher and Principal agree on criteria	63%	70%
Cluster III: Activities Following Classroom Observation		
1. Taacher and Principal meet for post-conference	96%	99%
2. Principal provides supportive comments	98%	98%
3. Principal gives a copy of report to teacher.	98%	98%
4. Principal allows teacher to reply to report	96%	96%
Cluster IV: Training and Competence of the Evaluator		
1. Principal is evaluated by supervisor	88%	93%
2. Principal has received training in evaluation.	79%	93%
3. Principal is knowledgeable in academic areas to be evaluated	84%	70%
4. Principal includes another teacher on an evaluation "team"	3%	30%
Cluster V: Conditions of the Evaluation		
1. Evaluation is more formative than summative.	19%	35%
2. Evaluation is diagnostic.	19%	48%
3. Evaluation is non-threatening.	58%	70%
4. Evaluation is a cooperative effort	22%	61%
5. Principal feels there is mutual trust.	53%	84%
6. Principal views the teacher more as a person than an employee.	42%	52%
7. Evaluation is an on-going process	38%	68%
8. Evaluation includes 3 or more observations	48%	59%
9. Evaluation leads to inservice training	37%	82%
Cluster VI: Purposes and Criteria		
1. Evaluation results are not used for tenure and termination	24%	37%
2. Teacher objectives are a basis for evaluation	43%	65%
3. Principal does not use a rating scale of competencies	63%	63%
4. Evaluation is based on substantive criteria.	85%	86%

Congruence Between Present Practice and Teacher-Preferred Evaluation Characteristics

The first research question to be answered was: "Do principals, in their evaluation of teachers, include procedures that are congruent with practices that have been identified in the literature as being associated with increased teacher satisfaction with evaluation?" To answer the question, each of the 28 teacher-preferred characteristics was placed in a "present practice" Scale of Congruence, from highest to lowest, based on the percentage of principals reporting that they include that characteristic as part of the evaluation system in their schools. (Table 3.2) On this scale, which was divided into 4 sections, High Congruence, Medium-High Congruence, Medium-Low Congruence and Low Congruence, the 28 characteristics arranged themselves in a fairly even distribution from High to Low congruence.

7. 84 percent of the principals who responded to the survey usually or always are knowledgeable in the academic areas which they evaluate.

8. 79 percent of the principals who responded to the survey have received inservice training in evaluation.

Medium-High Congruence

In Table 3.2, 6 teacher-preferred characteristics of evaluation appear in the "Medium-High" category of congruence with principals' present practice:

1. 63 percent of the principals who responded to the survey usually or always reach agreement with the teacher on the criteria for the evaluation prior to a classroom observation.

2. 63 percent of the principals who responded to the survey usually do not use a rating scale of personal and professional competencies to evaluate teachers.

3. 61 percent of the principals who responded to the survey usually or always reach agreement with the teacher regarding the time and place prior to a classroom observation.

4. 58 percent of the principals who responded to the survey reported that evaluation is not threatening to their teachers.

5. 54 percent of the principals who responded to the survey reported that teachers participate in designing the evaluation form used in their evaluation systems.

6. 53 percent of the principals who responded to the survey reported that there is mutual trust with their teachers regarding evaluation.

Medium-Low Congruence

In Table 3.2, 9 teacher-preferred characteristics of evaluation appear in the "Medium-Low" category of congruence with principals' present practice.

1. 48 percent of the principals who responded to the survey reported that their evaluation system includes 3 or more observations.

2. 43 percent of the principals who responded to the survey reported that they use teacher objectives as a basis for evaluation of teachers.

3. 42 percent of the principals who responded to the survey view the teacher more as a person than as an employee.

4. 41 percent of the principals who responded to the survey reported that they usually or always meet with the teacher for a conference prior to the observation.

5. 38 percent of the principals who responded to the survey reported that teachers participate in designing the evaluation process in their schools.

6. 38 percent of the principals who responded to the survey reported that teachers participate in establishing the goals and purposes of evluation in their schools.

7. 38 percent of the principals who responded to the survey reported that their evaluation system is an on-going process.

8. 37 percent of the principals who responded to the survey reported that evaluation results usually or always lead to specific in-service training.

9. 33 percent of the principals who responded to the survey reported that teachers participate in developing the policies which govern taacher evaluation.

Low Congruence

In Table 3.2, 5 teacher-preferred characteristics of evaluation appear in the "Low" category of congruence with principals' present practice:

1. 24 percent of the principals who responded to the survey do not use evaluation results for tenure and termination decisions.

2. 22 percent of the principals who responded to the survey reported that evaluation is a cooperative effort between the principals and the teacher.

3. 19 percent of the principals who responded to the survey reported that their evaluation system is more formative than summative.

4. 19 percent of the principals who responded to the survey reported that their evaluation system is diagnostic.

5. 3 percent of the principals who responded to the survey include another teacher on an evaluation team.

Summary

In present practice, the 28 evaluation characteristics associated with increased teacher satisfaction with evaluation range from 3 percent inclusion to 98 percent inclusion by the principals in the survey. Fourteen of the 28 characteristics are included at a rate which results in high or medium-high congruence (over 50 percent rate of inclusion) on the scale which was established. The remaining 14 characteristics are included at a rate which results in low or medium-low congruence (under 50 percent rate of inclusion) on the scale. Only 5 characteristics failed to reach an inclusion rate of at least 25 percent.

Conclusion: Principals who responded to the survey tend to include procedures in their evaluation systems that are congruent with practices associated with increased teacher satisfaction with evaluation.

Analysis

In the absence of satistical methods for analyzing the response data as they appear in Table 3.2, a number of additional conclusions can nevertheless be drawn by comparing the responses which appear on the High side of the congruence scale (over 50 percent rate of inclusion) with those which appear on the Low side (under 50 percent rate of inclusion). The 28 characteristics of evaluation in the study divide evenly on a Scale of Congruence, and the 14 evaluation characteristics with more than

a 50 percent rate of inclusion have features that can be compared and contrasted with the features of the 14 evaluation characteristics attaining less than a 50 percent rate of inclusion.

Concentration of Clusters

One notable feature of both the High and Low groups is a concentration of responses from particular clusters. For example, 7 of the 14 evaluation characteristics in the Low group are from Cluster V, Conditions of the Evaluation. Moreover, 3 of the 4 characteristics in Cluster I, Teacher Participation in Evaluation Decisions, appear in the Low group.

On the other hand, the High group includes all 4 of the characteristics from Cluster III, Activities Following Classroom Observation, as well as 3 of the 4 characteristics from Cluster IV, Training and Competence of the Evaluator, and 2 of the 3 characteristics from Cluster II, Activities Prior to Classroom Observation. Only Cluster VI, Basis of the Evaluation, is evenly divided between the High and Low groups.

Conclusions: (1) Principals' practice and teachers' preferences tend to be congruent regarding the activities prior to and following an observation, and regarding the training and competence of the principal.

(2) Principals and teachers tend to be less congruent regarding teacher participation in evaluation decisions, and regarding the overall conditions of the evaluation.

Procedural Versus Qualitative Characteristics

Another source of comparison is provided by noting which of the 28 characteristics of evaluation in the study are primarily procedural

and which are primarily qualitative (e.g., "preconference", which is procedural, versus "non-threatening" which is qualitative). Such a division yields 6 qualitative characteristics. Of the 6, 2 are in the High group and 4 are in the Low group:

High Congruence Qualitative Characteristics	
Non-threatening	(58%)
Mutual trust	(53%)

Low Congruence Qualitative Characteristics	
Formative	(19%)
Diagnostic	(19%)
Cooperative effort	(22%)
Views teacher as a person	(42%)

The 2 qualitative characteristics in the High group, "non-threatening" and "mutual trust", are both at the lower end of the High scale (58 percent and 53 percent rate of inclusion, respectively). Three of the qualitative characteristics in the Low group, "formative", "diagnostic", and "cooperative", are at the lower end of the Low scale (19 percent, 19 percent and 22 percent rate of inclusion respectively). One qualitative characteristic, "teacher as a person", is near the upper end of the Low scale (42 percent rate of inclusion).

Conclusions: (1) Principals in the survey report that they use evaluation procedures that are characterized by a relatively high level of mutual trust.

(2) A majority of the evaluation systems used by principals in the survey are reportedly non-threatening to the teachers.

(3) Evaluation procedures reported by principals in the survey are generally not very formative, not very diagnostic, and not characterized

by a spirit of cooperation between the principal and teacher.

Collegial Characteristics

A third source of comparison between the High group and the Low group are those evaluation characteristics which indicate a collegial relationship rather than a principal-dominated relationship. There are 12 such characteristics, of which 5 are in the High group and 7 are in the Low group:

High Congruence Collegial Characteristics

Postconference	(96%)
Agree on criteria	(63%)
Agree on time and place	(61%)
Participate in designing the form	(54%)
Mutual trust	(53%)

Low Congruence Collegial Characteristics

View teacher as a person	(42%)
Preconference	(41%)
Participate in designing process	(38%)
Participate in establishing goals	(38%)
Participate in developing policies	(33%)
Cooperative effort	(22%)
Formative	(19%)

Although more of the characteristics indicating collegiality appear in the Low category than in the High, the results are, nevertheless somewhat inconclusive. Three of the 12 characteristics, however, are essential to a collegial relationship. If they do not exist, the relationship can not be said to be collegial. Those 3 characteristics are "mutual trust", "cooperative effort", and "formative", and 2 of those characteristics appear in the lowest category on the congruency scale:

"cooperative effort" (22 percent inclusion) and "formative" (19 percent inclusion). The remainder of the 12 characteristics, while being indicators of collegiality, can, nevertheless, exist without a true collegial relationship. The implication is, then, that most principals in the study do not maintain a collegial relationship when they evaluate teachers. Question number 18E on the questionnaire, which was not included in the group of 28 responses to be analyzed, refers directly to the collegial relationship. On that question, only 33 percent of the principals who responded indicated that evaluation is characterized by a collegial relationship with the teacher.

Conclusion: Evaluation systems employed by principals in the survey are more principal-dominated than collegial in nature.

Discretionary Practices

While many of the 28 characteristics of evaluation included in the study are easily within the discretionary control of the principal, it is quite probable that others are not. Several of the characteristics, in fact, are probably specified in the professional agreement with the teachers' association (e.g., meeting for a post-observation conference, and providing a copy of the evaluation report to the teacher). Others are probably established by the superintendent and the board of education (e.g., teacher participation in developing evaluation policies). To obtain a clearer picture of the congruence which exists between the principals in the study and teachers' preferences, these characteristics of evaluation which are usually left to the discretion of the principal can be separated from those which might often be beyond the discretionary authority of the principal.

There are 15 characteristics of evaluation which are almost entirely discretionary to the principal, 5 in the High category, and 10 in the Low category:

High Congruence Discretionary Characteristics

Provides supportive comments	(98%)
Principal received evaluation training	(79%)
Agree on time and place	(61%)
Non-threatening	(58%)
Mutual trust	(53%)

Low Congruence Discretionary Characteristics

Three or more observations	(48%)
Teacher objectives as basis	(43%)
Views teacher as person	(42%)
Pre-observation conference	(41%)
On-going process	(38%)
Leads to in-service training	(37%)
Cooperative effort	(22%)
Formative	(19%)
Diagnostic	(19%)
Evaluation by "team"	(3%)

The results of this division are inconclusive. Although more of the discretionary characteristics appear in the Low category than in the High, 4 of those in the Low category have higher than a 40 percent rate of inclusion. Conversely, only 4 discretionary characteristics have lower than a 37 percent rate of inclusion. Three of those 4, however, "cooperative", "formative", and "diagnostic", are essential to the kind of collegial evaluation process which would be most satisfactory to teachers.

Of the 13 other evaluation characteristics, all of which are often non-discretionary, only 4 are in the Low category of the scale, while 9 are in the High category. Six are at the upper end of the scale, with a rate of inclusion which exceeds 75 percent. The implication is that principals tend to be more congruent with teachers' preferences when activities are

required of them, and less congruent when the decisions are left to their discretion.

Conclusion: If principals desire increased satisfaction from teachers regarding the evaluation process, the activities necessary to accomplish that are within their discretionary power to implement.

The NSPRA "Must" List

After reviewing statements from teachers' organizations across the country, the National School Public Relations Association (NSPRA) compiled a "must list" of 13 evaluation characteristics which were mentioned again and again in those statements.[1] All 13 were included in this survey. To determine how well the principals in the survey implement the list, those 13 characteristics have been separated out from the original 28.

High Congruence "Must List" Characteristics

Gives copy of report to teacher	(98%)
Post-observation conference	(96%)
Allows teacher to reply to report	(96%)
Principal is evaluated	(88%)
Principal received evaluation training	(79%)
Agree on criteria	(63%)
Agree on time and place	(61%)
Non-threatening	(58%)

Low Congruence "Must List" Characteristics

On-going process	(38%)
Leads to in-service training	(37%)
Not used for tenure and termination	(24%)
Cooperative effort	(22%)
Evaluation by "team"	(3%)

Overall, principals in the survey are doing rather well at implementing most of the characteristics seen as very important by the

[1]National School Public Relations Association, p. 57.

NSPRA. Eight of the 13 are in the High category, with 5 exceeding a 75 percent rate of inclusion. Only 3 are included by less than 25 percent of the principals. Those 3, however, are important. The issue of tenure and termination decisions, for instance, was first on the NSPRA list, with emphasis added to the condition that evaluation must have as its primary purpose the improvement of instruction, not the "legalistic purposes" of firing, determination of tenure, salary and promotion.

Conclusion: Principals, either because of discretionary decision-making or the successful negotiations of teachers' organizations, are implementing most of the highest-priority characteristics of evaluation as communicated by teachers' organizations, but there is room for improvement.

Difference Between Ideal and Present Practice

The second research question to be answered was, " Is there a difference between what a principal would ideally do and what he/she actually does in the process of evaluating teachers?

Principals in the study indicated that ideal evaluation systems would be quite different than present systems. In present practice the 28 characteristics of evaluation are split evenly above and below 50 percent on a Scale of Congruence based on the rate of inclusion of the characteristics. In ideal practice, only 4 characteristics obtained less than a 50 percent rate of importance, Of those 4, one obtained a 49 percent importance rate, and the other 3 were at a 30 percent or higher rate of importance (Table 3.3)

4. 96 percent of the principals who responded to the survey believe that it is moderately or very important to allow the teacher to reply to the report.

5. 98 percent of the principals who responded to the survey believe that the principal should be evaluated by his/her supervisor.

6. 93 percent of the principals who responded to the survey believe that the principal should receive in-service training.

7. 86 percent of the principals who responded to the survey believe that evaluation should be based on substantive criteria.

8. 84 percent of the principals who responded to the survey believe that there should be mutual trust.

9. 82 percent of the principals who responded to the survey believe that evaluation should lead to in-service training.

Medium-High Congruence

In Table 3.3, 15 characteristics of evaluation appear in the "Medium-High" category of congruence with principals' opinions of ideal practice:

1. 74 percent of the principals who responded to the survey believe that teachers should participate in establishing the goals and purposes of education.

2. 70 percent of the principals who responded to the survey believe that it is moderately or very important for the teacher and principal to agree on the criteria of the evaluation.

3. 70 percent of the principals who responded to the survey believe that it is moderately or very important for the principal to be knowledgeable in areas to be evaluated.

4. 70 percent of the principals who responded to the survey believe that evaluation should be non-threatening.

5. 68 percent of the principals who responded to the survey believe that evaluation should be an on-going process.

6. 68 percent of the principals who responded to the survey believe that teacher objectives should be a basis for evaluation.

7. 64 percent of the principals who responded to the survey believe that teachers should participate in designing the evaluation form.

8. 63 percent of the principals who responded to the survey believe that the principal should not use a rating scale of competencies.

9. 61 percent of the principals who responded to the survey believe that evaluation should be a cooperative effort.

10. 59 percent of the principals who responded to the survey believe that teachers should participate in designing the process of evaluation.

11. 59 percent of the principals who responded to the survey believe that evaluation should include 3 or more observations.

12. 56 percent of the principals who responded to the survey believe that it is moderately or very important for the teacher and principal to agree on the time and place for the observation.

13. 54 percent of the principals who responded to the survey believe that teachers should participate in developing evaluation policies.

14. 54 percent of the principals who responded to the survey believe that it is moderately or very important for the teacher and principal to meet for a preconference.

15. 52 percent of the principals who responded to the survey believe that the principal should view the teacher more as a person than an employee.

Medium-Low Congruence

In Table 3.3, 4 characteristics of evaluation appear in the "Medium-Low" category of congruence with principals' opinions of ideal practice:

1. 49 percent of the principals who responded to the survey believe that evaluation should be diagnostic.

2. 37 percent of the principals who responded to the survey believe that evaluation results should not be used for tenure and termination decisions.

3. 35 percent of the principals who responded to the survey believe that evaluation should be more formative than summative.

4. 30 percent of the principals who responded to the survey believe that it is moderately or very important for the principal to include another teacher on an evaluation "team".

Low Congruence

For ideal practice, there were no characteristics which obtained lower than a 25 percent rate of importance; therefore, none of the characteristics is in the Low category.

Comparison: Scale of Congruence (Present Practice) with Scale of Congruence (Ideal Practice)

Comparing the congruency information from present and ideal practice (Tables 3.2 and 3.3), it is evident that principals in this study hold opinions about ideal practices of teacher evaluation which, when compared with teacher-preferred characteristics, are noticeably more congruent than are present evaluation practices.

As noted, 20 characteristics of evaluation were rated moderately or very important in an ideal evaluation system by at least 50 percent of the principals responding to the survey, while only 14 characteristics are usually or always included in present evaluation systems by over 50 percent of those same principals. Conversely, only 4 characteristics are rated moderately or very important in an ideal evaluation system by less than 50 percent of the principals in the survey, while 14 characteristics attained less than a 50 percent rate of inclusion by those same principals.

The 4 characteristics ranked lowest on the ideal Scale of Congruence were among the 5 lowest ranked on the present practices Scale of Congruence, indicating that, although the rate of importance to ideal practice is higher than the rate of inclusion in present practice, priorities among the characteristics remain essentially the same. In fact, 7 of the 8 characteristics in the High category in present practice are also in the High category in ideal practice, and 5 of the 6 characteristics in the Medium-High category in present practice are also in the Medium-High category in ideal practice.

Conclusion: Although principals' opinions regarding ideal evaluation systems are more congruent with teachers' preferences than are the characteristics of actual practice, the rank of each characteristics by its rate of importance is very similar to its rank by rate of inclusion.

Comparison: Present Practice to Ideal Practice

In Chapter II, the 28 characteristics of evaluation associated with increased teacher satisfaction in the evaluation process were grouped into 6 clusters. They were again listed by cluster at the beginning of this chapter. In subsequent sections of this chapter, however, individual characteristics in the study were placed on two different Scales of Congruence without regard to cluster membership, one scale which ranked the characteristics according to the rate of inclusion in present practice; the other which ranked the characteristics by rate of importance to ideal practice. Comparing the information from both congruency scales indicated a difference between present practice and ideal practice, in the direction of higher congruence for ideal practice. These scales

revealed little information, however, regarding the six clusters which had been organized originally from the 28 characteristics.

Of importance to answering the second research question is a determination as to whether the difference between present and ideal practice, which has been noted, is statistically significant. Because 28 items are too many run separately on a *t*-test and obtain a valid statistical test of significance for the entire group, a more appropriate test is one which will consider the multiple variations of items within a cluster, and assess the statistical significance of the cluster as a whole. The test chosen was the Multivariate Analysis of Variance (MANOVA).

There are 4 variables in Cluster I, Teacher Participation in Evaluation Systems; (1) Teachers participate in designing the form, (2) Teachers participate in designing the process, (3) Teachers participate in developing policies, and (4) Teachers participate in establishing the goals. A MANOVA was run on the 4 items in Cluster I, comparing present practice and ideal practice. The multivariate *F* ratio comparing present practice with ideal practice was significant. ($F(4,77) = 7.54$; $p < .0001$.) Univariate *F* tests (1,80) indicated that teacher participation (1) in designing the process of evaluation, (2) in developing evaluation policies, and (3) in establishing the goals and purposes of evaluation, all contributed significantly to the multivariate *F* ratio at the $p < 01$ level. Teacher participation in designing the evaluation form did not contribute significantly.

The MANOVA indicates that for Cluster I principals believe teachers should ideally participate in evaluation decisions significantly more than they presently do, particularly in developing policies, establishing goals, and designing the process, each of which is a significant

contributor to the overall finding.

Cluster II, Activities Prior to Classroom Observation, contains three variables: (1) Teacher and principal meet for preconference ("preconference"), (2) Teacher and principal agree on time and place of observation ("time and place") and (3) Teacher and principal agree on criteria ("criteria"). A MANOVA was run on the 3 variables in Cluster II, comparing present practice and ideal practice, and a significant multivariate F ratio was obtained. ($F(3.76) = 3.54; p < .02$). Univariate F tests (1,78) indicated that when considered alone, none of the variables was significant at the $p < .05$ level. However, one of the variables, "time and place", approached significance ($p < .06$).

The significant multivariate F with the nonsignificant univariate F for each of the 3 variables indicates that principals' opinions regarding the importance of the 3 activities prior to classroom observation, in combination, differs significantly from the pattern and rate of inclusion of the 3 preobservation activities in present practice, with the greatest difference occurring in the variable "time and place". The unexpected finding of 3 variables which do not reach significance individually but do as a group is probably explained by the fact that 2 of the variables ("preconference" and "criteria") show positive difference between present and ideal practice, but the third ("time and place") shows a negative difference. The interaction among the 3 nonsignificant variables as they move in different directions produce a significant difference for the cluster, which demonstrates the power of a MANOVA to detect a significant multivariate difference among 3 non-significant variables.

Cluster III, Activities Following Classroom Observation, contains 4 variables: (1) teacher and principal meet for postconference, (2) principal provides supportive comments, (3) principal gives teacher a copy of report, and (4) principal allows teacher to reply to the report. A multivariate *F* ratio produced by running the MANOVA was not significant, probably because of the ceiling effect which occurred among all four variables.

Cluster IV, Training and Competence of the Evaluator, includes 4 variables: (1) principal is evaluated by supervisor ("principal evaluated"), (2) principal has received in-service training in evaluation ("principal in-serviced"), (3) principal is knowledgeable in academic areas to be evaluated ("principal knowledgeable"), and (4) principal includes another teacher on an evaluation "team" ("team"). A MANOVA was run on the 4 items comprising Cluster IV. A multivariate *F* ratio comparing present and ideal practice was significant. ($F(4,69) = 18.79$; $p < .0001$.)

Univariate *F* tests (1,72) indicated that "principal in-serviced" contributed significantly to the multivariate *F* ratio ($p < .0001$), as did "team" ($p < .0001$) and "principal evaluated" ($p < .05$). "Principal knowledgeable" did not contribute significantly.

The significant multivariate *F* for Cluster IV indicates that in matters related to the competence of the principal to provide quality evaluation, principals' opinions regarding the importance of the four variables to an ideal evaluation system is significantly different from the extent to which those variables reportedly are used in present evaluation systems. The greatest difference occurs, however, regarding

whether another teacher should be included in evaluating a teacher's performance. There was no significant difference regarding whether the principal needs to be expert in all academic areas which are to be evaluated.

Cluster V, Conditions of the Evaluation, contains 9 variables. (1) Evaluation more formative than summative, (2) Evaluation is diagnostic, (3) Evaluation is non-threatening, (4) Evaluation is a cooperative effort, (5) Principal feels there is mutual trust, (6) Principal views teacher as a person, (7) Evaluation is on-going process, (8) Evaluation include 3 or more observations, (9) Evaluation leads to in-service training. A MANOVA was run on the 9 items comprising Cluster V, producing a significant multivariate F ratio, when comparing present and ideal practice. ($F(9,51) = 12.87$; $p < .001$.). Univariate F tests (1,59) indicated that 8 of the variables contributed significantly to the multivariate F ratio ($p < .001$). One of the variables, "Principal views teacher as person", did not contribute significantly.

The significant multivariate F for Cluster V indicates that principals' opinions regarding the importance of at least 8 specific conditions of evaluation differs significantly from the number of principals who reported establishing such conditions in present practice. Specifically, principals' opinions of an ideal evaluation system are that it should be significantly more formative, more diagnostic, less threatening, more cooperative, characterized by more trust, less time-limited, include more observations, and lead to more in-service training than present systems.

Cluster VI, Basis of the Evaluation, contains 4 variables: (1) Evaluation results not used for tenure and termination ("termination"), (2) Teacher objectives are a basis for evaluation ("objectives"), (3) Principal does not use a rating scale of competencies ("rating scale"), and (4) Evaluation is based on substantive criteria ("criteria"). A MANOVA was run on the 4 items in Cluster VI, comparing present and ideal practice. The multivariate F ratio was significant. ($F(4,74) = 4.71$; $p < .002$.)

The univariate F tests (1,77) indicated that two of the variables contributed signficantly to the multivariate F ratio: "termination" ($p < .01$) and "objectives" ($p < .001$). The remaining two variables, "rating scale" and "criteria" did not contribute significantly.

The significant multivariate F ratio for Cluster VI indicates that principals' opinions regarding the importance of the 4 bases of evaluation, considering the variance of those opinions within the cluster, is significantly different from the bases of present evaluation systems as reported by principals in the survey.

Discussion

There is a significant difference between the characteristics of evaluation systems which principals in the study would ideally implement, and the characteristics of the evaluation systems which they presently report operating. In 5 of the 6 clusters of characteristics which were studied, the difference was significant. No difference occurred in Cluster III because a ceiling effect restricted the variance. Other than for the 4 activities following a classroom observation (Cluster III),

all of which exceeded a 96 percent rate of inclusion in present practice, principals in the study indicated that their conceptions of ideal evaluation systems were significantly different than the systems they presently operate.

First, according to the survey responses, principals believe that teachers should participate much more than at present in the decisions leading to implementation of an evaluation system. Second, a significant number of the principals surveyed indicated that the activities prior to an observation should be different than in present evaluation systems. The preconference is seen as more important than in present practice, as is agreeing on the criteria for evaluation. Reaching agreement with a teacher on the time and place of an observation, however, is less important than in present practice.

Third, opinions of the respondents regarding the competence of the evaluator, are also significantly different than present conditions. Many of the principals who are competent in most academic areas to be evaluated indicated that such competence is not important. A significant number of the respondents (approximately one-third) indicated that including a teacher on an evaluation team is important, while only 2 principals presently do so. These responses invite the conclusion that, according to the principals' in the study, it is quite important for the principal to be a competent evaluator. Conversely, it is less important for the principal to be competent in the subject areas being taught. However, to have a competent peer also participate in a performance evaluation results in a better evaluation

Fourth, principals' opinions about the ideal conditions of an evaluation differ significantly from present evaluation conditions, according to the survey. All 9 conditions of evaluation which were surveyed relate directly to a humanistic, collegial evaluation process, which focuses on the professional growth of the teacher. To summarize the responses of the principals in the study, an ideal evaluation system would be more formative, more diagnostic, less threatening, more cooperative, characterized by more trust, more of an open-ended, on-going process, based on more observations, and lead to more in-service training than present evaluation systems do. The conclusion is that principals' opinions of ideal evaluation systems are more humanistic and collegial than the systems they presently operate.

Fifth, the opinions of the principals in the study were significantly different than the evaluation procedures they presently employ in regard to the basis of evaluation. According to the survey responses, tenure and termination decisions should be less a result of evaluation than they now are, and teacher objectives should be utilized much more than at present. A rating scale of characteristics should be employed less, but evaluation systems should continue to be based on objective, substantive criteria.

Conclusion: In the opinion of the principals in the survey, evaluation should be more collegial, and should include teachers in planning and implementation, much more than at present.

Individual Characteristics Ranked by Magnitude of Difference

A significant difference has been noted between present practice and ideal practice in 5 of the 6 clusters of evaluation characteristics. Moreover, on a scale of congruence, the 28 characteristics of evaluation, when compared with teacher preferences, produced an overall higher level of congruence for principals' opinions of ideal evaluation practice then for principals' present evaluation practice. Some of the 28 individual characteristics which were studied, however, reflect a large difference between present and ideal practice, while others show little or no difference. In this section, the difference between ideal and present practice have been ranked (Table 3.4).

The greatest difference (45 percentage points) is in the area of teacher in-service as a result of evaluation. In addition, of the top 6 characteristics in the ranking, 5 are from Cluster V, Conditions of the Evaluation, indicating a significant difference between the principals believe about evaluation and how they presently implement evaluation systems.

Conclusion: The greatest difference overall between present practice and principals' opinions of ideal practice occurs among the evaluation characteristics which reflect a humanistic and collegial system.

TABLE 3.4 EVALUATION CHARACTERISTICS RANKED BY MAGNITUDE OF DIFFERENCE BETWEEN PRESENT PRACTICE AND IDEAL PRACTICE

Cluster Number	Rank	Characteristic	Present Practice	Ideal Practice	Difference
V-9	1	Leads to in-service training	37%	82%	45
V-4	2	Cooperative effort	22%	61%	39
I-4	3	Participate in establishing goals	38%	74%	36
V-5	4	Mutual trust	53%	84%	31
V-7	5	On-going process	38%	68%	30
V-2	6	Diagnostic	19%	48%	29
IV-4	7	Evaluation by "team"	3%	30%	27
VI-2	7	Teacher objectives as basis	43%	65%	22
I-2	9	Participate in designing process	38%	59%	21
I-3	10	Participate in developing policies	33%	54%	21
V-1	11	Formative	19%	35%	16
IV-2	12	Principal received evaluation training	79%	93%	14
II-2	13	Pre-observation conference	41%	54%	13
VI-1	14	Not Used for tenure and termination	24%	37%	13
V-3	15	Non-threatening	58%	70%	12
V-8	16	Three or more observations	48%	59%	11
I-1	17	Participate in designing form	54%	64%	10
V-6	18	Views teacher as person	42%	52%	10
VI-3	19	Does not use a rating scale	58%	68%	10
II-3	20	Agree on criteria	63%	70%	7
IV-1	21	Principal is evaluated	88%	93%	5
III-1	22	Post-observation conference	96%	99%	3
VI-4	23	Based on substantive criteria	85%	86%	1
III-2	24	Provides supportive comments	98¢	98%	0
III-3	25	Gives a copy of report to teacher	98%	98%	0
III-4	26	Allows teacher to reply to report	96%	96%	0
II-2	27	Agree on time and place	61%	56%	-5
IV-3	28	Principal knowledgeable in every area	84%	70%	-14

Analysis of Intervening Variables

The third research question to be answered by this study was, "What effect, if any, does each of the following have on research questions 1 and 2?"

A. Number of years as a teacher.
B. Number of years as a principal.
C. Number of years in education.
D. Highest degree earned.
E. Staff size.
F. Average income of families in the community.

A measure of congruence between ideal and present practice was obtained on 18 of the 28 pairs of variables: (1) preconference, (2) time and place, (3) criteria, (4) postconference, (5) supportive comments, (6) copy of report, (7) attach reply, (8) principal knowledgeable, (9) team, (10) formative, (11) diagnostic, (12) non-threatening, (13) cooperative, (14) mutual trust, (15) teacher as person, (16) on-going, (17) in-service, and (18) criteria. This measure of congruence has no relationship to the congruence between principals and teachers which is the central theme of the study. It is, rather, a measure of the consistency with which each principal responded to each of the two items in a pair (present vs. ideal). Only 18 of the 28 pairs were used in this particular analysis because the computation which created the congruency measure (present practice minus ideal practice, plus 10 to clear the minus sign) required the 6-point scale. The 10 variables not included in this analysis each had a 2 or 3 point scale.

The congruency measures on the 18 pairs of variables were analyzed by multiple discriminant analysis using the stepwise RAO V selection method. Discriminant analysis is a multivariate statistical procedure

that attempts to determine whether previously determined groups can be discriminated from one another based on certain specific variables. In this case, the survey responses were sorted into groups 6 different times, once for each of the 6 intervening variables. Discriminant analysis then attempts to classify each case into a group based on the statistical "description" developed by the analysis phase of the program. This classification is then compared to the known group membership of the sample, which provides an informal measure of the discriminent coefficients' validity.

From the survey responses, the principals in the study were first grouped by the number of years each had been a teacher. The range was 1 year through 22 years. From this range, 3 groups were arbitrarily formed: Group 1 (1 through 5 years), Group 2 (6 through 10 years), and Group 3 (over 10 years).

The next grouping was according to the number of years each had been a principal. The range was 1 year through 30 years. Three groups were formed: Group 1 (1 through 5 years), Group 2 (6 through 10 years), and Group 3 (over 10 years).

The third grouping was by the total years each had spent in education. The range was 9 years through 36 years. Three groups were formed: Group 1 (15 years and under), Group 2 (16 through 25 years), and Group 3 (over 25 years).

The fourth grouping was by the highest degree each had earned. There were 4 groups: M.A., M.A. plus 30 hours, M.A. plus 45 hours, and Doctorate.

The fifth grouping was done by the size of the staff each supervised. Three groups were formed: Group 1 (less than 16), Group 2 (16 through 25) and Group 3 (over 25).

The final grouping was by the average income of families in the school district, as reported by the principal. Four classifications were used: Group 1 (less than $15,000), Group 2 ($15,000 - $30,000), Group 3 ($30,000 - $45,000), and Group 4 (over $45,000).

The discriminant analysis was run on each of the above 6 groupings. Five were non-significant. The only grouping to reach significance was the number of years as a principal ($p < .05$). In that analysis, there were significant differences among the vectors of the means of the 3 groups (Wilks' $\lambda = 0.34$; $p < .04$).

In a discriminant analysis, the functions are derived in such a way that the discriminant scores are in standard (Z) score form and the absolute value of the coefficient indicates the relative contribution of each variable to the function(s). The maximum number of functions obtained is equal to the number of groups minus 1. When the group centroids for these data were plotted, clear separation occurred among the three groups on the discriminant function. The centroid locations for the group who had been principal for 5 years or less was 0.96432. For those who had been principal for 6 through 10 years, the centroid location was -1.48702, and for those who had been principal for over 10 years, the centroid location was 0.35179.

Six of the variables contributed heavily to the discriminant function:

Variable	Coefficient
Mutual trust	-1.43120
Evaluation leads to inservice	0.86591
Postconference	0.38708
Supportive comments	-0.34517
Teacher replies to evaluation	0.29651
Evaluation is non-threatening	0.29321

A review of the group means for each of these 6 variables reveals a curvilinear relationship, which is not useful to the study. Because the groups were originally organized in a linear progression according to number of years as a principal, a relationship which would have meaning should also form a linear progression.

Conclusion: The number of years a principal spent as a teacher, as a principal, or in education do not have any effect on the congruency between principals' evaluation practices and teachers' preferences. Neither do the education of the principal, the size of the school, or the average income of families in the district.

Analysis of District Documents

The fourth research question to be answered by this study was, "Are district-approved teacher evaluation policies and procedures congruent with practices that have been identified in the literature as being associated with increased teacher satisfaction with evaluation?"

In the cover letter which was sent with the survey instrument, principals were asked to send a copy of district-approved policies and procedures, as well as a copy of the evaluation form they used. On the front page of the survey instrument, principals were reminded to send those documents. In spite of that, only 28 principals sent the requested

evaluation documents with the questionnaire. Of these, 9 were duplicates, leaving 19 useful for the study. Only 8 of the 19 included district policies; the other 11 sent only evaluation forms. With so few documents to work with, no useful conclusions could be drawn regarding the congruency between board-approved policies and teacher preferences.

CHAPTER IV

ANALYSIS OF FINDINGS

Introduction

The first purpose of this study was to determine whether principals, when they evaluate teachers, incorporate methods which tend to result in greater teacher satisfaction with evaluation. A second purpose was to discover whether principals' concepts of ideal evaluation procedures are different than the procedures they presently report using, and whether those ideal procedures are closer to the way teachers prefer to be evaluated. Both purposes were accomplished. First, it was discovered that suburban elementary principals who responded to the survey, regardless of the number of years as a teacher, as a principal, or as an educator, and regardless of the size of the school they work in, the highest degree they have earned, or the economic conditions of the district they serve, reported that they tend to include procedures in their evaluation systems which are congruent with characteristics of evaluation that have been associated with increased teacher satisfaction with evaluation. Congruence is not consistent across all characteristics, and there are examples of very high and very low congruence. Nevertheless, the tendency toward overall agreement between principals' evaluation systems and teacher-preferred evaluation characteristics is clear.

Second, it was discovered that principals' concepts of the characteristics of ideal evaluation systems are significantly different than

the characteristics of the evaluation systems they presently report operating. The difference occurs primarily among the evaluation characteristics which would be associated with a humanistic, collegial system of evaluation. Moreover, the principals' concepts of ideal evaluation systems, as indicated in the survey, are closer to the way teachers prefer to be evaluated than are the evaluation systems reportedly in operation.

In addition to the two major findings, some other relationships emerged that, when considered along with the overall findings, help to elucidate the matter of congruence between present evaluation systems and teacher-preferred characteristics of evaluation. First, principals in the survey tend to be implementing most of the high-priority characteristics of evaluation which have been supported by teachers' organizations across the country.

Second, evaluation practices which produce the greatest congruence with teacher-preferred characteristics tend to be those that are non-discretionary, over which the principal has little control, e.g. providing a copy of the evaluation report. Conversely, the practices which produce the least congruence with teacher-preferred characteristics of evaluation tend to be those that are more discretionary, over which the principal has greater control, e.g., developing a cooperative, collegial evaluation program.

Third, principals in the survey tend to operate evaluation systems that are more principal-dominated than collegial in nature, although from the opinions indicated in the survey, principals tend to believe that teacher evaluation should be more collegial than it is.

Fourth, principals in the survey believe that teachers should be involved much more in the planning and implementation of evaluation systems than they are at present. That involvement should include developing policies as well as setting goals and designing programs.

Fifth, present evaluation systems operated by principals in the survey tend to be summative, and not very diagnostic, and therefore probably do not focus strongly on improvement of instruction.

These five secondary findings reveal that the congruence which seems to exist between principals' evaluation practices and teachers' preferences while substantial, is also limited. In the pages that follow, it will be demonstrated that the level of congruence that has been demonstrated by the study is limited to rather specific types of activities. Through an analysis and interpretation of the survey responses, it will be shown, for example, that principals' evaluation programs are far from being the collegial, cooperative ventures that teachers would like them to be. Prior to that analysis, however, two important observations will be made regarding the interpretability of the results. First, the manner in which teachers' preferences were determined will be discussed, in terms of both the strengths and limitations of that process. Next, there will be a discussion regarding the nature of congruency itself, with an indication of potential pitfalls in interpreting congruency.

Following that, a review of the findings will then confirm that congruency between teachers' preferences and principals' evaluation practices tends to be evident, although inconsistent, with congruency being highest among non-discretionary, non-collegial activities, such as providing a copy of the evaluation report and allowing the teacher to attach

a written reply. Further analysis will reveal that, in spite of substantial agreement regarding the items on the NSPRA "must list" congruence is quite low on two critical qualitative characteristics: the purpose for evaluation and the collegial nature of evaluation.

Following that analysis, the discussion will center on the nondiscretionary characteristics of the congruency between principals' evaluation and teachers' preferences. It will be shown that congruency is highest when principals have little discretionary decision-making power, such as holding a postconference, and lowest when principals have the greatest discretionary decision-making authority, such as creating a cooperative, collegial evaluation program.

Furthermore, the analysis will reveal that, because of a lack of clear operational definitions of terms, characteristics of evaluation which would suggest collegiality in the evaluation process, like the postconference and mutual trust, when considered along with other responses, may suggest a principal-dominated system. Other characteristics, such as teacher participation in evaluation decisions, confirm that evaluation programs, as indicated by the principals in the survey, tend to be administrator-dominated.

Finally, the analysis will reveal that principals' opinions regarding ideal evaluation practices are much more congruent with teachers' preferences than are presently reported evaluation practices, and concurrently are more collegial and cooperative.

Teacher-Preferred Characteristics

To ascertain how teachers prefer to be evaluated, an extensive literature review was conducted. From the publications that were reviewed, a list of 28 evaluation characteristics was compiled, each of which was supported by at least 2 sources as a characteristic of evaluation which has been associated with greater satisfaction on the part of teachers toward evaluation. In the course of searching through the literature, a common thought was repeatedly expressed. When teachers are satisfied with the way their performance is evaluated, they are more likely to benefit from the evaluation through improved instructional skills, and they tend to be more satisfied with other aspects of their work. Unfortunately, however, according to most writers teachers tend to be unhappy with the quality of evaluation they receive; so much so, in fact, that teachers' organizations have been attempting, often successfully, to negotiate key components of evaluation into professional contracts.

The 28 characteristics of evaluation that finally emerged from the review of the literature reflected the results of a number of research studies as well as statements made by local and national teachers' organizations regarding teacher evaluation. The list is a compilation of conclusions and opinions from the sources reviewed. Nowhere in the literature was an identical list of teacher-preferred characteristics discovered.

Because no measurement has been made of teachers' actual agreement with all 28 characteristics from the list, congruence with present evaluation systems is, by necessity, an inferred measure. It is understood, for example, that teachers' organizations which were wurveyed by the

NSPRA, although representing the majority of teachers, do not necessarily represent the view of all teachers. In that same vein, studies that report certain preferred characteristics of evaluation were often small and limited in scope, and generalization of the results is therefore subject to question. The opinion poll by the NEA, moreover, tended to uncover conflicting attitudes and opinions on the part of teachers. For these reasons, at least two sources supported each characteristic before it was included in the study. The 28 characteristics which were collected for the present study represent, therefore, as close an approximation as possible to a summary of the important characteristics teachers most prefer in evaluation. Significantly, the list is compatible with major writers on the topic of teacher evaluation, and is consistent with the currently popular concepts of collegiality, professional growth, and instructional excellence. In the final analysis, however, most of the characteristics on the list, apart from being supported by research studies and statements from teachers' organizations, just make good sense.

Congruency

Once the list of 28 preferred characteristics of evaluation had been established and organized into 6 clusters, this study set out to determine whether the evaluation practices presently reported as employed by principals are congruent with those 28 preferred characteristics. Once again, it is recognized, however, that congruency is inherently a difficult concept to measure. It is possible for two individuals, for example, to agree completely on a particular concept while each, though speaking

words exactly like the other, maintains a personal definition of the concept quite dissimilar to the other. It is possible for that same circumstance to have happened with the present questionnaire. The respondents may have held entirely different concepts than those which were intended. In that case, congruency might appear to have occurred, when in fact it did not. Conversely, two individuals might appear to disagree on a concept, because each is using words quite different from the other, when in fact the concepts they hold are very similar. Congruency may seem to be missing, when in fact it is present. Such may also be the case with the questionnaire which was used in the survey: some words such as "diagnostic" or "summative" may not have conveyed the same ideas to all respondents. With this limitation in mind, the survey instrument which was developed attempted to communicate as accurately as possible both the concepts and the language of the characteristics of evaluation shown in the literature review to be preferred by teachers.

For these reasons, however, developing the survey instrument was not an easy task. In a few cases, the concept which was originally expressed in the literature was not clearly written, and the language that was then chosen for the question on the survey instrument represented, by necessity, a personal interpretation of the original concept. An example is the term "subjective survey sheet". One author used the identical words "subjective survey sheet", while another, attempting to describe the same idea, used the words "descriptive method", and a third called it a "scaled rating". These three ideas, though different, seemed closely related. The item on the questionnaire which attempted to cover all three terms read "rating scale of professional and personal

characteristics", which tried to capture and combine "subjective"," descriptive", and "scaled rating". To the extent the effort was successful, the possibility that congruency was being measured was enhanced. The margin for error, however, is admittedly quite large.

Most items on the questionnaire, however, were able to use language nearly identical to that in the literature, reducing the opportunity for error in interpretation. On some of those items, though, it was still difficult to know whether congruency was accurately being measured, because of the possibility of other types of errors. For example, principals who reported that they shared mutual trust with their teachers were inferring that trust. Such reporting was highly subjective, yet was considered necessary to the scope of the study. What was important was not the accuracy of the statement, but rather the combined impressions of many principals as to the level of trust associated with evaluation in their schools.

An additional source of potential error was the clarity of the concepts themselves. For example, the term postconference can convey anything from a brief chat after an observation to a lengthy, formal conferance designed to provide supportive feedback for the purpose of improving specific instructional skills. Clearer operational definitions of terms, therefore, might have reduced the high response rate in cluster III, and provided greater variance of responses on each of the items.

Regardless of the clarity of the survey items, however, another potential source of error is principal bias. Because it is assumed that principals are familiar with the literature, they might be expected to report more teacher-oriented procedures than they actually use, thereby inflating the level of congruence for present practice. That same bias,

however, would likely inflate their ideal responses, resulting in a fairly stable difference between ideal and present practice.

Keeping in mind these difficulties with measuring congruency, the results of the survey did establish a rather high level of agreement between principals and teachers regarding the 28 evaluation characteristics which were studied. In the pages which follow, that congruency will be analyzed and interpreted in light of the inherent limitations already discussed, as well as the implications frmm the study results.

Conclusions

A number of conclusions were drawn from the presentation of the response data in Chapter III. All are related to the central conclusion: principals who were surveyed tend to report including procedures in their evaluation systems that are congruent with practices associated with increased teacher satisfaction with evaluation. The congruence was far from perfect; only 14 of the 28 characteristics which were studied exceeded a 50 percent inclusion rate (Table 3.2). Several of the characteristics, however, attained nearly 100 percent inclusion, and several were in the 35 to 45 percent range. Thus, there was a noticeable tendency toward agreement between principals' evaluation practices and the teacher-preferred characteristics of evaluation expressed in the literature.

Agreement, as was shown in Table 3.2, was not consistent. It tended to be strong in some areas, weak in others. For example, congruency was highest among some of the characteristics on the list which could be classified as procedural, such as holding a postconference and

giving a copy of the evaluation report to the teacher. Conversely, congruency tended to be lower among the characteristics which could be classified as qualitative, such as collegiality, cooperation, and a diagnostic, formative style of evaluation. The implications of these tendencies are very apparent when the results of the survey are compared, as they are in the next section, with a high priority list of evaluation characteristics expressed by teachers' organizations, and compiled by the NSPRA.

The NSPRA "Must List"

After surveying teachers' organizations and school districts across the country, the NSPRA in 1974 published a report which included, among other things, a "must list" of evaluation procedures and characteristics. In that list, the NSPRA summarized 10 "salient points... emphasized again and again by teachers' organizations" regarding evaluation procedures.[1]

Those 10 points were translated into 13 survey items for the present study in which it was discovered that 8 were being implemented by over 50 percent of the principals, and 5 were being implemented by fewer than 50 percent of the principals. The conclusion was that principals in the study tend to be implementing most of the high-priority characteristics of evaluation supported by teachers' organizations across the country. If the individual responses which led to that conclusion are closely analyzed, however, other important implications become evident.

---

[1]National School Public Relations Association, p. 57.

When the 5 specific items which generated low congruence are separated from the rest of the list, 2 items immediately stand out. Although the NSPRA list was not organized by priority, these 2 items were highlighted on the list, and could easily be considered the most important. The first of the two items was given prominence on the list by being placed first, and also by italics emphasizing its urgency:

> The purpose of teacher evaluation must be clearly understood to be improvement of instruction, not for formal, legalistic purposes of firing, determination of tenure, salary and promotion (emphasis in original).[2]

Principals in the survey, however, indicated rather low congruence with this statement. In spite of the fact that almost 94 percent of the principals agreed that improvement of instruction is a purpose of evaluation, 76 percent indicated that evaluation is used for tenure and termination decisions, which conflicts directly with the position of teachers' organizations. This finding is consistent with other parts of the NSPRA report, however. For that report, the NSPRA surveyed school districts as well as teachers' organizations, and found school districts reporting that evaluation results must be used to make personnel decisions such as tenure and termination. The NSPRA report concluded that there is direct conflict between teachers' organizations and district administrators on this point. Teachers want evaluation to be limited to the improvement of instructional skills, while administrators express the need for evaluation to have a dual purpose: personnel decisions and improvement of instruction.[3] The paradox is that the position

---

[2] Ibid.
[3] Ibid., p. 56.

of each group on this issue holds merit. Teachers' organizations are correct in asserting that evaluation which is to be used to make a decision as significant as termination has little change of engendering cooperation, trust and collegiality. School districts are correct in holding the position that some teachers are not worthy of continued employment or of being granted tenure, and formal evaluation procedures are the only way to document ineffective teaching skills.

The second characteristic of evaluation emphasized by the NSPRA as being of crucial importance is also a qualitative feature and even more closely related to the characteristics of collegiality than the first:

> Above all, evaluation must take place in a constructive and non-threatening atmosphere. The teacher must feel that improvement of his performance is a cooperative effort involving him, his evaluators, and others on the school staff. No matter how well designed--in the abstract--an evaluation program may seem, if it is perceived by teachers as negative or punitive, it will not improve teaching but will lower teacher effectiveness because of teacher fears and lowered morale (emphasis added).[1]

A number of studies have verified that teachers tend to feel threatened by evaluation--over 50 percent in Jensen's study,[2] over 70 percent in the study by Young and Heichberger.[3] Other studies have demonstrated that trust is enhanced by collegial styles of evaluation which regard the teacher as a person, and focus on the improvement of skills. Few principals in the present study, however, seem to implement collegial styles of evaluation, for it is just those areas where congruence is the lowest. Only 22 percent reported that evaluation is a

---

[1]Ibid.
[2]Jensen, p. 137.
[3]Young and Heichberger, p. 14.

cooperative effort. Only one-third indicated that evaluation is collegial. Even though more than half indicated that evaluation is not threatening, and nearly three-quarters indicated evaluation is not punitive, in light of the NSPRA report and other studies which have been cited, one wonders whether their reporting is accurate.

Consequently, on both of the qualitative evaluation characteristics to which the NSPRA imparted a sense of urgency, the principals in the survey group have indicated an inconsistent response, and a rather low level of congruence. Although teachers' organizations advocate a collegial approach to evaluation, less than a third of the principals in the study do so. Although most teachers are adamant that evaluation should not be used for tenure and termination decisions, over 70 percent of the principals in the study do so. On the two key qualitative characteristics most urgently expressed on the NSPRA list, the purpose of evaluation and the collegial nature of evaluation, there is wide separation between principals' present practice and teachers' preferences. Although overall congruence between the characteristics of evaluation advocated by teachers' organizations and the procedures of evaluation practiced by principals in the study group seems to be rather high, the appearance of collegiality is misleading.

Discretionary Characteristics

One of the conclusions from the data analysis in Chapter III was that evaluation characteristics which result in high congruence between principals' evaluation practices and teachers' preferences tend to be non-discretionary activities. That is, the principal has little control

over many of the characteristics which attained high congruence. For example, the negotiated agreement between a school board and the teachers' association often includes a clause on evaluation which establishes both the criteria for the evaluation and the evaluation form. These and other activities are often negotiated into the contract because teachers consider them to be high-priority characteristics of the evaluation process, and do not wish to leave them to the discretion of the principal. Other characteristics of the evaluation process may not be negotiated into the contract, but are often decided by the superintendent. For example, the superintendent probably decides whether teachers in the district will participate in establishing the policies which govern the evaluation process. Altogether, 13 of the 28 evaluation characteristics in the present study are non-discretionary--that is, often required by someone other than the principal. Nine of those 13 non-discretionary characteristics attained high congruence between teachers' preferences and principals' evaluation practices.

Conversely, the 15 remaining characteristics, over which principals tend to have discretionary control, such as making evaluation an on-going, diagnostic process, tended to have low inclusion rates by principals in the study. The conclusion is clear that activities which are left to the principals' discretion are implemented less frequently than those which are not. Furthermore, activities left to the principals' discretion tend to be those which are characteristics of a more collegial system, such as encouraging evaluation to be a cooperative effort, increasing the diagnostic function of evaluation, and creating more formative evaluation procedures, resulting in evaluation systems that are more administrator-dominated than collegial in the schools served by the respondents. The implication is, however, that if

principals desire collegial, cooperative, and formative evaluation systems, it is within their discretionary power to implement them. In so doing, their evaluation programs would become more highly congruent with teachers' expectations for evaluation. The implication for teachers, however, is even clearer. If teachers desire evaluation systems that are responsive to their expectations of collegiality, participation, cooperation, trust and professional growth, it will likely be necessary for teachers' organizations to negotiate the components of that kind of system into professional contracts.

Administrator-Dominated Evaluation Systems

Thus far, an analysis of the survey responses has led to the conclusion that principals' present evaluation practices and teachers' preferred evaluation characteristics tend to be congruent, although the congruence is not consistent across all characteristics. It has been demonstrated that the characteristics which tend to be most congruent are generally non-discretionary and primarily procedural, while those which tend to be incongruent are generally discretionary and primarily qualitative. This section discusses the third attribute which distinguishes congruent characteristics from incongruent characteristics in the present study, one which has been alluded to in previous sections: many of the characteristics which tend to be incongruent are associated with collegial, cooperative evaluation systems. As a result, present evaluation systems, rather than being cooperative activities, seem to be dominated by administrators.

Research studies have indicated that teachers are more satisfied and have more confidence in evaluation when it is characterized by a cooperative, collegial atmosphere. Yet several items in the present study indicate that the responding principals tend to report operating evaluation systems that are not collegial in nature, but rather tend to be dominated by the principal. Only about one-third of the principals, for example, indicated that teachers have input into evaluation decisions. Only about 40 percent indicated that they hold a pre-observation conference with the teacher. And there are indications (that will be discussed presently) that both pre- and post-observation conferences by the principals in the study might not be the cooperative, give-and-take activities typical of a collegial system. The discussion which follows highlights several components of present evaluation systems which, according to the survey results, seem to be dominated by administrators.

Teacher Participation in Evaluation Decisions

During the past two decades, teachers' organizations have dedicated significant negotiating efforts to increasing the level of participation by teachers in substantive evaluation decisions. Several research studies which were reviewed in Chapter II support the idea that teachers who participate in developing the process in which their performance will be evaluated are more likely to be satisfied with the outcomes of the evaluation. The NSPRA report indicates that teachers want to be involved in designing the evaluation process, in establishing the policies, goals, and objectives of evaluation, and in designing the evaluation form. In all areas other than designing the form, however, teacher participation

in such evaluation decisions, although having gained a solid foothold, is still lacking in two-thirds of the schools in the survey, according to the respondents.

Yet, even this relatively low level of participation may be higher than in other sections of the country. Between 30 and 40 percent of the principals in the present study indicated that teachers participate in the substantive decisions of an evaluation process. Comparing these results with Tobia's study of Pennsylvania teachers, in which he found little or no teacher involvement in evaluation decisions anywhere in the state, it is apparent that the level of involvement of teachers in the schools covered by the present study, though not high, is higher than might be expected.[1] The NSPRA report, however, noted a strong trend toward teacher participation in evaluation decisions in 1974, over 10 years prior to the present study. Whether the present findings are indicative of that trend is impossible to know, but there is a clear implication that they are.

The survey results also point to the fact that teachers' organizations are successfully negotiating teacher participation into professional contracts. Three of the respondents spontaneously noted on the survey form, for example, that teachers in their districts participate in evaluation decisions through the collective bargaining agreement. If the results of the present study are indicative of a trend toward teacher involvement in evaluation systems, then the trend is a healthy one, for numerous studies cited in Chapter II have concluded that teacher

---

[1]Tobia, abstract.

participation in establishing evaluation procedures leads to higher morale, better attitudes, and greater benefits from evaluation. Nevertheless, only about one-third of the principals in the present survey reported that teachers are involved in substantive evaluation decisions. Thus, in spite of a noted trend to the contrary, the majority of the principals in the study seem to operate evaluation systems which do not encourage teacher participation.

Pre-Observation and Post-Observation Conference

Pre-observation and post-observation conferences are major components of any collegial evaluation system. Both are integral parts of the clinical supervision model and the MBO model. The preconference is an activity which enhances the collegial nature of the evaluation process and fosters a spirit of trust and cooperation between the teacher and the evaluator.[1] As conceptualized by Cogan, Goldhammer and others, the preconference is a formalized activity during which specific exchanges are to take place between the participants. Whether the respondents to the present survey had that concept in mind when they responded to the question is impossible to know. It is likely that some did. It is more likely, however, that the concepts of a preconference held by the responding principals ranged from the classic formalized activity which was described, to a very informal and brief conference for the purpose of establishing a few details about the observation. Nevertheless, whatever the concept, only 41 percent of the principals in the survey indicated that they usually hold a preconference. It is not surprising

---

[1]Manatt, p. 5.

that the preconference, with its time requirements did not receive overwhelming support from the principals in the survey group, who are usually under no obligation to hold a preconference.

The post-observation conference, however, is included in evaluation systems by nearly all of the respondents.[1] The post-observation conference, according to Manatt, is seen by several leading writers as the most important for changing teachers' behavior. In practice, however, the postconference, like the preconference, can range from an informal, brief conversation during which the evaluation form is given to the teacher, to a formalized, collegial, cooperative conference during which important information and feedback is exchanged in the pursuit of improving instructional skills. No definition was given to the term postconference in the survey instrument; it was simply referred to by name. Nevertheless, from other responses on the survey, several characteristics of both pre- and postconferences which were probably conceptualized by principals in the study can be inferred.

For example, 98 percent of the principals in the study indicated that they provide supportive comments during the evaluation process. Riddile's study concluded that such supportive feedback contributes to a perception of effective evaluation.[2] However, only 22 percent of the respondents indicated that evaluation is a cooperative effort and only a third indicated that the teacher is treated as a colleague during the evaluation. Both Goldhammer and Cogan emphasized that simply going

---

[1] Ibid., p. 5.
[2] Riddile.

through the 5 steps of clinical supervision, which includes a pre- and postconference, is not enough for a quality evaluation. There must be a spirit of colleagueship and mutuality in the relationship.[1] Moreover, according to Manatt, teachers prefer an open and democratic approach during a conference.[2] In spite of the fact that supportive comments are provided during the process, neither colleagueship nor a democratic approach seems to exist in the majority of the conferences held by principals in the study.

A review of research studies by Bolton revealed that teachers accept feedback more readily when the focus is on improving performance.[3] Only 19 percent of the respondents, however, indicated that their evaluation programs are diagnostic or formative, making it unlikely that the focus of the postconference is on effective improvement of instruction.

Consideration of the response rates of each of the above items, therefore, leads to a realization that the concepts of pre- and postconferences held by a majority of the principals in the study are probably not consistent with the kind of conference envisioned by major writers such as Cogan, Redfern, Manatt, and others. In the majority of the evaluation conferences being conducted by the respondents to the survey there would appear to be little collegial give-and-take in the pursuit of improving instruction. As a result, pre- and postconferences, as reported by principals in the survey, rather than being indicators of a collegial evaluation system, more likely tend to be non-collegial components of administrator-dominated evaluation systems, in more than two-

---

[1]Reavis, p. 580.
[2]Manatt, p. 13.
[3]Bolton, p. 31.

thirds of the schools covered by the survey. This finding is consistent with that of Manatt who reports that conferences, by teachers' own accounts, do little to produce growth for teachers.[1]

Mutual Trust

Ideally, a collegial evaluation system should be marked by a high level of trust between the participants. Only half of the principals who responded to the survey, however, indicated that mutual trust exists in their evaluation programs, and there is reason to doubt the accuracy of even that level, when other responses are taken into consideration. For example, when asked to characterize the evaluation systems in their schools, more than 59 percent of the principals in the study indicated that evaluation is an administrative task. Only 22 percent indicated that it is primarily a cooperative effort. These findings are contrary to research studies cited in Chapter II, which concluded that teachers who share input into the evaluation process tend to be more willing to be evaluated and tend to feel more confident about evaluation outcomes than those who are denied input.[2] Other studies indicated that teachers who participate as colleagues rather than subordinates in a performance evaluation tend to place more trust in the process and the evaluator.[3] Nearly 60 percent of the principals in the present study, however, when characterizing their evaluation systems, indicated that evaluation programs under their direction tend to treat teachers more as

---

[1]Manatt, p. 13.

[2]Houston; Henderson, p. 63; Miller; Paulin, p. 9; Tobia, abstract.

[3]Alfonso and Goldberry, p. 106.

employees than individuals. Therefore, one would expect that 60 to 70 percent of the teachers who are evaluated by the principals in the study should feel some lack of trust in the evaluation program and be somewhat unwilling to participate. This is contradictory to the responses of 54 percent of the principals, who indicated that evaluation in their schools is characterized by mutual trust.

Although the reason for the dissonance which was found among the responses is beyond the scope of this study, one can speculate that either several of the principals in the survey group inspire trust in themselves even when their evaluation systems do not encourage trust, or they are inaccurate in their inferrences of teachers' true feelings. Whatever the cause, the response on the mutual trust is inconsistent with other responses, and a more objective reporting would discover a much lower level of trust than that reported by the respondents. Similarly, the pre- and post-observation conferences were shown to be misleading as indicators of collegiality. These findings, coupled with a response that indicated a rather low level of teacher participation in evaluation decisions, point to low collegiality and high principal-dominance in the majority of the evaluation programs being conducted by principals in the survey.

Ideal Practice Versus Present Practice

After determining the level of congruence between principals' reported evaluation practices and teachers' preferences, the second purpose of the study was to determine the difference between the responding principals' concepts of ideal evaluation practices and the evaluation programs that they presently report operating. A significant difference was noted, particularly in 2 related areas: teacher participation in evaluation decisions, and collegiality in the evaluation process.

As discussed previously, only about one-third of the principals in the study indicated that teachers participate in substantive evaluation decisions. Fifty to 75 percent, however, believe that teachers should ideally participate in some or all of those decisions. Such a cooperative approach to evaluation would lead to improved relationships between teachers and administrators, as well as an increased sense of trust by the teachers, and a greater willingness by the teachers to be evaluated, according to several studies which have been cited. In short, greater participation by teachers in the substantive decisions of evaluation, as advocated by a majority of the principals in the study, would enhance a spirit of collegiality and cooperation among teachers and principals. This is consistent with other responses on the survey, in which a majority of the respondents recommended that ideal evaluation systems be characterized by collegiality, mutual trust, and cooperation. Over 60 percent of the respondents indicated that the teacher should ideally be treated as a colleague during the evaluation, but only 33 percent reported that the teacher is treated as a colleague during present evaluation programs. Over 60 percent of the respondents indicated that the

evaluation process should ideally be a cooperative effort, but only 22 percent reported that evaluation is actually a cooperative effort during present evaluation programs. Other responses indicative of collegiality revealed similar significant differences between present practice and ideal practice: mutual trust, 53 percent present practice, 84 percent ideal practice; non-threatening, 58 percent present practice, 70 percent ideal practice.

After research studies have indicated that collegial systems of evaluation are superior to administrator-dominated systems, and when a majority of the principals in the study have indicated that they recognize the importance of teacher participation and increased cooperation, one can only speculate as to why the majority of principals in the study do not implement more collegial, participatory evaluation programs. Several possible reasons exist for not implementing that kind of evaluation system: time constraints, constraints from the superintendent or from district policies, or a lack of training on the part of the principal. A fourth reason might be a lack of confidence in teachers to be effective partners in an evaluation process.

Several studies of teacher evaluation practices have determined that lack of time is often a significant problem in implementing some of the more collegial evaluation models of evaluation, such as MBO and clinical supervision. Time might be the factor causing the difference between present and ideal practices in the present study.

Another reasonable explanation for the differences might be that the superintendent or the school district requires certain practices which make collegial systems more difficult to operate. In the interviews,

however, each interviewee was asked whether he or she felt constrained by the superintendent or by district policies in the implementation of teacher evaluation. No one indicated such pressure, although only 3 of the interviewees were seen to be operating systems which were truly independent of, or significantly divergent from district-wide evaluation systems. In all 3 cases, a great deal of information was volunteered by the principals, as were specific forms which were noted to be their own, while in other interviews, the responses generally referred to systems which more obviously originated in the district central office. Although these were only personal impressions, and have no statistical significance to the study, the clear impression was that most of the principals who were interviewed were operating evaluation systems which were not of their own development, and had chosen not to expand on those systems. Therefore, it is quite possible that constraints, either explicit or implicit, from the central office result in the implementation of evaluation systems which are less collegial than principals in the survey ideally think they should be.

Evaluator training is essential to the implementation of a successful collegial evaluation system. It is possible that the majority of the principals require such training before they would feel confident enough to implement more collegial evaluation systems. Conversely, principals in the survey might feel that teachers require training before they are competent enough to participate in a more cooperative system of evaluation. Whatever the cause might be, the fact remains that principals in the survey tend to operate systems that are significantly less collegial and humanistic than either they or teachers believe they should be.

INTERVENING VARIABLES

Principals' responses were consistent across all six factors which could potentially have effected the outcomes of the study. Although a reasonable expectation would have been that years of experience as a principal or as a teacher might have altered either the principals' performance or their opinions of ideal practice, such was not the case. Further, it might have been expected that the size of a school's professional staff might effect the conditions of the evaluation system. This did not occur either. In addition, neither economic factors of the district nor the educational level of the principal effected the results of the study.

CHAPTER V

SUMMARY AND RECOMMENDATIONS

One hundred thirty-eight suburban elementary principals were mailed questionnaires in which they were asked to indicate the frequency with which they include each of 28 evaluation characteristics in their teacher evaluation programs. These characteristics have been shown by a literature review to be associated with higher teacher satisfaction with evaluation. Additionally, the principals were asked their opinion regarding the importance of each characteristic to an ideal system of evaluation. Eighty-one principals responded. An additional 22 principals were interviewed face-to-face, but the results of the interviews were not included in the data analysis.

Response frequencies were analyzed according to a Scale of Congruence, which ranged from 0 percent rate of inclusion in present evaluation systems to 100 percent, and from 0 percent rate of importance in ideal evaluation systems to 100 percent. The 28 teacher-preferred characteristics of evaluation were organized into 6 distinct clusters: I Teacher Participation in Evaluation Decisions, II Activities Prior to Classroom Observation, III Activities Following Classroom Observation, IV Training and Competence of the Evaluator, V Conditions of the Evaluation, VI Purposes and Criteria of the Evaluation. The differences between present and ideal practice in each of the clusters was analyzed using the Multi-variate Analysis of Variance (MANOVA). The effects of 6 intervening

variables (Years as a teacher, Years as a principal, Years in education, Size of staff, Highest degree earned, and Average family income in the district) were analyzed using a Multiple Discriminent Analysis.

CONCLUSIONS

1. Suburban elementary principals who responded to the survey, regardless of the number of years as a teacher, as a principal, or as an educator, and regardless of the size of the school they work in, the highest degree they have earned, or the economic conditions of the district they serve, reported that they tend to include procedures in their evaluation systems which are congruent with characteristics of evaluation that have been associated with increased teacher satisfaction with evaluation. Congruence is not consistent across all characteristics, and there are examples of very high and very low congruence. Nevertheless, the tendency toward overall agreement between principals' evaluation systems and teacher-preferred evaluation characteristics is clear.

2. It was discovered that principals' concepts of the characteristics of ideal evaluation systems are significantly different than the characteristics of the evaluation systems they presently report operating. The difference occurs primarily among the evaluation characteristics which would be associated with a humanistic, collegial system of evaluation. Moreover, the principals' concepts of ideal evaluation systems, as indicated in the survey, are closer to the way teachers prefer to be evaluated than are the evaluation systems reportedly in operation.

3. Principals in the survey reported that they tend to be implementing most of the high-priority characteristics of evaluation which have been supported by teachers' organizations across the country.

4. Evaluation practices which produce the greatest congruence with teacher-preferred characteristics tend to be those that are non-discretionary, over which the principal has little control, e.g., providing

a copy of the evaluation report. Conversely, the practices which produce the least congruence with teacher-preferred characteristics of evaluation tend to be those that are more discretionary, over which the principal has greater control, e.g., developing a cooperative, collegial evaluation program.

5. Principals in the survey reported that they tend to operate evaluation systems that are more principal-dominated than collegial in nature, although from the opinions indicated in the survey, principals tend to believe that teacher evaluation should be more collegial than it is.

6. Principals in the survey believe that teachers should be involved much more in the planning and implementation of evaluation systems than they are at present. That involvement should include developing policies as well as setting goals and designing programs.

7. Present evaluation systems reported to be operated by principals in the survey tend to be summative, and not very diagnostic, and therefore probably do not focus strongly on improvement of instruction.

RECOMMENDATIONS

1. School districts who are serious about developing high quality evaluation systems should acknowledge the dilemma of the conflicting purposes of teacher evaluation. One solution would be to develop a two-tier evaluation system, dedicating one tier to the improvement of instructional skills, and calling it "professional growth", while maintaining the non-collegial-sounding title "evaluation" for a second tier reserved for the pragmatic personnel decisions of tenure and continued employment. Such a system is consistent with the responses of the principals in the present survey, who indicated that teacher evaluation has the dual purpose of improvement of instruction and tenure and termination decisions. In the absence of such a two-tier system, conflict of purpose will likely continue and congruence between evaluation practices and teachers' preferences will remain low.

2. Both principals and teachers would benefit from training in the process of evaluation. A key ingredient in the development of excellent school systems is high-quality supervision, which can only come about through training. Such training would be pragmatic and practical, recognizing the limitations of time and the natural resistance of both teachers and principals to become involved in a time-consuming, somewhat threatening process.

3. School districts should develop a process whereby teachers can participate actively in establishing and maintaining the policies, the goals, and the procedures of evaluation. This participation should exceed simple review and ratification of district-developed plans, and in fact should involve collegial planning and the establishment of mutual goals.

4. District administrators must recognize the relationship between high quality supervision and educational excellence by providing the time, the training, the resources, and the leadership necessary to implement a true collegial system of professional growth supervision and evaluation in their schools.

Implications for Futher Study

1. The present survey discovered a significant difference between present practice and principals' concepts of ideal practice regarding teacher evaluation, although no determination of the reason for this difference was made. Further study is needed to determine why principals tend not to implement evaluation programs that reflect their concepts of the ideal. Several possible reasons were offered. Perhaps additional study can verify whether any, all, or none of these were the true reasons for the difference.

2. Because the present research is limited to a group of primarily suburban elementary principals, it would be of interest to expand the research to include a larger cross-section of principals. Additionally, it would be expected that high school principals, because of the nature of their jobs, might respond quite differently to many of the questions. A separate study, therefore, concentrating on high school principals, could produce complementary results to the present study.

3. One of the limitations of the present study is that concepts such as preconference and postconference were not defined, allowing the respondents to maintain their individual conceptions as they responded to the survey. Therefore, although congruence appeared to occur on many items, more precise definition of concepts might have dispelled the appearance of congruence. Conversely, more precision might have verified congruence. Nevertheless, additional research that would define with some precision the concepts which were studied and would measure more precisely the implementation of those concepts is necessary before there can be generalization of the results.

4. One of the conclusions of the study was that teachers' preferences and principals' evaluation practices are more congruent in activities that are non-discretionary. The source of most decisions impacting on those non-discretionary activities is the district superintendent. This study did not attempt to discover the nature of that impact. Further study is needed to determine the congruence between principals and superintendents regarding evaluation issues, and whether that relationship impacts significantly on the congruence between teachers' preferences and principals' evaluation practices.

5. The conclusion that high congruence occurs among non-discretionary activities was made on the basis of inference and non-scientific judgments. Although useful to the purposes of the study, such inferential information limits the extent to which valid conclusions can be made. Further research should focus on the degree of a principal's autonomy surrounding evaluation issues, and the true impact of that autonomy on the evaluation process.

6. Finally, one of the questions on the survey sought to determine whether the principals in the study had received training in evaluation. The question was rudimentary and did not attempt to define the nature or quality of the training. It would be helpful to know how the quality and amount of training would effect congruency between principals' evaluation practices and teachers' preferences. Further study is needed to determine the impact of specific training in evaluation on the study results.

BIBLIOGRAPHY

Ackerson, Gary Lee. "A Comparison of Criteria Used in the Selection and the Evaluation of Teachers." Ph.D. Dissertation, University of Iowa 1977.

Ahnell, I.V., and Hawn, Horace C. "Self-Evaluation Through Video-Tape Recordings." In Observation Methods in the Classroom, pp. 61-64. Edited by Charles W. Beegle and Richard M. Brandt. Washington, D.C.; Association for Supervision and Curriculum Development, 1973.

Alfonso, Robert J., and Goldsberry, Lee. "Colleagueship in Supervision." In Supervision of Teaching, pp. 90-107. Edited by Thomas J. Sergiovanni. Alexandria, Va.: Association for Supervision and Curriculum Development, 1982.

Allen, Paul M.; Barnes, William D.; Reece, Jerald L.; Roberson, E. Wayne. Teacher Self-Appraisal. Washington, Ohio: Charles A. Jones Publishing Co., 1970.

Arreola, Raoul A. "Students Can Distinguish Between Personality and Content/Organization in Rating Teachers." Phi Delta Kappan 65 (November 1983): 222-223.

Bailey, Gerald D. Teacher Self-Assessment: A Means For Improving Classroom Instruction. Washington, D.C.: National Education Association, 1981.

Baker, Lambert W. "An Analysis of Some Assumed Predictors of Success in Teaching." Report to the California Educational Research Association, Los Angeles, 14 March 1969.

Barnes, Barbara Kininessi, "An Identification of Perceived Inservice Training Needs of Florida Public School Principals Relative to Teacher Evaluation." Ph.D. Dissertation, Florida state University, 1983.

Barr, Arvil S.; Eustice, David E.; and Noe, Edward J. "The Measurement and Prediction of Teacher Efficiency." Review of Educational Research 25 (June 1955): 261-269.

Barr, Arvil S. "Measurement of Teacher Characteristics and Prediction of Teacher Efficiency." Review of Educational Research 22 (June 1952): 171-172.

Barr, Arvil S.; Burt, William H.; Brueckner, Leo J. Supervision. New York: Appleton-Century-Crofts, Inc., 1938.

Barr, Arvil S. The Measurement and Prediction of Teacher Effectiveness. Madison, Wisconsin: Dunbar Publications, 1967.

Barr, Arvil S. "Measurement and Prediction of Teacher Efficiency." Journal of Experimental Education 16 (1948): 203-283.

Bassin, William M. "A Note on the Biases in Students' Evaluations of Instructors." Journal of Experimental Education 43 (February 1974): 16-17.

Bauer, Shirley S. "Perceptions of Selected Idaho Educators Regarding Supervision." Ph.D. Dissertation, University of Idaho, 1975.

Beard, Larry Dean. "Practices and Attitudes of Iowa High School Principals Regarding Teacher Evaluation Criteria." Ed.D. Dissertation, Drake University, 1975.

Beegle, Charles W., and Brandt, Richard M., ed. Observational Methods in the Classroom. Washington, D.C.: Association for Supervision and Curriculum Development, 1973.

Blackbourn, Richard Lee. "The Relationship Between Teachers' Perceptions of Supervisory Behaviors and Their Attitudes Toward a Post-Evaluative Conference." Ed.D. Dissertation, Mississippi State University, 1983.

Bloom, Benjamin S. "Testing Cognitive Ability and Achievement." In Handbook of Research on Teaching, pp. 379-397. Edited by N. L. Gage. Chicago: Rand McNally, 1963.

Blumberg, Arthur. Supervisors and Teachers: A Private Cold War. Berkeley, California: McCutchen Pub. Co., 1974.

Blumberg, Arthur, and Amidon, Edmund. "Teacher Perceptions of Supervisor-Teacher Interaction." Administrator's Notebook 14 (September 1965).

Blumberg, Arthur; Amidon,Edmund; and Weber, W. "Supervisor-Teacher Interaction as Seen by Supervisors." Unpublished manuscript. Philadelphia: Temple University, 1967.

Bodine, Richard. "Teachers' Self-Assessment." In School Evaluation: The Politics and Process, pp. 169-219. Edited by Ernest R. House. Berkeley, California: McCutchen Publishing Co., 1973.

Bolton, Dale L. Teacher Evaluation. PREP Reports, no. 21. Washington, D.C.: HEW, Office of Education, National Center for Education Communication, 1972.

Borich, Gary D.; Malitz, David; and Kugle, Cherry L. "Convergent and Discriminant Validity of Five Classroom Observation Systems: Testing A Model." Journal of Educational Psychology 70 (1978): 119-28.

Brophy, Jere E. "Stability of Teacher Effectiveness." American Educational Research Journal 10 (Summer 1973): 245-52.

Brungardt, Karl Edmund, "A Procedural Model for the Improvement of Instruction Through the Evaluation/Supervision of Teachers." Ed.D. Dissertation, University of Minnesota, 1979.

Bryant, Gerald, and Haack, Frank. "Appraisal: Peer-Centered and Administration-Centered." Educational Leadership (May 1977): 608-612.

Burnett, Richard Mulkey. "An Evaluation of Certain Criteria Used In Selecting Teachers for Public Schools." Ed.D. dissertation, University of Texas, 1966.

Campbell, Ona L. "The Relationship Between Eight Situational Factors and Low Scores on the Leadership Dimensions of Instructional Superisors." Doctoral dissertation, North Texas State College, 1961.

Carpenter, Dennis Arthur. "A Study of Teachers' Perceptions of Two Teacher Evaluation Processes." Ed.D. dissertation, George Peabody College, 1981.

Cartlidge, Arthur James. "A Study of Self and Principals' Evaluation of Beginning Teachers in Selected Mississippi Public Schools," Ed.D. dissertation, Mississippi State University, 1983.

Christner, C.A. "The J. R. Syndrome: Administrator Bias in Teacher Evaluation." presented at the Annual Meeting of the American Educational Research Association, April 1981.

Claye, Clifton M. "Lola Gets What Lola Wants From Supervision." The Journal of Educational Research 56 (1963): 358.

Cogan, Morris L. Clinical Supervision. Boston: Houghton Mifflin, 1983.

Cole, Charles C. "A Comparison of Two Methods of Teacher Evaluation." Ed.D. dissertation, North Texas State University, 1977.

The Committee on IOTA. Measuring Teacher Competence: Research Backgrounds and Current Practice. California Teachers Association, N.D.

Cooley, John. "The Role of Personal Controlability in Teacher Receptivity to Evaluation." Ed.D. dissertation, University of California, Santa Barbara, 1980.

Cookson, Peter W., Jr. "Teacher Evaluation in Independent Schools." Independent School 39 (May 1980): 49-54.

Cornett, Joe D. "Effectiveness of Three Selective Admissions Criteria in Predicting Performance of First-Year Teachers." Journal Educational Research 62 (February 1969): 247-50.

Cornwell, C. D. "Statistical Treatment of Data from Student Teaching Evaluation Questionnaires." Journal of Chemical Education 51 (March 1974): 155-160.

Crenshaw, Harrison M. II, and Hoyle, John R. "The Principals' Headache--Teacher Evaluation." NASSP Bulletin 6 (February 1981): 37-45.

Davis, Barbara Irene. "The Status of Teacher Evaluation Practices in Arizona and a Proposed Model." Ed.D. dissertation, University of Arixona, 1974.

Davis, Betty Gale Mullen. "The Relationship Between School Climate and the Congruency of Perceptions of Elementary School Principals and Teachers Concerning Teacher Evaluation." Ed.D. dissertation, University of Mississippi, 1982.

Detzner, Daniel F. "Teacher Evaluation: A Self-Appraisal Method: The University of Minnesota General College Studies 11 (1974-75).

Dieter, Don L. "Teacher Evaluation: A Strategy for Science Teachers:" Paper presented at the Fourth Annual Meeting of the North Carolina Science Teachers Association, Asheville, N.C., November 1972.

Dubin, Robert, and Tarveggia, Thomas. The Teaching Learning Paradox. Eugene, Oregon: Center for Advanced Study of Educational Administration, 1968.

Eads, Albert Edward, Jr. "A Study of the Attitudes of Teachers Toward a 'Supervision by Objectives' Teacher Evaluation Model." P.h.D. dissertation, University of South Carolina, 1974.

Eckard, Pamela J., and McElhinney, James H. "Teacher Evaluation and Educational Accountability." Educational Leadership (May 1977): 613-618.

Ecker, Homer. "How Valid are Expert Opinions about Effective Teaching?" Phi Delta Kappan 62 (October 1980): 131-134.

Educational Testing Service. Teachers Do Make a Difference. Princeton: Educational Testing Service, 1976.

Egan, Leona Rust. "Identifying the Effective Teacher: An Evaluation of Criteria Used in Selecting U.S. Navy Instructors." Ed.D. dissertation, American University, 1964.

Eisner, Ei-liot W. "An Artistic Approach to Supervision." In Supervision of Teaching, pp. 53-56. Edited by Thomas J. Sergiovanni. Alexandria, Va.: Association for Supervision and Curriculum Development, 1982.

Epstein, Charlotte, "Student Grade Teachers? Some Faculty Attitudes." Community and Junior College Journal 44 (April 1974): 33.

Fagan, Leonora Perry. "Teacher Evaluation: Factors Influencing the Perceptions of Evaluators: Ph.D. dissertation, University of Toronto, 1983.

Fallman, John; Lavely, Carolyn; Silverman, Stuart; and Merica, John. "Student Raters' Referents in Rating College Teaching Effectiveness." Journal of Psychology 86 (March 1974): 247-249.

Farris, John Alvin, "Teacher Evaluation Process in Evangelical Christian Schools." Ed.D. dissertation, University of the Pacific, 1983.

Ferguson, Victor Simon. "Principals' Perceptions Concerning Teacher Evaluation in the State of Washington." Ed.D. dissertation, Seattle, University, 1981.

Florida Educational Research and Development Council. "Standards and Procedures Used for Evaluating Classroom Instruction of Annual Contract Teachers." FERDC Research Bulletin 1 (May 1965): 4-6.

Fouke, Linda Graves. "A Comparative Study of Teacher Evaluations Written by Administrators." Ed.D. dissertation, Columbia University Teachers College, 1983.

Fox, Ronald. "Student Evaluation of Teacher as a Measure of Teacher Behavior and Teacher Impact on Students." Journal of Educational Research 77 (September-October 1983): 16-21.

Fraser, Ken P. "Supervisory Behavior and Teacher Satisfaction." Ed.D. dissertation, Montana State University, 1979.

Frudden, Sally, and Manatt, Richard P. "Lesson Analysis--The Neglected Key to Teacher Performance Evaluation." Occasional paper no. 82-5, School Improvement Model Project. Ames, Iowa: College of Education Iowa State University, 1982.

Gage, N.L. "Desirable Behaviors of Teacher." Urban Education 1 (Winter 1965): 85-95.

Garmen, Noreen B. "The Clinical Approach to Supervision." In Supervision of Teaching, pp.35-52. Edited by Thomas J. Sergiovanni. Alexandria, Va.: Association for Supervision and Curriculum Development, 1982.

Gasper, Francine Marie Adams. "Teachers' Perceptions of Increased Professional Competence as an Outcome of the Stull Evaluation Process." Ph.D. dissertation, The Fielding Institute, 1980.

Glass, Gene V. "Teacher Effectiveness." In Evaluating Educational Performance, pp. 11-32. Edited by Herbert J. Walberg. Berkeley, California: McCutchan Publishing, 1974.

Glass, Gene, "A Paradox About Excellence of Schools and the People in Them." The Educational Researcher 4 (March 1975): 9-13.

Goldhammer, Robert; Anderson, Robert H.; and Krajowski, Robert J. Clinical Supervision: Special Methods for the Supervision of Teachers. New York: Holt, Rinehart and Winston, 1980.

Goldstein, William. "An Enlightened Approach to Supervising Teachers." The Clearinghouse 46 (March 1972): 391-394.

Grant, Stephen, and Carvell, Robert. "A Survey of Elementary School Principals and Teachers: Teacher Evaluation Criteria." Education 100 (Spring 1980): 223-226.

Grossman, Barney. "Teachers' Methodological Emphasis and Their Evaluation of Supervisory Practices." Ed.D. dissertation, Rutgers University, 1967.

Group for the Study of Effective teaching. "Teaching Effectiveness Evaluation Project: Final Report." Chapel Hill, North Carolina: School of Education, University of North Carolina at Chapel Hill,' 1983.

Hamachek, Don. "Characteristics of Good Teachers and Implications for Teacher Education." Phi Delta Kappan 50 (February 1969): 341-344.

Hansen, J. Merrell. "School Effectiveness = Teacher Effectiveness." High School Journal 64 (February 1981): 222-226.

Harris, Ben M. "Teacher Evaluation as a Development System." ERIC #ED224800, 1983.

Harris, Mary B. "Sex Role Stereotypes and Teacher Evaluation." Journal of Educational Psychology 67 (December 1975): 751-756.

Hauge, Barbara Jean Swenson. "A Study of Teacher Evaluation Practices and Perceived Attitudes of Those Practices by Elementary School Principals in the Twin Cities Metropolitan Area." Ed.D. dissertation, University of Minnesota, 1981.

Heitzman, William Ray, and Starpoli, Charles. "Teacher Characteristics and Successful Teaching." Education 95 (Spring 1975): 298-300.

Henderson, Lester F. "Elementary Teachers' Satisfaction and Morale and Perceived Participation in Decision-Making." Ed.D. dissertation, University of Arkansas, 1967.

Hendrix, Clifford L. Jr. "A Study of Supervisory Practices with References to Classroom Observations and Conferences as Perceived by Teachers, Principals and Supervisors in Tennessee." Ed.D. dissertation, University of Tennessee, 1976.

Heuss, R nald. "Teacher Evaluation Purposes, Procedures and Instrument in Texas Public Schools." Ed.D. dissertation, Baylor University, 1984.

Hickman, George Augustus. "A Study of Teacher Evaluation Systems in the Province of Newfoundland and Labrador." Ed.D. dissertation, University of Toronoto, 1983.

Holley, Freda M. "Personnel Evaluation: Essentials for Success." In Merit Pay and Evaluation, Hot Topics Series 1983-84, pp.6-9. Bloomington, Indiana: Phi Delta Kappa Center on Evaluation Development and Research, 1983.

Holley, Freda M., and Hickman, Randall C. "Research on Teacher Evalutaion: Needs and Realities." Paper presented at the Annual Meeting of the American Educational Research Association, Los Angeles, April 1981.

Holmes, David S. "Effects of Grades and Disconfirmed Grade Expectancies on Students' Evaluation of Their Instructor." Journal of Educational Psychology 63 (April 1972): 130-133.

Hosford, Phillip, and Neuenfeldt, John. "Teacher Evaluation Through Videotape: Hope or Heresy?" Educational Leadership 36 (March 1979): 418-422.

Hosford, Philip L., and Martin, Jeannette V. "Historical Analysis of Videotape Usage in Predicting Teacher Effectiveness." College of Education Dialogue Series. Las Cruces, New Mexico: New Mexico State University, 1980.

Houston, Samuel Lawson. "An Analysis of Perceptions of Tennessee Public School Teachers, Building-Level Administrators and Central Office Administrators Toward Teacher Evaluation." Ed.D. dissertation, University of Tennessee, 1981.

Hoyt, Donald P. "Identifying Effective Teaching Behaviors." Final Report. Manhattan: Kansas State University, December 1969.

Hyman, Ronald T. "Supervision and the Ideal Teacher-Student Relationship." In Selected Articles for Elementary School Principals, pp. 67-72. Washington, D.C.: National Education Association, Deptmart of Elementary School Principals, 1968.

Ingils, Chester R. "Let's Do Away with Teacher Evaluation." The Clearing House 44 (April 1970): 451-456.

Irvine, Jacqueline Jordon. "The Accuracy of Pre-Service Teachers' Assessments of Their Classroom Behavior." Journal of Research and Development in Education 17 (Fall 1983): 25-31.

Jensen, Mary A. "How Teachers View Teacher Evaluation." Education 102 (Winter 1981): 130-137.

Jones, Sheridan Davis. "A Model for Identifying Evaluation Procedures that have a Positive Influence on Teacher Attitude." Ph.D. dissertation, Oregon State University, 1981.

Jones, Thomas W. "The Validity of Self-Evaluation of Competence by Special Education Teacher Trainees." Texas Tech Journal of Education 9 (Winter 1982): 23-27.

Johnson, I.T. "Why Supervise Teachers?" School Management 15 (October 1971):34.

King, Richard A. "Reliable Rating Sheet: A Key to Effective Teacher Evaluation." NASSP Bulletin 62 (December 1978): 21-26.

Kinney, Lucien B. "Developing Instruments to Direct Appraisal of Teacher Effectiveness." Mimeographed Local Collection of the ERIC Clearinghouse. (n.d.).

Kronowitz, Ellen, and Finney, Victoria. "Student Teachers' Performance." California Journal of Teacher Education 10 (Winter 1983): 84-101.

Larsen, Edwin M. "Students' Criteria for Responses to Teaching Evaluation Questionnaires." Journal of Chemical Education 51 (March 1974): 163-165.

Leese, Joseph. "Teacher Assessment and Consensus Preference." NASSP Bulletin 65 (November 1981): 19-27.

Levin, Benjy. "Teacher Evaluation--A Review of Research." Educational Leadership 37 (December 1979): 240-245.

Levin, Harry A. "New Perspective on Teacher Competence Research." Harvard Educational Review 24 (1954): 98-105.

Levine, Judith R. "When Colleagues Judge Colleagues." Teaching of Psychology 11 (February 1984): 38-39.

Lewis, James Jr. Appraising Teacher Performance. West Nyack, N.Y.: Parker Publishing Co. 1973.

Lieberman, Seymour. "The Effect of Changes in Roles on the Attitudes of Role Occupants." Human Relations 9 (November 1956): 401.

McCall, William S. Measurement of Teacher Merit. Bulletin no. 284. Raleigh, North Carolina: State Supt. of Public Instruction, April 1952.

McFaul, Shirley A., and Cooper, James M. "Peer Clinical Supervision: Theory vs. Reality." Educational Leadership 41 (April 1984):4-9.

McNeil, John D. "Concomitants of using Behavioral Objectives in the Assessment of Teacher Effectiveness." The Journal of Experimental Education 36 (Fall 1967): 69-74.

McNeil, John D. "Performance Tests: Assessing Teachers of Reading--A Proposal." The Reading Teacher 25 (April 1972): 622-627.

McNeil, John D. "Politics of Teacher Evaluation." In Handbook of Teacher Evaluation, chapter 15. Edited by Jason Millman. Beverly Hills, Calif.: Sage Publications, 1981.

McNeil, John D. "A Scientific Approach to Supervision." In Supervision of Teaching, pp. 18-24. Edited by Thomas J. Sergiovanni. Alexandria, Virginia: Association for Supervision and Curriculum Development, 1982.

McNeill, John, and Popham, James. "The Assessment of Teacher Competence." In Second Handbook of Research on Teaching, pp. 218-244. Edited by Robert M. W. Travers. Chicago: Rand McNally, 1973.

McNergney, Robert, and Satterstrom, Lyn. "Teacher Characteristics and Teacher Performance." Contemporary Educational Psychology 9 (January, 1984): 19-24.

Manatt, Richard P. "Teacher Performance Evaluation--Practical Application of Research." Occasional Paper no. 82-1. Ames, Iowa; Iowa State University College of Education, January 1982.

Manatt, Richard; Hidlebaugh, Everitt; and Palmer, Kenneth. "Improved Teacher Performance Evaluation with Improved Rating Scales: NASSP Bulletin 60 (September 1976): 21-23.

Margules, Morton, and Keith, Virginia M. "Study of Supervisors' Ratings of Most Effective and Least Effective Industrial Arts Teachers on Three Competency Dimensions." The Journal of Experimental Education 32 (Summer 1969): 37-44.

Marsh, Herbert W.; Overall, J.J.; and Kesler, Steven P. "Validity of Student Evaluations of Instructional Effectiveness: A Comparison of Faculty Self-Evaluations and Evaluations by Their Students." Journal of Educational Psychology 71 (April 1979): 149-160.

Medley, Donald M; Coker, Homer; Coker, Joan G; Lorentz, Jeffrey L.; Soar, Robert S.; Spauding, Robert L. "Assessing Teacher Performance from Observed Competency Indicators Defined by Classroom Teachers." Journal of Educational Research 74 (March-April 1981): 197-216.

Menges, Robert J. "The New Reporters: Students Rate Instruction." In Evaluating Learning and Teaching, pp. 59-72. Edited by C. Robert Pace. San Francisco: Jossey-Bass, 1973.

Meyer, D. Eugene. "A Survey of Public and Private Schools in Illinois Regarding Faculty, Administrative, and Supportive Personnel Evaluations, Procedures, and Instruments Used." Illinois School Research and Development 18 (Spring 1983): 15-20.

Miller, Duane Dean. "The Development of a Process Model to Utilize Peer Opinion for Teacher Evaluation in Selected Public Schools in Nebraska." Ed.D. dissertation, University of Nebraska, 1974.

Miller, Mary Michaelle. "The Evaluation Styles, Methods and Modes Preferred by Middle School Teachers and Administrators in an Urban Setting." Ed.D. dissertation, University of San Francisco, 1981.

Miller, Patricia Hayes. "A Survey of the Design and Implementation of Teacher Evaluation Systems in Tennessee Public Schools. " Ed.D. dissertation, University of Tennessee, 1980.

Millman, Jason, ed. Handbook of Teacher Evaluation. Beverly Hills, Calif.: Sage Publications, 1981.

Mirus, Rolf. "Some Implications of Student Evaluation of Teachers." Journal of Economic Education 5 (February 1973): 35-37.

Mitzel, H.E. "Teacher Effectiveness." In The Encyclopedia of Educational Research, 3rd ed. Edited by Chester W. Harris. New York; MacMillan Co., 1960.

Moeller, Gerald H., and Mahan, David J. The Faculty Team. Chicago: Science Research Association, 1971.

Moritz, William, and Martin-Reynolds, JoAnne. "Split-Screen Videotaping: The Genie in the Bottle." Educational Leadership 37 (February 1980): 396-399.

Morrison, Edward J. "Performance Observation as an Approach to Teacher Evaluation." Educational Horizons 52 (Summer 1974): 167-172.

Morrow, James R., Jr. "Some Statistics Regarding the Reliability and Validity of Student Ratings of Teachers." Research Quarterly. 48 (May 1977): 372-375.

Morsh, Joseph and Wilder, Eleanor, "Identifying the Effective Instructor: A Review of the Quantitative Studies, 1900-1952." Research Bulletin AFPTRC-T-54-44. Texas Lackland Air Force Base, 1954.

Musella, Donald. "Improving Teacher Evaluation." Journal of Teacher Education 21 (Spring 1970): 15-21.

Musella, Donald. "Open-Closed-Mindedness as Related to the Rating of Teachers by Elementary School Principals." Journal of Experimental Eduation 35 (Spring 1967): 75-79.

Muth, Rodney. "Power, Conflict and Consensus: Implications for Teacher Evaluation and School Climate." NJEA Review 54 (October 1980): 19-20.

National Education Association. The Elementary School Principalship in 1968. Washington, D.C.: National Educational Association, Department of Elementary School Principals, 1968.

National Education Association. Selected Articles for Elementary School Principals. Washington, D.C.: National Education Association, Department of Elementary School Principals, 1968.

National Education Association, "Evaluation of Teaching Competence." In NEA Research Bulletin, pp. 67-74. Washington, D.C.: National Education Association, 1969.

National School Public Relations Association. Evaluating Teachers for Professional Growth. Current Trends in School Policies and Programs, Arlington, Va.: National School Public Relations Association, 1974.

Neagley, Ross L., and Evans, N. Dean. Handbook for Effective Supervision of Instruction, 3rd Ed. Englewood Cliffs, New Jersey: Prentice-Hall, 1980.

Negley, Harold H. "Effectiveness in Supervision of Social Studies in Relation to the Extent of Authoritarianism in the Practice of Supervisors." Ed.D. dissertation, Indiana University, 1962.

New England School Development Council. "Teacher Competence and the Relation to Salary." Cambridge, Massachusetts.: New England School Development Council, 1956.

Newfield, John. "Accuracy of Teacher Reports: Reports and Observations of Specific Classroom Behavior." Journal of Educational Research 74 (November-December 1980): 78-82.

Noonan, Joann Alice. "An Analysis of Perceptions of Michigan Superintendents, Principals, and Teachers in Regard to Present Versus Preferred Teacher Evaluation Systems." Ed.D. dissertation, Western Michigan University, 1981.

Okey, James R., and Humphrey, Donald W. "Measuring Teacher Competence." Paper presented at the Annual Meeting of the National Association for research in Science Teaching, Chicago, April, 1974.

Ornstein, Allan C. "How Good are Teachers in Effective Student Outcomes?" NASSP Bulletin 67 (January 1983): 73-80.

Osborne, Grace S., and Hurlburt, Allan S. "Credibility Gap in Supervision." School and Society 99 (November 1971): 415-417.

Ovard, Glen F. "Teacher Effectiveness and Accountability." NASSP Bulletin 59 (January 1975): 87-94.

Pace, C. Robert , ed. Evaluating Learning and Teaching. New Directions for High Education series (Winter). San Francisco: Jossey-Bass, 1973.

Paulin, Pauline. "The Politics of Evaluation at the Local Level: A View Through Teacher's Perspectives." Paper presented at the Annual Meeting of the American Educational Research Association, Los Angeles, 13-17 April 1981.

Paulin, Pauline Jane. "The Influence of Teacher Autonomy on Resistance to Authority." Ph.D. dissertation, University of California, 1980.

Peck, Robert P., and Tucker , James A. "Research on Teacher Education." In Second Handbook of Research on Teaching, pp. 940-978. Edited by Robert M. W. Travers, Chicago; Rand McNally, 1973.

Pederson,K. George. "Improving Teacher Effectiveness." Education Canada 15 (Fall 1975): 12-20.

Perrone, Michael. "In Connecticut: What Principals Think of Statute on Evaluation.: NASSP Bulletin 62 (December 1978): 71-75.

Peters, C. D., and VanVoorhis, W. R. Statistical Procedures and Their Mathematical Bases. New York: McGraw-Hill, 1940.

Peterson, Penelope L., and Walberg, Herbert J. Research on Teaching: Concepts, Findings and Implications. Berkely, California, McCutchan Publishing Co. 1979.

Popham, W. James. "Found: A Practical Procedure to Appraise Teacher Achievement in the Classroom." In Accountability for Teachers and School Administrators, pp. 75-77. Edited by Allan C. Ornstein. Belmont, California: Fearon Publishing, 1973.

Popam, W. James. Evaluating Instruction. Englewood Cliffs, N.J. Prentice-Hall, 1973.

Popham, W. James. "Performance Tests of Teaching Proficiency: Rationale, Development, and Validations." Educational Research Journal 8 (January 1971): 105-117.

Povey, R. M. "A Comparison of Three Methods of Evaluating Teaching Performance in a College of Education." British Journal of Educational Psychology 45 (1975): 279-285.

Queer, Glenn. "An Analysis of Teacher Rating Scales: A National Survey." Washington, D.C.: U.S. Department of Health, Education and Welfare, June 1969.

Ramsay, Russell G. "Teachers' Perceptions of the Design and Implementation of Teacher Evaluation Systems in Tennessee Public Schools." Ed.D. dissertation, University of Tennessee, 1980.

Reavis, Charles A. "Clinical Supervision: A Review of the Research." Educational Leadership 35 (April 1978): 580-584.

Redfern, George B. Evaluating Teachers and Admnistrators: A Performance Objectives Approach. Boulder, Colorado: Westview Press, 1980.

Richards, William Howard III. "The Effect of a Professional Growth and Evaluation Cycle Upon Experienced Teacher Attitudes." Ed.D. dissertation, Boston, University, 1983.

Riddile, Marianne Margrave Webb. "Teachers' Perceptions of an Evaluation by Objectives Process in a Large Suburban School District." Ed.D. dissertation, George Peabody College for Teachers, 1981.

Robinson, John J. "The Written Classroom Observation Report: Its Use and Effectiveness in Connecticut Senior High Schools." Ed.D. disser tation, Hofstra University, 1975.

Robinson, John J. and Lee, John H. Jr. "Evaluation: Can We Agree?" NASSP Bulletin 62 (December 1978): 15-20.

Rogers, Carl R. "Significant Learning: In Therapy and Education." Educational Leadership 16 (January 1959): 232-242.

Rosenshine, Barak. "Evaluation of Classroom Instruction." Review of Educational Research 40 (April 1970): 279-300.

Rosenshine, Barak, and Furst, Norma F. "Research on Teacher Performance Criteria." In Research in Teacher Education: A Symposium, pp. 37-72. Edited by B. Othanel Smith. Englewood Cliffs, New Jersey: Prentice-Hall, 1971.

Ryans, David G. "Characteristics of Teachers: Their Descriptions, Comparison and Appraisal." Washington, D. C.: American Council on Education, 1960.

Saadeh, Ibrahim, Q., "Teacher Effectiveness of Classroom Efficiency: A New Direction in Evaluation of Teaching." Journal of Teacher Educational 21 (Spring 1970): 73-89.

Sandefur, J. T., and Bressler, Alex A. Classroom Observation Systems in Preparing School Personnel. Washington, D.C.: ERIC Clearinghouse on Teacher Education, 1970).

Sapone, Carmelo V. "Appraisal and Evaluation Systems: Perceptions of Administrators, Teachers." NASSP Bulletin 65 (February 1981): 25-30.

Sapone, Carmelo V. "Appraisal and Evaluation Systems: What Are the Perceptions of Educators, Board Members?" NASSP Bulletin 66 (December 1982): 46-51.

Scandrett, Terrence Glen. "A Comparison of Summative and Formative Teacher Evaluation at a Rural Missouri High School." Ed.D. dissertation, Louis University. 1983.

Schaff, M. E., and Siebring, B. R. "What Do Chemistry Professors Think About Evaluation of Instruction?" Journal of Clinical Education 51 (March 1974): 152-154.

Schwab, Joseph J., and Klinckmann, Evelyn. "Discussion in the Teaching of BSCS Biology." In Biology Teachers' Handbook, pp. 429-448. Edited by Joseph J. Schwab. New York: John Wiley and Sons, 1963.

Searles, William E., and Ng, Raymond W. M. "A Comparison of Teacher and Principal Perceptions of an Outstanding Biology Teacher." Journal of Research in Science Teaching 19 (September 1982): 487-495.

Selden, David. Evaluate Teachers? AFT QUEST Paper No. 4. Washington, D.C.: American Federatoon of Teachers, 1969.

Sergiovanni, Thomas J. "Reforming Teacher Evaluation: Naturalistic Alternatives." Educational Leadership 34 (May 1977): 602-607.

Shannon, Thomas A. "Teacher Evaluation: Some Points to Ponder." In Merit Pay and Evaluation pp. 99-100. Hot Topics Series 1983-84. Bloomington, Indiana: Phi Delta Kappa Center on Evaluation Development and Research, 1983.

Sihota, Sohan S., and Singhania, Ram P. "On Teaching Effectiveness." Journal of Business Education 57 (November 1981): 53-56.

Simon, Alan E. "Peer Supervision: An Alternative." Paper presented at the Annual Conference of The Association for Supervision and Curriculum Development, Detroit, 3 March 1979.

Smith, Fred M., and Gremillion J. B. Teacher Effectiveness. Baton Rouge, La.: Louisana State University, College of Education, 1971.

Spahr, David H. "Evaluating Teacher Effectiveness by Intervisitation." In Selected Articles for Elementary School Principals, pp. 211-215. Washington, D.C.: National Education Association, Dept. of Elementary School Principals, 1968.

Stephens, J. M. The Process of Schooling. New York: Holt, Rinehart, 1967.

Stow, Shirley B. "Using Effectiveness Research in Teacher Evaluation." Educational Leadership 37 (October 1979): 55-56.

Stronck, David R. " A Comparison of Peer and Pupil Evaluations of Lessons Taught by Preservice Biology Teacher." Science Education 60 (January-March 1976): 217-221.

Stulac, Josef, F. II. "Procedure and Results of Measurement in the South Carolina Assessments of Performance in Teachings." Paper presented at the American Educational Research Association Annual Meeting, New York, 19 March 1982.

Sudman, Seymour. Applied Sampling. New York: Academic Press, 1976.

Sullivan, Cheryl Granade. Clinical Supervision: A State of the Art Review. Washington, D.C.: Association for Supervision and Curriculum Development, 1980.

S ummers, Anita, A. and Wolfe, Barbara L. Which School Resources Help Learning: Efficiency and Equity in Philadelphia Public Schools. Philadelphia: Federal Reserve Bank of Philadelphia, 1975.

Symmonds, Percival M. "The Measurement of Teaching Efficiency in High Schools." Educational Administration and Supervision 13 (April 1927): 217-231

Taylor, James Benteley: "Rating Scales as Measures of Clinical Judgement: A Method for Increasing Scale Reliability and Sentivity.". Educational and Psychological Measurement 28 (Autumn 1968): 747-766.

Tirrell, Frederick John. "An Investigation of the Current Perceptions and Ideal Expectations of Senior High School Principals and Teachers Regarding the Principal's Role in Teacher Evaluation." Ed.D. dissertation, Boston College, 1982.

Tobia, Edward Francis. "The Relationship Between Teacher Participation in the Development of a Teacher Evaluation Process and Teacher Attitude Toward Evaluation.: Ed.D. dissertation, Temple University, 1984.

Tollefson, Nona. "Selected Student Variables and Perceived Teacher Effectiveness." Education 94 (September-October 1973): 30-35.

Tomblin, Elizabeth Alene Risinger. "Effects of Participatory and Nonparticipatory Methods of Teacher Evaluation on Selected Teacher Variables." Ph.D. dissertation, University of Colorado at Bouldor, 1976.

Travers. Robert M. W. Second Handbook of Research on Teaching. Chicago: Rand McNally, 1973.

Waimon, Morton D., and Ramsayer, Gary C. "Effects of Video Feedback on the Ability to Evaluate Teaching." Journal of Teacher Education 21 (Spring 1970): 92-95.

Wanat, Stanley F. "Performance Tests: Assessing Teachers of Reading--A Response." The Reading Teacher 25 (April 1972): 623-633.

Washington, Eva. Expert Teacher Action Program. Belmont, Calif.: Fearon Publishers, 1971.

Wiles, Kimball. Supervision for Better Schools. New York: Prentice Hall, 1953.

Willard, David Thomas, "An Assessment of the Effects of a Staff Evaluation Model Developed from Douglas MacGregor's Theory 'Y' Upon the Attitudes of Teacher and Supervisors Toward Evaluation." Ph.D. dissertation, St. Louis University, 1979.

Wolf, Robert L. "How Teachers Feel Toward Evaluation." In School Evaluation: The Politics and the Process, pp. 156-168. Edited by Ernest R. House. Berkeley, Calif.: McCutchan Publishing, 1973.

Woody, Clifford. "Some Observations of Methods of Research in the Appraisal of Teaching.: Educational Administration and Supervision 29 (January, 1943): 1-8.

Wright, Clifford J. and Nutall, Graham. "Relationships Between Teacher Behavior and Pupil Achievement in Three Experimental Elementary Science Lessons." American Educational Research Journal 7 (November 1979): 477-491.

Yevish, Irving A. "The Observation Fallacy." Educational Forum 32 (January 1968): 171-175.

Young, James M., and Heichberger, Robert L. "Teachers' Perceptions of an Effective School Supervision and Evaluation Program.: Education 96 (Fall 1975): 10-19.

Zalby, Leon W. "Student-Faculty Evaluation," Science 183 (March 1974): 1267-1270.

APPENDIX A

Remember: Please send board-approved evaluation policies and your evaluation forms.

A. How many years were you a teacher? _____

B. How many years have you been a principal? _____

C. How many total years have you worked in education? _____

D. What is the highest degree you have attained?

___ BA ___ MA ___ MA+30 ___ MA+45 ___ Doctorate

E. Please estimate the average household income of the community served by your school:

1	2	3	4	5
under $15,000	$15,001-30,000	$30,001-45,000	$45,001-60,000	over $60,000

F. Of the families in the community served by your school, approximately what percentage falls in each of the above income categories:

_____% _____% _____% _____% _____%
1 2 3 4 5

G. What is the size of your teaching staff? (Circle the appropriate number)

5-10 11-15 16-20 21-25 26-30 over 30

H. Have you received in-service training in evaluating teachers?

____ Yes ____ No ____ No, but I would like to

I. Are you usually evaluated by your supervisor?

____ Yes ____ No ____ No, but I would like to be

J. How many formal observations do you usually make when you evaluate a teacher?

0 1 2 3 4 5 more than 5

K. What is the average length of a classroom observation when you evaluate a teacher? (in minutes)

____ minutes or ___ class period(s) of ___ minutes

____ other _____

Please indicate how often each of the following occurs during teacher evaluation in your school (circle the appropriate number):

Key:	1 always	2 usually	3 sometimes	4 seldom	5 rarely	6 never

1. Classroom observation is a significant part of the process of evaluating teachers.

   1 2 3 4 5 6

2. Before a classroom observation, the teacher and supervisor agree on the time and place of the observation.

   1 2 3 4 5 6

3. Before an observation, the teacher and supervisor agree on the criteria for the evaluation.

   1 2 3 4 5 6

4. Before the observation, the teacher and supervisor agree on the instrument to be used in the evaluation.

   1 2 3 4 5 6

5. Before the observation, the supervisor meets with the teacher for a pre-observation conference.

   1 2 3 4 5 6

6. Following the evaluation, the supervisor provides the teacher with a copy of the evaluation report.

   1 2 3 4 5 6

7. Following the observation, the teacher and supervisor hold an evaluation conference.

   1 2 3 4 5 6

8. The teacher is allowed to attach a reply to the evaluator's report.

   1 2 3 4 5 6

9. During the evaluation process the evaluator provides supportive comments based on observed strengths and weaknesses.

   1 2 3 4 5 6

Key:	1	2	3	4	5	6
	always	usually	sometimes	seldom	rarely	never

10. The overall results of the teacher evaluation process lead to specific in-service training designed to improve classroom teaching.

1 2 3 4 5 6

11. The supervisor is knowledgeable in the academic area he/she is expected to evaluate.

1 2 3 4 5 6

12. Teacher evaluation is done by a "team", including one person skilled in the teacher's specialty.

1 2 3 4 5 6

13. Teacher evaluation is based on objective, substantive criteria.

1 2 3 4 5 6

14. Teacher evaluation includes a process for establishing specific goals and objectives for the teacher.

1 2 3 4 5 6

---

15. Which of the following are purposes of teacher evaluation in your school?

___professional growth
___program evaluation
___improvement of instruction
___reward superior performance
___tenure and termination decisions
___teacher accountability
___modification of teacher assignments
___other _____

16. Which of the following form the basis of teacher evaluation in your school?

___ the "clinical cycle" of supervision
___ rating scale of teacher competencies
___ specific teacher objectives
___ rating scale of professional and personal characteristics
___ narrative description of teacher performance
___ other _____

17. Teachers in your school participate in which of the following (if any)?

___ designing the teacher evaluation form
___ establishing the goals and purposes of evaluation
___ developing the policies which govern teacher evaluation
___ designing the process of teacher evaluation

18. How would you characterize the evaluation process in your school?

objective	1	2	3	4	5	6	subjective
summative	1	2	3	4	5	6	formative
diagnostic	1	2	3	4	5	6	descriptive
cooperative effort	1	2	3	4	5	6	administrative task
teacher as subordinate	1	2	3	4	5	6	teacher as colleague
on-going	1	2	3	4	5	6	specific time frame
punitive	1	2	3	4	5	6	non-punitive
mutual trust	1	2	3	4	5	6	suspicion
teacher as employee	1	2	3	4	5	6	teacher as person
threatening	1	2	3	4	5	6	non-threatening
satisfactory	1	2	3	4	5	6	unsatisfactory

Please give your professional opinion regarding the relative importance unimportance of each of the following to teacher evaluation:

Key:	1	2	3	4	5	6
	very important	moderately important	mildly important	mildly unimportant	moderately unimportant	very unimporta

19. Classroom observation is a significant part of the process of evaluating teachers.

1 2 3 4 5 6

20. Before a classroom observation, the teacher and supervisor agree on the time and place of the observation.

1 2 3 4 5 6

21. Before an observation, the teacher and supervisor agree on the criteria for the evaluation.

1 2 3 4 5 6

22. Before the observation, the teacher and supervisor agree on the instrument to be used in the evaluation.

1 2 3 4 5 6

23. Before the observation, the supervisor meets with the teacher for a pre-observation conference.

1 2 3 4 5 6

Key:	1	2	3	4	5	6
	very important	moderately important	mildly important	mildly unimportant	moderately unimportant	very unimportant

24. Following the evaluation, the supervisor provides the teacher with a copy of the evaluation report.

1 2 3 4 5 6

25. Following the observation, the teacher and supervisor hold an evaluation conference.

1 2 3 4 5 6

26. The teacher is allowed to attach a reply to the evaluator's report.

1 2 3 4 5 6

27. During the evaluation process, the evaluator provides supportive comments based on observed strengths and weaknesses.

1 2 3 4 5 6

28. The overall results of the teacher evaluation process lead to specific in-service training designed to improve classroom teaching.

1 2 3 4 5 6

29. The supervisor is knowledgeable in the academic area he/she is expected to evaluate.

1 2 3 4 5 6

30. Teacher evaluation is done by a "team", including one person skilled in the teacher's specialty.

1 2 3 4 5 6

31. Teacher evaluation is based on objective, substantive criteria.

1 2 3 4 5 6

32. Teacher evaluation includes a process for establishing specific goals and objectives for the teacher.

1 2 3 4 5 6

33. In your professional opinion, which of the following should be purposes of teacher evaluation?

___program evaluation
___professional growth
___improvement of instruction
___teacher accountability
___tenure and termination decisions
___reward superior performance
___modification of teacher assignment
___other_____

34. In your professional opinion, which of the following should form the basis of teacher evaluation?

___ narrative description of teacher performance
___ rating scale of teacher competencies
___ specific teacher objectives
___ rating scale of professional and personal characteristics
___ the "clinical cycle" of supervision
___ other _____

35. In your professional opinion, which of the following (if any) should teachers participate in?

___ developing the policies which govern teacher evaluation
___ designing the teacher evaluation form
___ designing the process of teacher evaluation
___ establishing the goals and purposes of evaluation

36. In your professional opinion, how would you ideally characterize the process of teacher evaluation?

subjective	1	2	3	4	5	6	objective
summative	1	2	3	4	5	6	formative
descriptive	1	2	3	4	5	6	diagnostic
cooperative effort	1	2	3	4	5	6	administrative task
teacher as colleague	1	2	3	4	5	6	teacher as subordinat
specific time frame	1	2	3	4	5	6	on-going
non-punitive	1	2	3	4	5	6	punitive
mutual trust	1	2	3	4	5	6	suspicion
teacher as employee	1	2	3	4	5	6	teacher as person
non-threatening	1	2	3	4	5	6	threatening
satisfactory	1	2	3	4	5	6	unsatisfactory

37. In your professional opinion, how many formal observations is ideal when evaluating a teacher?

0 1 2 3 4 5 more than 5

38. In your professional opinion, what is the ideal length of a classroom observation?

___ minutes or ___ one full class period

___ other _____

APPENDIX B

INTERVIEW QUESTIONS

1. Are your teachers satisfied with the present evaluation system?

2. Do your teachers express any specific displeasures with your evaluation system?

3. Are you satisfied with your evaluation system?

4. If you could change any parts of your evaluation system, what would you change?

5. Are you constrained by the superintendent, by the district, or by the teacher's contract in any way in the manner in which you like to evaluate teachers?

APPROVAL SHEET

The dissertation submitted by Kenneth E. Upshaw has been read and approved by the following committee:

Dr. M. P. Heller, Professor, and Director of Administrative Services, Department of Educational Administration and Supervision, Loyola University of Chicago.

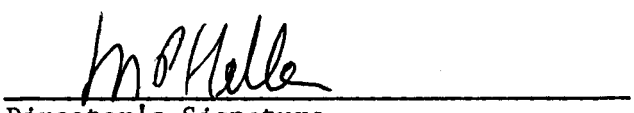
Dr. Martha Wynne, Assistant Professor, Department of Foundations of Education, Loyola University of Chicago.

Dr. Phillip Carlin, Associate Professor, and Chairman of the Department of Educational Administration and Supervision, Loyola University of Chicago.

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the dissertation is now given final approval by the Committee with reference to content and form.

The dissertation is therefore accepted in partial fulfillment of the requirements for the degree of Doctor of Education.

December 16, 1985
Date


Director's Signature