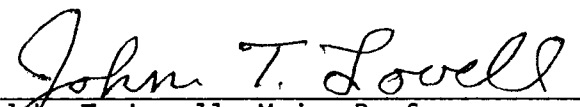
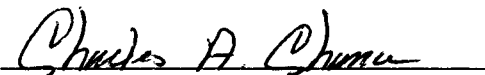
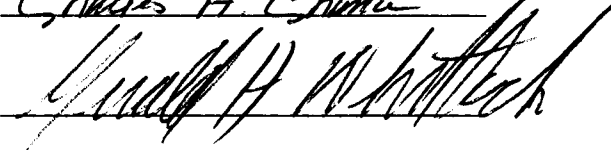
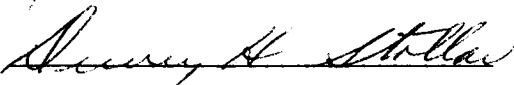


To the Graduate Council:

I am submitting herewith a dissertation written by Samuel Lawson Houston entitled "An Analysis of Perceptions of Tennessee Public School Teachers, Building-Level Administrators and Central Office Administrators Toward Teacher Evaluation." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Education, with a major in Educational Administration and Supervision.


John T. Lovell, Major Professor

We have read this dissertation
and recommend its acceptance:

Accepted for the Council:


Vice Chancellor
Graduate Studies and Research

AN ANALYSIS OF PERCEPTIONS OF TENNESSEE PUBLIC SCHOOL TEACHERS,
BUILDING-LEVEL ADMINISTRATORS AND CENTRAL OFFICE
ADMINISTRATORS TOWARD TEACHER EVALUATION

A Dissertation
Presented for the
Doctor of Education
Degree
The University of Tennessee, Knoxville

Samuel Lawson Houston
March 1981

3049553

Copyright by Samuel Lawson Houston 1981
All Rights Reserved

ACKNOWLEDGMENTS

The writer wishes to express his gratitude to Dr. John T. Lovell, his major professor, for his support, guidance, and friendship; to Dr. Dewey Stollar, Dr. Charles Chance, and Dr. Gerald Whitlock for their counsel and valuable suggestions.

The writer is especially grateful to his wife, Jane, and his two sons, Sam and Matt, without whose patience, understanding, and sacrifice this accomplishment would not have been possible.

ABSTRACT

This study had as its purpose the comparison and analysis of the descriptive data collected by Patricia Hayes Miller (University of Tennessee, June, 1980) concerning the perceptions of 314 central office and 261 building-level administrators toward teacher evaluation, and by Russell G. Ramsay (University of Tennessee, August, 1980) concerning the perceptions of 243 teachers toward the facets of teacher evaluation explored by the same questionnaire used by Patricia Miller. Responses by randomly drawn members of each of the three segments mentioned revealed their perceptions of the purposes of teacher evaluation, methods of evaluation, and evaluation priorities.

The major findings of the study were:

1. The purposes for which teacher evaluation was being conducted in Tennessee were perceived in a distinctly different manner by administrators and teachers.
2. Significant differences existed among teachers, building-level administrators and central office administrators as to the desirability of the various methods of teacher evaluation explored by the "Teacher Evaluation Questionnaire."
3. There was a significant lack of agreement among teachers, building-level administrators and central office administrators regarding:

- a. the number of hours spent per teacher in the evaluation process,
 - b. the number of hours per teacher which should be spent in the evaluation process,
 - c. the number of observations performed per teacher and the number of observations which should be performed during an evaluation year, and
 - d. the present average length of an observation and the optimum length of an observation.
4. A significant discrepancy existed between the perceptions of teachers and administrators as to the degree of teacher involvement in the development and implementation of the evaluation process.
5. Teachers perceived their overall evaluation process to be significantly less acceptable than did administrators.
6. Evidence pointed to a more positive perception of the overall evaluation process by teachers as a function of their participation in the development of the evaluation procedures and instruments.
7. Evidence indicated a significant positive correlation between teachers' perceptions of their overall evaluation process and the extent of their participation in the implementation of the process.
8. At this point in time, the method utilized in the evaluation did not influence Tennessee teachers' perceptions of the purposes for teacher evaluation.

9. Evidence pointed to a significant positive correlation between perceptions of the overall evaluation process and the increased incidence and length of classroom observations.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
Statement of the Problem	2
Purpose of the Study	2
Questions to be Answered	3
Importance of the Study.	5
Assumptions.	7
Delimitations of the Study	7
Definition of Terms.	7
Procedure.	10
Measurement.	11
Analysis of the Data	13
Organization of the Study.	15
II. A REVIEW OF RELATED LITERATURE AND RESEARCH.	17
History.	17
Purposes	22
Criteria	26
Methods of Evaluation.	31
Perceptual Factors Affecting Evaluation.	43
III. PRESENTATION AND ANALYSIS OF THE DATA.	50
Introduction	50
Purposes for Teacher Evaluation.	50
Methods of Evaluation.	54

CHAPTER	PAGE
IV. SUMMARY, FINDINGS, CONCLUSIONS AND RECOMMENDATIONS	81
Introduction	81
Summary.	81
Findings	86
Conclusions.	103
Recommendations.	104
LIST OF REFERENCES.	106
APPENDIX.	116
VITA.	127

LIST OF TABLES

TABLE	PAGE
I. Frequencies of Ranks for School and Personal Purposes of Teacher Evaluation as Perceived by Tennessee Teachers (T), Building-Level Administrators (BL) and Central Office Administrators (CO) During the 1978-79 School Year	52
II. Mean Responses of Central Office Administrators (CO), Building-Level Administrators (BL) and Teachers (T) in Tennessee Concerning the Desirability of Various Methods of Evaluation as Measured in Section III of the Teacher Evaluation Questionnaire in Use During the 1978-79 School Year.	56
III. Mean Responses of Teachers (T), Building-Level Administrators (BL) and Central Office Administrators (CO) in Tennessee Concerning the Time Spent in Teacher Evaluation and Observations Accompanying an Evaluation as Perceived During the 1978-79 School Year	63
IV. Proportions of Teachers (T), Building-Level Administrators (BL) and Central Office Administrators (CO) in Tennessee Responding Affirmatively to the First Seven Items in Part IV-G of the Teacher Evaluation Questionnaire During the 1978-79 School Year.	66
V. Analysis of Variance of the Overall Ratings Given to the Teacher Evaluation Process by Teachers, Central Office Administrators, and Building-Level Administrators as Implemented During the 1978-79 School Year.	70
VI. Tennessee Teachers' Rankings of Purposes of Teacher Evaluation According to the Method of Evaluation Used During the 1978-79 School Year	74
VII. Building-Level Administrators' Rankings of Purposes for Teacher Evaluation According to the Method of Evaluation Utilized During the 1978-79 School Year	76

TABLE	PAGE
VIII. Central Office Administrators' Rankings of Purposes of Teacher Evaluation According to the Method of Evaluation Utilized During the 1978-79 School Year	78
IX-A. Analysis of Variance Data Supporting Table II, Page 56.	121
X-A. Analysis of Variance Data Supporting Table III, Page 63.	125
XI-A. Significant Z Scores Supporting Table IV, Page 66.	126

LIST OF FIGURES

FIGURE	PAGE
I. Graphic Representation of the Regression Equation Developed for Tennessee Teachers' Overall Ratings of Their Evaluation Procedures as a Function of Their Participation Scores as Assessed During the 1978-79 School Year	73

CHAPTER I

INTRODUCTION

Performance evaluation has significant implications for the achievement of organizational goals. It should provide an information base from which personnel decisions can be made. From the standpoint of the educational organization, performance evaluation should also facilitate the improvement of instruction.

Instructional improvement occurs most often through the analysis of instructional performance within the context of a collegial relationship between the teacher and the instructional supervisor. A prime factor in the development of this type of relationship is a mutual understanding of the purposes of the process.

Vroom's expectancy theory of motivation has implications for the evaluation process and its relation to instructional improvement. Participants must perceive that the evaluation process will lead to the improvement of instruction. Moreover, the teacher must discern sufficient personal utilities or satisfactions coming from the evaluation process to motivate him/her to participate actively in it.¹ If teacher evaluation is to serve as a conduit for instructional improvement in Tennessee, differences in the perceptions of administrators and teachers must be assessed and factors which account for these differences must be identified.

I. STATEMENT OF THE PROBLEM

In 1974, the Tennessee State Board of Education mandated the development and implementation of formal performance evaluation procedures for professional personnel in all school districts within the state.² Although the various evaluation plans have been on file with the State Department since that time, no data base has been established through which evaluation practices can be assessed and improved.

II. PURPOSE OF THE STUDY

In a study recently completed by Patricia Miller, data were collected which described the perceptions of building-level and central office administrators toward methods and purposes of teacher evaluation and the role of the administrator in the evaluation process. The extent of participation in the development and implementation of the process was also assessed. A parallel study by Russell Ramsay described the perceptions of teachers toward these factors.³

The purpose of this study was to identify and analyze differences in perceptions of administrators and teachers as described in the Miller and Ramsay studies and to examine the relationships between the utilization of various methods of evaluation, extent of participation in development and implementation, purposes of performance evaluation, and perceived adequacy of the system.

III. QUESTIONS TO BE ANSWERED

1. Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding:
 - a. the purposes of teacher evaluation?
 - b. the desirability of the teacher checklist, record of critical incidents, parental input, self-evaluation, evaluation by objectives, competency tests for teachers, student test data, classroom observation, and instructional conferences as methods of teacher evaluation?
 - c. the number of hours spent per teacher in the evaluation process?
 - d. the number of hours which should be spent in the teacher evaluation process?
 - e. the number of observations performed per teacher during an evaluation year?
 - f. the average length of an observation?
 - g. the number of times a teacher should be observed during an evaluation year?
 - h. the optimum average length of an observation?
 - i. the opportunity afforded the teacher to see and review his/her evaluation documents?
 - j. the right of the teacher to make written comments or disagreements concerning his/her evaluations?

- k. whether the teacher receives a copy of the final evaluation that is placed in the personnel file?
 - l. whether teachers are informed of the evaluation process?
 - m. the involvement of teachers in the development of evaluation instruments?
 - n. the involvement of teachers in the development of evaluation procedures?
 - o. whether teachers are informed of the evaluation appeal process?
 - p. the ratings given to the overall evaluation processes?
2. Are there significant differences in the ratings of the overall evaluation procedures as a function of the teachers' involvement in the development of the procedure and/or their involvement in the development of evaluation instruments?
 3. Are there significant differences in the ratings of the overall evaluation procedures as a function of the extent of the teachers' participation in the implementation of the procedure?
 4. Within the three groups--teachers, building-level administrators, and central office administrators--are there significant differences in the perceived school purposes of evaluation as a function of the methods utilized in the procedure?

5. Are there significant differences in the teachers' ratings of the overall evaluation process as a function of:
- a. the number of hours perceived to have been spent per teacher in the evaluation process?
 - b. the number of times a teacher was observed during an evaluation period?
 - c. the perceived average length of the observations?

IV. IMPORTANCE OF THE STUDY

Beginning with the movement for documentation of the effectiveness of federally funded programs in the 1960's educators have been grappling with demands for accountability from federal and state governments, communities, advocacy groups, and taxpayers at large. In 1974, the Tennessee State Board of Education reacted to pressure for accountability by mandating the development of formal performance evaluation procedures for professional personnel in all school districts. The intent of the Board was evident in the following excerpt from the Rules, Regulations and Standards:

Local boards of education shall develop evaluative procedures for all professional school personnel during the 1973-74 school term. The evaluative procedures shall be designed for the purpose of improving the instructional program. The Evaluative Criteria shall be filed with the Commissioner of Education by June 1, 1974.⁴

A further measure of instructional performance was instituted by the State Board during the 1977-78 school year in the form of competency testing for students. Ostensibly, the purpose of the tests

was to assess the skills of students. Since the attainment of skills may reflect the quality of instruction, these test results may also serve as an assessment of the quality of instruction in Tennessee schools. In essence, because the quality of instruction is a function of teacher performance, the Board has addressed another form of teacher evaluation.

In the Fall of 1979, a series of legislative hearings were conducted by a subcommittee of the Education Committee of the Tennessee House of Representatives for the purpose of investigating means of improving instruction in public schools. Among the factors under investigation were teacher and administrator evaluation. The tenor of the hearings was that performance evaluation must become more effective in improving teacher performance and, in turn, improving instruction.⁵

Recent data which describe present performance evaluation practices in Tennessee have been provided to the State Department of Education through the Miller and Ramsay studies. A description in and of itself does not afford the State Department with a comprehensive assessment of the extent to which the various procedures are perceived to fulfill the organizational purposes for which they were designed. Only after significant differences in the perceptions of administrators and teachers toward the purposes and methodology of teacher evaluation have been identified and relationships have been explored can these data serve to make teacher evaluation a viable tool for instructional improvement in Tennessee.

V. ASSUMPTIONS

1. The respondents in the Miller and Ramsay studies responded candidly regarding their perceptions of the evaluation process.
2. The respondents depicted the evaluation methods presently in use and the extent of their participation in the development and implementation of the process candidly.
3. The respondents had certain perceptions of the adequacy of the evaluation process which served as valid and useful data.

VI. DELIMITATIONS OF THE STUDY

1. The perceptions of administrators and teachers toward teacher evaluation were delimited to the data provided by the respondents in the Miller and Ramsay studies.
2. Analysis of relationships will involve only those participants who responded "yes" or "no" on the items concerning participation in and overall adequacy of the evaluation process.
3. Respondents' perceptions represent reactions to a variety of teacher evaluation practices over the state.

VII. DEFINITION OF TERMS

The following terms are defined to eliminate any confusion:

Adequacy. Perceptions of adequacy are a function of the overall evaluation process given in response to the final item of the questionnaire.

Building-level administrator. The person(s) at the local school responsible for teacher evaluation; this position may include the principal, assistant principal, or curriculum principal/coordinator.

Diagnostic evaluation. The assessment of teacher personality and attitudes via biographical questionnaires, interviews, attitude and preference inventories, and checklists throughout the school year.

Evaluation by objectives. The identification of goals and activities by teachers and administrators for the purpose of professional growth and improvement.

Formative evaluation. The assessment of observable and discrete behaviors through observation coding systems and inference rating scales throughout the school year.

Job targets. Goals identified by teachers and/or administrators at the beginning or any time during the school year; these may or may not be based on a needs assessment.

Observation. Collection of data by an observer relative to the teacher's evaluation during a class session.

Observation instrument. A form used by an observer to make an objective and/or evaluative judgment of teaching performance, classroom environment, student interaction, and/or methods of instruction.

Observer. The teacher or administrator identified to collect data on teacher and/or student behaviors that occur in the classroom and other educational settings.

Peer observation. Classroom observations of a teacher by another teacher.

Personnel officer. The system-wide person responsible for processing personnel decisions including recruitment, selection, tenure, transfers, and dismissals.

Pre-observation conference. Discussions between the teacher and observer to outline the details of the classroom observation prior to visitation.

Post-observation conference. Discussions between the teacher and observer after the classroom observation has occurred to report and analyze data and provide feedback to the teacher.

Self-observation. Observation of a teacher by a student in the class and/or school during a class session.

Summative evaluation. Cumulative ratings of teacher performance and measures of pupil gain collected during the school year. This evaluation is conducted at the end of the school year.

Supervisor. The system-wide person responsible for instructional improvement, curriculum development, and teacher evaluation; this position may include the assistant superintendent, supervisor, director, consultant, specialist, or coordinator as identified by the school system.

Teacher evaluation. The assessment of teachers conducted for purposes of improvement of instruction, staff development, tenure, merit pay, transfer, promotion, or dismissal.

Teacher evaluation system. The step-by-step process involved in the assessment of teachers; other synonymous terms include teacher evaluation system and teacher evaluation framework.

VIII. PROCEDURE

This study was a sequel to the descriptive studies conducted by Patricia Miller, assessing the perceptions of building-level and central office administrators toward teacher evaluation, and Russell Ramsay, who conducted the same assessment for teachers. From a population of 1,826 Tennessee building-level administrators, a random sample of 319 was selected to take part in the Miller study. Usable returns were received from 261 principals, assistant principals, curriculum principals, and others. The central office administrator population consisted of some 148 Tennessee superintendents, 627 supervisors, assistant superintendents, coordinators and others who had responsibility for teacher evaluation. The total population of superintendents was surveyed, while responses were solicited from a random sample of 240 members of the supervisory group. Usable returns were received from 314 central office administrators.

In the Ramsay study, a random sample of 380 Tennessee teachers was drawn from a population of 40,565. Of the teachers surveyed, 243 yielded usable returns.

The data provided by these two studies were synthesized to determine the existence of significant differences of the types expressed in the "Questions to be Answered" section. Further analysis was conducted to determine the existence of relationships between the overall assessment of evaluation procedures by administrators and teachers and participation in the development and implementation of these procedures.

IX. MEASUREMENT

The questionnaire used in the Miller and Ramsay studies was developed through an extensive review of current literature in the area of teacher evaluation as well as a review of the evaluation systems submitted to the Tennessee Commissioner of Education as required in the 1974 regulation.⁶ The instrument was reviewed by a panel of experts which included prominent scholars and practitioners. Members of the panel included Dr. Ben Harris, University of Texas; Dr. Morris Cogan, University of Pittsburgh; Dr. Dale Bolton, University of Washington; Dr. George Redfern, Educational Management Consultant; Dr. Madeline Hunter, University of California, Los Angeles; Dr. Harold Drummond, University of Chicago; Dr. Benjamin Bloom, University of Chicago; Dr. Robert Smallridge, Superintendent, Oak Ridge Schools; Dr. Ann Roney, Supervisor, Knoxville City Schools; A. D. Hancock, Principal, Metro-Nashville Schools; Chester LaFever, Principal, Metro-Nashville Schools; and Dr. Marshall Perritt, Deputy Superintendent, Memphis City Schools.

Data were collected under three major headings (see Appendix). Within the "School System Information" section, respondents were asked to choose from a list of 12 purposes for teacher evaluation the 3 which were perceived to be the most important within their particular school system and to rank the 3 choices from 1 to 3, where 1 is the most important. They were also asked to choose and rank the three purposes which were considered personally to be most important. Further questions in this section elicited a ranking of the

various functions of the principal, perceptions of those presently involved in the evaluation process, the person with primary responsibility for evaluation, the evaluation process, and the person who should have primary responsibility for teacher evaluation.

Through the items within the "Methods of Evaluation" section, respondents checked the methods of evaluation presently in use. Methods included in this section were: teacher checklists, records of critical events, parental input, self-evaluation, evaluation by objectives, competency tests for teachers, student test data, classroom observation, observation-related conferences, and instructional conferences. Respondents were also asked to rate each item in the section regarding its desirability as a means of teacher evaluation. Items were rated on a five-point scale ranging from very desirable (one) to very undesirable (four). A rating of five indicated insufficient information.

In the final section, entitled "Evaluation Priorities," respondents indicated their perceptions of the number of hours spent yearly with each teacher in the evaluation process and the number of hours which should be spent with each teacher yearly in this process. Perceptions of the actual and ideal number of classroom observations and the actual and optimum amounts of time spent in this activity were elicited. Following this subsection were a series of seven questions exploring the opportunity of teachers to see and review their evaluation documents, the right of teachers to respond in writing to evaluations, whether the teacher received copies of the evaluation documents, whether the teacher is informed of the evaluation and/or appeal processes, and

whether the teacher is involved in the development of the evaluation procedure and/or the evaluation instruments. Participants in the two studies responded either "yes" or "no" to each of these questions.

A final question required the respondent to rate the overall evaluation process in his/her school system as either good, fair, or poor.

X. ANALYSIS OF THE DATA

Differences in the perceptions of administrators and teachers concerning the school and personal purposes of evaluation were analyzed through the Spearman Rank Order Correlation procedure. This procedure was also used to assess significant differences in the perceived purposes of evaluation as a function of the evaluation methods utilized. The Spearman Correlation Coefficient gave strong evidence of a lack of agreement between the rankings of the administrator and teacher groups in question.

The analysis of variance procedure was used to compare the responses of teachers, building-level administrators and central office administrators concerning the desirability of various methods of teacher evaluation, time spent in the evaluation process, the role of classroom observation in evaluation, teachers' opportunity to see and review evaluation documents, and the ratings given to the overall evaluation process (questions 1-b, 1-c, and 1-p). A significance level of .05 was maintained, and Duncan's Multiple Range Test was utilized to determine the location of significant differences. The test for significant

differences between proportions was employed to examine the responses of the three focal groups concerning the teacher's right to make written comments on disagreements concerning the evaluation, the receiving of a copy of the final evaluation by the teacher, whether teachers are informed of the evaluation process, the involvement of teachers in the development of evaluation procedures and instruments, and whether teachers were informed of the evaluation appeal process (questions 1-j and 1-o). This test was deemed appropriate because the responses were limited to "yes," "no," and "I don't know" and, therefore, must be treated as nominal level data. A significance level of .05 was utilized.

The "students t" test was used to analyze teachers' ratings of their overall evaluation procedures as a function of their perceived involvement in the development of evaluation procedures and evaluation instruments (question 2). Differences in ratings were considered to be significant at the .05 level.

In order to determine the extent of the teachers' participation in the implementation of the evaluation process (question 3), a summed score for participation was derived by combining the responses to the following questions:

1. Do teachers have an opportunity to see and review their evaluation documents?
2. Do teachers have the right to make written comments of disagreements concerning their evaluation?
3. Do teachers receive a copy of the final evaluation that is placed in their personnel file?

4. Are teachers informed of the evaluation process?
5. Are teachers informed of the evaluation appeal process?

For the first question, a response of "yes, always" or "yes, if they wish" was given a score of one and responses of "not usually" or "no, confidential" were scored as zero. The resulting summed scores ranged from zero, indicating no participation, to five, reflecting a high degree of participation. Regression analysis was then performed using the teachers' ratings of the overall evaluation process as the dependent variable and the summed participation score as the independent variable. Pursuant to the development of a regression equation, an analysis of variance was performed and the significance of the slope and intercept of the least squares line was determined through the use of t tests. In conjunction with this procedure, a Pearson Product Moment Correlation was calculated to further assess the relationship between the variables.

Regression analysis was also used to gather evidence concerning teachers' ratings of the overall evaluation process as a function of time spent in the evaluation process and as a function of the number and length of observations accompanying an evaluation (question 5). As in the analysis of the previous question, a Pearson Correlation Coefficient was calculated.

XI. ORGANIZATION OF THE STUDY

The study was divided into four chapters as follows:

Chapter I contains the general introduction, statement of the problem, purpose of the study, questions to be answered, importance of the study, assumptions, limitations, and organization of the study.

Chapter II contains a review of related literature in the area of teacher evaluation for personnel decisions.

Chapter III is the presentation and analysis of the data.

Chapter IV contains the summary, findings, conclusions, and any recommendations based on the research conducted.

CHAPTER II

A REVIEW OF RELATED LITERATURE AND RESEARCH

This chapter is an analysis and synthesis of the literature related to teacher evaluation. Data for this chapter were obtained from numerous research studies, books, pamphlets, theses and periodic literature. The chapter was organized to include a historical overview of the development of teacher evaluation, research and theory concerning purposes and criteria for evaluation, methods of teacher evaluation and perceptual factors influencing the evaluation process.

I. HISTORY

"All teachers are evaluated. Regardless of how formal the evaluation system is, what evidence is collected and analyzed or how often formal reports are written, teachers are evaluated and evaluated often."⁷ In this brief statement, Bolton has articulated the inevitability of teacher evaluation in the educational organization.

Popham believed that, historically, evaluation has been viewed quite properly as an integral activity of a rational approach to life.⁸ During the formative years of public education in the United States, teacher evaluation was performed by members of the lay community. In 1642, the governor and Company of Massachusetts Bay mandated that town officers spend a portion of their time working with parents, school masters, and their children, giving attention to "the calling and

employment of the children of the community, especially with respect to their ability to read, to develop religiosity, and to obey the laws of the country."⁹ Whithall and Wood ascribed a connotation of inspection to visits made by community leaders to teachers' classrooms during this period.¹⁰

In 1709, The Commission of the City of Boston delegated the responsibility for scrutinizing the methods of teaching and their resultant proficiency to a committee of citizens. The committee was instructed to make frequent, direct inspections and to give information about the methods of teaching and the proficiency of learning.¹¹

The practice of lay teacher evaluation continued into the early 1800's; however, the tone of the citizens' visits had changed from inspection to stimulation of the teachers' desire to improve teaching practice. As more emphasis was placed on the supportive aspects of supervision, the lay evaluation deferred to the professional school supervisors for judgments of teacher effectiveness.¹² Davis marked the origin of formal evaluation of professionals in school practice of the late 19th Century, as well as in the efficiency movement of the early 20th Century. The focus on efficiency during this period was credited to Frederick Taylor's scientific management movement.¹³ Smith attributed the use of industrial techniques during the era of scientific management to the growing complexity of the educational organization.

As the functions of the school expanded and school budgets mounted, it was natural to apply to the operations of the school the procedures for upgrading efficiency that had been successful in industry.¹⁴

One of the earliest pieces of research in teacher evaluation was conducted by J. L. Merriam in 1905. Since past academic performance was a major consideration in the selection of teachers, Merriam's study attempted to establish a relationship between this criterion and teacher performance. His findings indicated a negligible relationship between scholarship and teaching ability. Even practice teaching performance was found to be only "slightly prophetic" of future effectiveness.¹⁵

The report, "A Tentative Scheme of the Measurement of Teaching Efficiency," presented by Edward C. Elliott at a 1910 convention of city superintendents in Washington, D. C. was indicative of early administrative interest in obtaining a practical method of measuring teaching efficiency. Elliott established seven categories of teacher efficiency including physical efficiency, projected efficiency, dynamic efficiency and achieved efficiency. To each of these categories Elliott awarded points so that the sum of all 7 maximum point values was 100. This first attempt at quantifying teacher effectiveness proved inconclusive.¹⁶

In 1912, the measurement movement was recognized by the organized teaching profession in the first report to the National Council of Education of its Committee on Standards of Tests for Measuring the Efficiency of Schools or School Systems. The committee was chaired by George D. Strayer.¹⁷ Studies of Littler, Buellfield, and others illustrated the early emphasis on teacher failure rather than improvement of performance. Deficiencies in discipline, knowledge of subject matter,

effort, and intelligence, among others, were identified as common causes of teacher failure.¹⁸ Beecher believed that the low level of training, poor wages, and the inadequacy or total absence of certification standards precipitated this rather negative focus.¹⁹

Interest in teacher evaluation during the 1920's was of a more positive nature. In his study of qualities related to success in teaching, Knight attempted to isolate significant qualities of effective teaching and to determine methods of measuring these qualities. Knight found little relationship between scholarship and general teaching ability; however, professional tests could serve only as a "partial indication of teaching success," as could in-service professional study. Intelligence was deemed significant for high school teachers. A further contribution of Knight was to call attention to dangers in the use of rating scales, in particular the possibility of a "halo" effect.²⁰

In 1929, A. S. Barr released his study of "Characteristic Differences of Good and Poor Teachers." Good and poor teachers were identified according to the level of consistency of superintendents' ratings with those of state inspectors. Although Barr characterized evaluation practices as unreliable and "of doubtful validity," his survey of desirable teacher characteristics included items such as ability to stimulate interest, effective organization of subject matter and provision for individual differences, all of which continue to be addressed by current evaluation instruments.²¹

Assessment of the reliability and validity of pupil ratings of teachers was the purpose of Bryan's study. Bryan used the

"chance-halves" method of correlation to assess reliability and assumed a level of validity from the resulting coefficients, since he was only trying to establish the validity of pupil ratings of teachers as a measure of pupil opinion.²² The high degree of reliability achieved in Bryan's study led Beecher to conclude that pupil ratings discriminated among teachers more than ratings by administrators.²³

During the 1940's, research by Baxter,²⁴ Reaves and Cooper,²⁵ and others focused on improvement of rating scales. Lewis' "Traditional Approach," still the predominant practice, evolved from these beginnings. The traditional approach is defined by Lewis as "the semi-annual ritual of writing narrative reports and/or checklist evaluations on teachers."²⁶

"Progressive education" came under fire during the 1950's. Critics of the period attributed the "intellectually inferior" state of public school instruction to such soft-line concepts. The launching of the Sputnik in 1957 heightened public criticism and resulted in a number of government-instituted curriculum projects, particularly in the sciences.²⁷

Federal legislation of the 1960's provided funding for various innovative social and educational programs. On the heels of these funds came demands for documentation of program effectiveness. Because many of these national programs dealt directly with the schools, the accountability demanded of them automatically raised questions about the teachers who played a prominent role in their implementation.²⁸

Public demands for accountability in the 1970's manifested in definitive teacher evaluation legislation such as California's Stull Act,

enacted in 1971, and New York's "Fair Dismissal" Act of 1972.²⁹

Growing pressure for accountability has also created an intensive search by school districts for improved ways to evaluate teachers.

One such effort was the development of an evaluation instrument known as TAI (Teacher Appraisal Instrument) described by Hunter in her research. Using this instrument, teachers are observed and are evaluated on criteria which focus on the following five questions:

1. Is there an instructional objective?
2. Is the objective appropriate?
3. Was the objective achieved?
4. What was done to facilitate learning?
5. What was done that interfered with learning?³⁰

Summary

Historically, the quest for effective evaluation of teaching performance can be traced to the inception of formal education in this country. Trends in philosophy, methodology, and purpose have been influenced by economics, politics, education, business, and industry. Unfortunately the interests of laymen, legislators, and professional educators have not produced a universally accepted set of guidelines for this critical function.

II. PURPOSES

In his discussion of "performance control systems," Mintzberg pointed out two organizational purposes which such systems could serve.

As a measurement tool the performance control system would be used simply to signal a performance deficiency. Mintzberg also recognized that performance evaluation could serve as a motivation to elicit higher levels of performance.³¹ Levin articulated two similar purposes related to the educational organization.

Teacher evaluation can serve one or both of two purposes. The first is to guide decisions about hiring, retention, and promotion. The second is to help improve teaching.³²

Within these broad classifications a diversity of purposes has been expressed. Sergiovanni viewed the improvement of instruction as the sole purpose of teacher evaluation.³³ Frank perceived the foremost goal of staff evaluation to be to provide individuals with constructive feedback.³⁴ The National School Public Relations Association concurred with Frank's opinion and linked the instructional improvement to the quality of feedback provided.³⁵

Styles and Camp linked personal development with instructional improvement in the following statement:

The objective of evaluation should be teacher growth, in scope and depth, and enhancement of one's latitude and special abilities in a fashion which contributes to structural function and educational and personal development.³⁶

In the same vein were Redfern's three purposes:

1. assessment of the status and quality of teaching performance,
2. identification of those aspects of performance which are below standard and need improvement, and
3. stimulation of the growth and development of the individual.³⁷

Castetter believed that the primary purpose of teacher evaluation was to facilitate change in individual behavior in order to achieve personal and organizational goals. He further stated that evaluation should support decisions on salary and wage increases, transfers, dismissals, and should communicate information to appraisees through a review of their progress.³⁸ Castetter's more administrative posture is similar to that of Cascio, who saw evaluation as:

1. a basis for personnel decisions,
2. a criterion in personnel research,
3. a predictor of future performance,
4. an aid in establishing training objectives,
5. a personnel development tool, and
6. a means of providing concrete feedback to employees.³⁹

In descending order of frequency, Porter, Lawler and Hackman reported the purposes of performance evaluation as business, counseling, promotion, training and development considering retention or discharge and salary administration or merit increase.⁴⁰

In outlining the purposes served by evaluation, Bolton assumed a more integrated stance by including statements related to instructional improvement and administrative decision making. He also felt that evaluation served as a basis for the teacher's career planning and could facilitate self-evaluation.⁴¹ The purposes revealed by a 1977 survey of 375 school districts would appear to corroborate Bolton's observations. In descending order of frequency the school districts identified as purposes of teacher evaluation: improvement of teaching

performance, appointment decisions for probationary teachers, tenure recommendations, dismissal of unsatisfactory tenure teachers, promotion, qualifying for salary increment decisions regarding reductions-in-force, and other pay-related decisions.⁴²

Perceptions of Purposes for Evaluation

Howsam suggested that the perceptions of teachers concerning the purpose of evaluation played a large part in determining their reactions toward the total evaluation process. To the extent that teacher evaluation is perceived to be for self-realization purposes, there is likely to be a relative absence of conflict.⁴³ Zelenak and Snider concluded that teachers who felt that evaluation was for instructional purposes were supportive of the process while those who felt that evaluation was utilized for administrative purposes (teacher tenure, promotion, dismissal, assignment, salary, and permanent record files) tended to regard the teacher evaluation process negatively.⁴⁴

McNally advised that great care should be taken to guard against teacher evaluation being perceived as a distinctly administrative activity, separate from and unrelated to the major concern of the school: the fostering and guidance of children's learning.⁴⁵

Popham and McNeil perceived a discrepancy of purpose between administrators and instructional supervisors or teachers. The administrator seeks knowledge of teacher effectiveness so as to improve decisions concerning hiring or firing a teacher, while the instructional supervisor or teacher wishes to determine which instructional

procedures are most likely to prove useful in achieving certain instructional ends.⁴⁶ One reason for this discrepancy may be that administrators view evaluation in terms of job orientation rather than career development, as Styles and Camp contended.⁴⁷ Arasian presented evidence to support the hypothesis that teachers do not think of accountability in the same rational, efficiency-oriented terms as their administrative superiors. Among classroom teachers accountability is perceived primarily as a method of pointing the finger of blame at individuals.⁴⁸ This perception is congruent with Stenmoe's findings. Undertaking evaluation for the purpose of assisting the teacher to improve teaching competency was favored by 96 percent of the responding teachers while only 54 percent of the same group supported evaluation for the purpose of dismissing poor teachers.⁴⁹

III. CRITERIA

There is a definite division in the literature concerning the criteria for teacher evaluation. The source of this division is the continuing controversy between those who support teaching behavior as evaluation criteria and those who favor the use of measures of pupil gain.

Bolton aligned himself with those favoring the assessment of teacher behaviors when he defined criteria for evaluation purposes as the behavior and results of behavior actually attained by a teacher on the job.⁵⁰ Ryans, an earlier proponent of teacher behavior criteria, believed that the criteria for evaluating teaching should be spelled

out in terms of actual teacher behaviors that could serve as benchmarks against which assessments of individual teacher behavior could be compared. In his 1960 study, Ryans set forth three patterns of teaching behavior:

Pattern X--warm, friendly, understanding vs. aloof, egocentric, and restricted teacher behavior.

Pattern Y--responsible, business-like, systematic vs. evading, unplanned, "slipshod" teacher behavior.

Pattern Z--stimulating, imaginative, surgent vs. dull, routine, teacher behavior.⁵¹

In support of process-oriented teacher behavior criteria Smith stated, "The best we can hope for is some sort of specificity in evaluation based upon scientific knowledge of the effects of various forms of teaching behavior."⁵²

A recent study of teacher evaluation in Tennessee concluded that all criteria for evaluation in use in the state could be categorized as teaching effectiveness and classroom management, personal attributes, professional attitudes and relationships, and out-of-class responsibilities. The diversity of criteria within these categories, however, leads to the conclusion that there is no uniformity of specific criteria among districts in Tennessee.⁵³

Although the position taken by Dunn and Dunn placed them among the supporters of teacher behavior criteria, the scope of behaviors to be evaluated was somewhat different:

It has become increasingly evident that teaching effectiveness is neither dependent on personality characteristics nor on the pooled judgment of experts. Rather, teaching effectiveness can be determined by the quality of the performance of specific teaching tasks, namely: (1) diagnosis, (2) prescription development, (3) guidance of students through the learning process, and (4) evaluation. Done properly these tasks should result in academic growth in the student.⁵⁴

Among those who find fault with process orientation as an evaluative criterion is Howsam who asserted that the only justifiable criterion for teacher evaluation is pupil growth.⁵⁵ Fattu agreed that the obvious measure of teacher effectiveness is the ability of the teacher to bring about changes in the behavior of the students.⁵⁶

One of the most outspoken critics of process-oriented evaluative criteria is James Popham. McNeil and Popham cited as the single most important deficiency in research training effectiveness the failure to use outcome measures as a criterion.⁵⁷ Popham believed that the test of a teacher's effectiveness was his accomplishment of the goals of education as displayed in the pupil's behavior following instruction. Regarding the use of teacher-behavior criteria:

It is indefensible because it is focused on the wrong things, namely, what the teacher does. 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.⁵⁸

In order to meet Popham's standards, criterion measures must:

1. differentiate among teachers,
2. assess learner growth,
3. yield data uncontaminated by required inferences,
4. adapt to teacher's goal preferences,
5. present equivalent stimulus situations (allow for differences in teaching situations), and
6. contain heuristic data categories (data which illuminate the nature of instructional tactics).⁵⁹

Although few, if any, educators would not hold as their primary objective

the intellectual, social and emotional growth of the pupil, the following statement by Hansen reflects the concern of many toward the use of pupil growth as the sole evaluative criterion:

Obviously, good teaching results in student learning. But again, the human, interpersonal, institutional, methods and means, and the content influence dramatically the predictability of that which is to occur.⁶⁰

A more eclectic position toward criteria was taken by Borich who recommended that information be gathered from multiple sources to assure the accuracy and usefulness of the instructional, administrative and planning decisions which are to be based on the appropriate data.⁶¹ Lawrence proposed that these decisions be based on:

1. data gathered by systematic observation instruments,
2. samples of pupil products and descriptions of pupil achievement attributable to the teacher,
3. data gathered by diagnostic tools that measure change in pupil attitudes, perceptions of self and others, motivations and feelings, and
4. records of concrete accomplishments of the teacher according to stated criteria.⁶²

Bolton believed that criteria determined for teacher performance should take into consideration both the personal attributes the teacher brings to the job and the situational constraints placed on him by the job.⁶³

Sergiovanni perceived a distinction between the criteria utilized in what he refers to as the scientific method and those employed in the naturalistic approach to teacher evaluation. In the scientific method

of evaluation, the stated intent of the teacher and the predetermined objective for students were the critical factors. The naturalistic approach is concerned with:

the implicit assumptions and guiding platform statements which teachers bring to the class, the manner in which these assumptions and platform statements are articulated into classroom activities and practices, and the implications and effects of these activities and practices.⁶⁴

McNally postulated that a knowledge and understanding of the criteria by which they are evaluated can be helpful to teachers in making their own self-evaluations, in becoming more clearly aware of the characteristics of good teaching and in planning their own efforts to improve. When choosing criteria, McNally advised that attention be given to their reliability and validity. Validity was defined as the degree to which the factors evaluated are important to the learning of children and to the successful functioning of the school. Reliability was said to be the degree to which different evaluators agree in their evaluations made of a teacher on the same criteria or the extent of agreement of a series of evaluations of the same teacher by one evaluation.⁶⁵

Levin defined validity, more succinctly, as the extent to which an evaluation measures what it is supposed to measure. Reliability was the extent to which an evaluation produces consistent results at different times for different people.⁶⁶

McNally suggested that no single set of criteria would suffice in a variety of settings. Criteria should be related to the needs and conditions of the particular situation.⁶⁷ Hyman proposed that the

supervisor and his colleagues draw up a set of criteria which fulfilled their specific needs.⁶⁸

Summary

There is almost universal agreement that teachers and administrators must concur on acceptable purposes and criteria for teacher evaluation in order for the process to be of value; however, no such agreement is evident in the literature. While the majority of purposes gleaned from the literature can be categorized loosely as either administrative or instructional, no consensus has been reached as to which one, or what combination of those should be emphasized. A similar diversity exists concerning the establishment of criteria for teacher evaluation. There was a clear division between those who favored evaluation of the teaching process, or teacher-behavior criteria, and those who contended that only outcome, or pupil growth measures, represented valid criteria. A third school of thought favored multiple criteria comprised of both process and outcome measures. Most of the literature concerning purposes and criteria for teacher evaluation expressed preferences rather than positions taken on the basis of definitive research in the area.

IV. METHODS OF EVALUATION

Castetter classified ranking, man-to-man comparison, grading, graphic scales, checklists, forced choice methods, and critical incident techniques as traditional or conventional approaches to teacher

evaluation. Forces which impinge on the effectiveness of these approaches included organizational changes, social changes, economic changes, client reaction, and personnel reaction. In Castetter's estimation, a prime factor in the relative ineffectiveness of the traditional approaches has been the failure to link evaluation procedures to organizational purposes, unit objectives and position goals.⁶⁹

The use of teacher checklists and/or rating scales is widespread. In an analysis of teacher evaluation programs in 70 school districts in 38 states, Ingils found that the primary means of evaluation was the rating of teachers in the general categories of personal characteristics, teaching techniques, and professional attitudes.⁷⁰ Lewis defined the trait approach as the appraisal of performance based on the extent to which the teacher exhibited certain characteristics thought to be essential for good teaching. The first phase of this process involves the identification of various traits, characteristics or behaviors believed to be related to good teaching. From these identified factors, a list of descriptive statements is developed. The degree to which the teacher is judged to have displayed the trait or behavior represented by the descriptive statement is often indicated on a rating scale.⁷¹

Barr questioned the reliability of a standardized rating procedure:

Where a number of subjects are employed as is usually the case, drawn from a number of different school systems, or even from the same school system, the teachers rated are generally not rated on the same performance.⁷²

Even with well-developed scales and highly trained raters, Barr contended that one must assume a large amount of transfer from one set of tasks to another.⁷³

Bolton stated that checklists and rating scales allowed the observer to consider clues from a variety of sources before making a judgment. Since these instruments normally are filled out infrequently, however, the delay in recording information could cause errors. Bolton believed that clustering of ratings, wherein several items have identical ratings, may be an indication of leniency, harshness or indecisiveness on the part of the rater. Mediocre ratings were seen as difficult to justify and almost void of useful information for the improvement of performance.⁷⁴

Mouly identified disadvantages of rating scales as halo effects (the tendency to award ratings based on a global judgment of the individual), central tendency (rating all employees in the "middle of the road" ranges), generosity or leniency errors, and a lack of common referent for scoring degrees such as "excellent" or "seldom." He added that the standards used were often not clear in the minds of the individual raters.⁷⁵

McNeil and Popham found that halo, lack of operational definitions, failure to control for sampling of teacher behavior, and the effect of the observer on teacher behavior made rating scales of doubtful use in the hands of school administrators, supervisors and fellow teachers.⁷⁶

In conjunction with halo, leniency, and control tendency, Cascio defined logical or proximity errors as those arising from rating

a particular item because of its nearness, or similarity, to another item.⁷⁷ According to Cascio, the primary disadvantage of rating scales lies in the fact that they give complete control to the rater, thereby offering no safeguards against the above-mentioned errors. Positive aspects of this tool listed by the author were:

Rating scales are less time consuming to develop and administer.

Quantitative results can be determined from rating scales.

Rating scales are dimensionally homogeneous.

Since they are standardized, ratings are comparable across individuals.⁷⁸

Donald Thomas found several problems inherent in the use of rating scales. He stated that the traits and behaviors usually identified are peripheral to actual components of the teaching process. Thomas also felt that this type of evaluation tended to produce stereotypical behavior in teachers and was, therefore, counterproductive. Thomas' synthesis of the research led to the conclusion that, taken individually or in clusters, the necessity of specific traits for successful teaching performance had not been determined.⁷⁹ Mannatt, Palmer, and Hidlebaugh addressed this concern in their research wherein a series of items which expressed teaching behaviors were validated for evaluative purposes. Through statistical analysis, they were able to draw, from a total of 360 items, 94 which could discriminate between high and mediocre teacher performance. Items from this pool used in developing the instrument were categorized as "productive teaching techniques," "positive interpersonal relations," "organized structure classroom management," "intellectual stimulation," and "desirable out-of-class behavior."⁸⁰

Griffith felt that checklists:

1. directed attention to aspects of a lesson which might be otherwise overlooked,
2. lent objectivity to an evaluator's observations,
3. provided a permanent record which is quick and easy to make, and
4. helped a teacher to analyze his or her own lesson and to determine what a supervisor considers important.

Intelligent use of checklists, however, required an awareness that they disregard differences in form and purposes of lessons, make no attempt to weigh factors according to significance, are often superficial, and may supplant careful, reflective judgment by the supervisor.⁸¹

Not all of the concerns about the use of checklists and rating scales deal with reliability, validity and rater error. The American Federation of Teachers passed a resolution at its 1973 convention condemning the use of rating scales because it was felt that they nurtured the exercise of political pressure and created disharmony among the members of a school's staff.⁸²

Lewis' primary concern for the use of the trait approach involved its effect on the principal-teacher relationship. The evaluator is placed in a position of having to do critical review of the teacher's performance based primarily on his/her judgment. Lewis cited research which pointed out negative effects of criticism on the attitude and motivation of the person being evaluated. In the majority of cases, the person being evaluated reacted defensively and was not motivated to improve his performance.⁸³

Wiles and Lovell found rating to be unsatisfactory as the sole procedure for teacher evaluation. Although it produced a judgment which could be used by administrators, it prevented the teacher from asking for needed help and the supervisor from seeking a normal teaching situation.⁸⁴

Sergiovanni and Starratt, who preferred the naturalistic approach to evaluation, believed that rating scales and other teacher evaluation instruments often prevented the evaluator from understanding classroom events and prevented them and the teacher from becoming personally involved in the evaluation process.⁸⁵

The critical incident technique developed by Flanagan for use in the Air Force has been adapted to teacher evaluation procedures in some schools.⁸⁶ Although not widely used in education, the critical incident technique has been viewed with interest in recent years. Critical incidents are reports by knowledgeable observers of things employees did that were especially effective or ineffective in accomplishing parts of their jobs. Positive aspects of this technique include the provision of information on both the static and dynamic aspects of a job, the focus on actual job behavior rather than vaguely defined traits, and the ability to provide meaningful feedback to the evaluatee. As these incidents are collected over time, they can serve a diagnostic purpose within the organization, assisting in the identification of objectives for training programs. Among the drawbacks to the critical incident technique are the lack of participation by raters in the setting of behavioral standards; the relative difficulty in

gratification, thus making comparisons difficult; and the time involved in record keeping, which Cascio characterized as burdensome.⁸⁷

Self-evaluation involves the improvement of instruction through having teachers reflect on their own teaching and modify it accordingly.⁸⁸ The most significant advantage of self-evaluation is that the teacher has the opportunity for improvement without external threat. Despite the potential for instructional improvement underlying this advantage, Bolton stated that self-evaluation may not serve all the purposes of evaluation. Standards used for self-evaluation may not relate readily to outside criteria or to the needs of the school district.⁸⁹

McNeil and Popham found few studies indicating that teachers are self-directing in their learning. They further concluded that there are negligible relationships between self-rating and other criteria, such as student ratings and measures of pupil gain.⁹⁰ Neely detected little enthusiasm for self-evaluation among the Oregon teachers involved in his survey.⁹¹ Wolf postulated that the effects of self-evaluation cannot be separated from teachers' overall attitudes toward evaluation. Teachers surveyed replied that they were not encouraged to evaluate their classroom behavior, indicating that self-evaluation is unlikely to occur or to be productive.⁹²

Evaluation by objectives involves the identification of areas of performance in need of improvement, usually through an interactive process between principal and teacher. These areas of improvement are operationalized through the establishment of job targets. Job targets

include the statement of a performance objective, a summary of the manner in which the objective is to be achieved, and means of measuring the degree of achievement. This evaluation tool came to education from the business world in the form of Management by Objectives (MBO). The principles which are basic to MBO were popularized by Drucker in the 1950's.⁹³ Through the work of Odiorne and others, MBO is perhaps the most widely heralded evaluation technique in business and industry, particularly in the evaluation of managerial personnel. In his definition of Management by Objectives, Odiorne stated that the goals of the organization are jointly defined by the participants in the evaluation process. Each individual's major areas of responsibility are then defined in terms of the results expected of him, and these measures are used as guides for operating the unit and assessing the contribution of each of its members.⁹⁴

Reddin defined MBO in terms of his concept of effectiveness:

The establishment of effectiveness areas and effectiveness standards for managerial positions and the periodic conversion of these with measurable, time-bonded objectives linked vertically and horizontally and with future planning.⁹⁵

Lewis related MBO to the educational organization in the form of his "results approach." The role of the principal-evaluator in the results approach was that of "coach-counselor" rather than the judge or critic.⁹⁶ In terms of organizational benefits, Lewis believed that MBO assigned and distributed management risks to all responsible school officials. Further, the use of this process decreases the need for close expression of personnel and provides a systematic means of tailoring the training and development program to the individual needs of the staff members.⁹⁷

Bolton cited two basic assumptions which comprise the basis for MBO:

1. The clearer the idea a person has of what he is trying to accomplish, the greater his chances of success.
2. Progress can be measured only in terms of goals.⁹⁸

Redfern outlined a process for teacher evaluation which was also an adaptation of MBO.⁹⁹ While the National School Public Relations Association survey showed that the "Redfern Method" was in use in a number of districts, a major problem encountered in this approach was the identification and establishment of realistic job targets.¹⁰⁰ In assessing the use of the Redfern Method over a seven-year span in the Hyde Park schools, Gray and Burns advanced two reasons why MBO may not be as successful in educational organizations. The reward system in education, specifically the salary and advancement of staff members, is not tied to the evaluation of performance and the span of control or supervision in education is much wider than in industry.¹⁰¹

Many of the public demands for measures of accountability in the public schools focus on the use of student test data. Although measures of pupil gain represent a direct indication of desired educational outcomes, Bolton voiced several concerns about their use in evaluating teacher effectiveness. Difficulty in controlling the situation weakens the ability to attribute pupil growth to the behavior of the teacher. Short-term academic gains may be sought at the expense of long-range attitudinal and behavioral growth.¹⁰²

Barr reminded that teachers have great flexibility in choosing purposes, means and methods of instruction and delivered the following

admonition: "Regardless of the validating data reported in test manuals, the tests used in developing the pupil gain criteria will have varying degrees of operational validity"103

When testing the effects of a program of evaluation by student achievement of instructional objectives, Popham placed importance on the reliability and discrimination of the pre- and post-tests to be utilized.¹⁰⁴ While standardized tests may uncover gross instructional deficiencies, Glass emphasized that such instruments were not designed to reveal the variety of ways in which teaching and learning could be "creative, formally opportunistic, and uniquely meaningful to students."¹⁰⁵ Millman cited four reasons why standardized test data were not optimal for program assessment:

1. National and state norm-referenced tests are often not matched with local objectives for instruction.
2. Norm-referenced tests often sacrifice subject matter representativeness in order to obtain maximum discriminability.
3. Content of these tests lacks precise description.
4. These tests lack diagnostic capability.¹⁰⁶

Haefele identified three techniques which utilized standardized test data, each with evident drawbacks. Year-end comparison of pupil performance against established norms did not take into account the differences in scores among schools serving diverse socioeconomic populations. Measuring pupil gain as a function of past student performance under a specific teacher disregarded the varying ability

levels of children from year to year.¹⁰⁷ Research by Daniel Kelpack disclosed diverse standardized test performance between seemingly matched or comparable schools, thus casting doubt on the validity of this practice.¹⁰⁸

Popham cautioned that the main mission of norm-referenced tests was to detect variability among examinees; consequently, the tests are described very generally as to content and skills being measured. This characteristic made norm-referenced tests invalid as a means of organizing instruction.¹⁰⁹ Levin concurred and also questioned the stability of pupil gain for a particular teacher over time.¹¹⁰

Stow failed to achieve significant correlations between student gain scores on criterion-referenced tests and composite scores of teacher performance; however, she discovered that performance appraisal centering on teacher effectiveness criteria coincided with extraordinary results in standardized test scores. The more teachers and administrators studied the research of teaching, the more willing they became to use direct instruction activities as at least part of the district's performance criteria.¹¹¹

Although student evaluation of teaching is prevalent in higher education, it has been received with ambivalence in the public schools. A study by Glen Ovard indicated that, when asked who should be involved in evaluation, teachers preferred everyone over students. Principals and supervisors were most preferred, garnering positive responses from 66.90 percent and 48.25 percent, respectively, of those teachers surveyed. Pupils' involvement was rated positively by only 14.19 percent of the respondents.¹¹²

Eastridge found disagreement among studies regarding the validity of student evaluations in elementary and secondary schools. William Beck is quoted as seeing a difference in the perceived dimensions of merit in teaching between college and elementary students. Perceptions of secondary pupils differed from both. Support for student evaluation came from Baralson, et al., who believed that evaluation procedures which had been successful with college students were applicable to high school students' perceptions as well.¹¹³

Concerns about student evaluation identified by Levin included possible rater bias in favor of "easy teachers," the impact of class size, and teacher reputation. Although several reliable instruments exist, Levin viewed validity as a continuing problem.¹¹⁴ Reiterating the validity question, Haefele stated that student ratings had not been linked with positive student outcomes.¹¹⁵

Despite the division of opinion concerning student evaluation, Barnette and Thompson contended that their study provided preliminary evidence that secondary teachers seem to improve aspects of their instruction more frequently as a result of evaluations by students.¹¹⁶ McNeil and Popham ascribed unquestionable validity to student ratings as indicators of predispositions toward a particular teacher or course.¹¹⁷

Summary

An examination of the literature concerning methodology for evaluation revealed an abundance of criticism of present and past

practices but a paucity of agreement as to the appropriate means for improvement. Those who have done research concerning the effects and/or the acceptance of the various methods admit to a lack of conclusive evidence and call for more intensive examination.

V. PERCEPTUAL FACTORS AFFECTING EVALUATION

The success or effectiveness of a teacher evaluation program can be affected by the perceptions of the various participants. A comparison of the findings of Becker and Hoyle reflected the changing perceptions of administrators toward evaluation. Of the 291 principals interviewed by Becker,¹¹⁸ only 3 showed concern for problems involving evaluation. Hoyle found that teacher evaluation was ranked the second most critical administrative problem by a group of 250 administrators and supervisors.¹¹⁹ The literature indicated several areas of perceptual difference existed among central office administrators, building level administrators and teachers.

Sergiovanni asked supervisors to write three statements which expressed personal feelings about teacher evaluation:

1. a public statement to be shared with teachers, board members, and parents;
2. a confidential statement to be shared only with trusted colleagues; and
3. a private statement which would not be shared readily.

It was postulated that the private and/or the confidential statements would contain doubts about the credibility of the evaluation process.¹²⁰

Results of a 1963 National Education Association survey revealed that, although three out of four superintendents and principals expressed confidence in their school system's program of evaluation, over half of the responding teachers did not share this perception.¹²¹ Somewhat contradictory results were obtained by Stemnock, who reported that 90 percent of teacher-respondents approved of regular evaluation of teachers.¹²² Haefele also felt that teachers favored evaluation, so long as the information collected through the process was utilized as feedback to aid in the teachers' self-evaluation.¹²³

Specific areas of disagreement between teachers and supervisors were identified by McAfee. Disagreement occurred concerning perceptions of who set the objectives for courses taught and the extent to which the objectives were met. Responses of the two groups over the entire questionnaire led to the conclusion that teachers and supervisors "do not often agree as to their estimates of the teachers' abilities, professional skills, attitudes, public relations, and knowledge and use of various methods and techniques of teaching."¹²⁴

Although teachers tended to approve of the concept of evaluation, perceptions of current evaluation practices were generally negative. Teachers responding to the 1963 National Education Association (NEA) study judged their evaluations to be inaccurate with no observable change in behavior resulting from the process.¹²⁵

Reavis concluded that teachers generally did not like the traditional methods of supervision and evaluation,¹²⁶ and half or less of the Connecticut teachers studied by Robinson found their evaluations

useful to them.¹²⁷ Nearly three quarters of the teachers responding in the Barnette and Thompson study stated that no instructional change occurred as a result of their performance evaluations.¹²⁸ Deal and Celotti were unable to attribute any significant increase in the individualization of instruction to evaluation practices at the building or the district level. These authors suggested that the chief contribution of their research was to elaborate further examination of administrative practice.¹²⁹

There is strong evidence in the literature that evaluation is threatening to the person being evaluated. Threat was one of the concepts examined by French, Kay, and Meyer as part of the General Electric studies. The authors concluded that annual appraisals were threatening to self-esteem. The greater number of threats observed in the appraisal interview, the greater the amount of defensive behavior by the appraisee.¹³⁰

Although an advocate of pupil gain as a criterion for teacher evaluation, Popham recognized the possibility that teachers might feel genuinely threatened by the prospect of having their competence assessed by this method.¹³¹ Wolf suggested that most teachers did not view evaluation as being in their interests.¹³² Larry Wicks, an NEA Uniserv Director, qualified this view:

Individual teachers do not believe that they have all the answers on how to teach. They want assistance and constructive guidance from those who can offer help, but when so-called constructive help has a destructive outcome, they then become rightfully skeptical of the entire process.¹³³

Berger stated that local and national teacher organizations had taken the position that professionals should be doing the evaluation and

the process should be given top priority in terms of quality and time.¹³⁴ Teachers' main fear is that they will be fired for frivolous reasons. The fact that evaluations have been carried out by administrators capable of changing the teacher's status--retention, tenure, promotion--formed the basis for Bushman's hypothesis concerning teachers' perceptions of threat.¹³⁵ Salik believed that evaluating teachers on their performance impeded the growth of the helping relationship between principal and teacher. The teacher is many times uncertain about the principal's means for judging him.¹³⁶ Kult also viewed the attitudes and motives of the principal as the primary impediment to the improvement of teacher evaluation.¹³⁷

Popham cited the variety of perceptions of good teaching held by raters as the cause of confusion and inconsistency in the most prevalent teacher evaluation procedures.¹³⁸ Bridges discovered that principals continued to prefer to employ teachers on the basis of expressive characteristics rather than those pointed out by teacher competency research as being "instrumental to teaching competency."¹³⁹

Aside from the preferences and predispositions of principals and other administrators, the inherent problem of role conflict may impede the effectiveness of the evaluator. Owens identified a common source of role conflict coming from the expectation that the administrator would be empathetic and understanding in dealing with his subordinates and, concurrently, will be expected to enforce the rules of the organization.

Many administrators feel this sort of conflict when they zealously attempt to build trust, confidence, and high morale in the teaching staff and then are required to conduct a formal evaluation procedure that seems to be in conflict with these goals.¹⁴⁰

Bolton posited a conflict whenever both accountability and assistance are expected from the same evaluation.¹⁴¹ Lewis agreed that the "coach-counselor" role of the evaluator was in conflict with the judgmental aspects of the traditional practice.¹⁴² Porter, Lawler, and Hackman perceived major conflicts between the goals of the individual, principally the goal of realizing extrinsic rewards from the organization, and the organizational goals of seeking counseling and planning information and seeking information on which to base rewards.¹⁴³ Tjosvold defined cooperative conflicts as those in which there is agreement among the parties concerning the goal to be achieved, but where disagreement exists as to the best means for achievement. Certain role-related conflicts, including the goal-means type, may be aggravated by the emphasis on specialization in teacher-preparation and by "competitive interests" arising from the wish for organizational costs and benefits to be distributed in a manner favorable to a particular group or individual.¹⁴⁴

Cascio delineated two fundamental requirements for the successful use of any performance evaluation system: acceptability to both rater and ratee and relevancy. Acceptability is enhanced markedly, according to Cascio, when both raters and ratees participate fully in system development.¹⁴⁵ Tobin viewed involvement as the key to successful evaluation¹⁴⁶ as did Frank who listed staff involvement as the first ingredient of a successful evaluation process.¹⁴⁷ In McNally's estimation, an evaluation program that is perceived as a cooperative venture of professionals is far more likely to have constructive results.¹⁴⁸

Industrial research by French, Kay and Meyer associated participation in the evaluation process with favorable attitudes toward the job and with high levels of performance for those employees with a high need for independence. Higher levels of participation also resulted in a greater degree of mutual understanding and acceptance of job goals. It was further determined that the setting of specific goals resulted in over twice as much improvement in performance as unilateral criticism without goal setting. Motivation was found to be greater when periodic follow-up conferences were held.¹⁴⁹

Poliakoff applied these principles to teacher evaluation in his prescription for partnership, which included self-evaluation by the teacher, evaluation by the principal, and follow-up consultation after all observations.¹⁵⁰ Whithall and Wood stated that participation in goal-setting and choice of evaluation procedures resulted in greater commitment to aims and greater responsibility for outcomes.¹⁵¹

Summary

An examination of the literature on teacher evaluation, such as the one in the previous chapter indicated a continual interest in the evaluation of teaching performance through the years. The influence of business and industry is evident throughout the history of formal teacher evaluation. Early efforts at increasing teaching efficiency sprang from Taylor's scientific management movement. The human relations movement in which teachers' satisfaction was emphasized, paralleled the same movement in industry. Efforts to find valid means of relating

pupil gain to teaching performance are forms of the industrial product orientation. Finally, the current participatory trends emanated from Management by Objectives, an approach developed and widely used in the private sector.

Although an evolution of theory and opinion is evident, translation of current trends into practice is somewhat lacking. The literature is replete with proposals for performance criteria; however, no consensus can be discerned. A similar situation exists with respect to purposes and methodology for the evaluation process. Writers seem to agree, however, on three points:

1. Evaluation is an organizational necessity.
2. In order to be effective, an evaluation process must be accepted by both evaluators and evaluatees.
3. These two groups must approach performance evaluation with a commonality of purpose.

Much has been written about the value of participation in the development and implementation of the evaluation process. Unfortunately, as in the other areas addressed above, the opinions, experiences, and general common sense of the writers was forced to stand alone because of a relative dirth of generalizable research on this critical topic within the educational context.

CHAPTER III

PRESENTATION AND ANALYSIS OF THE DATA

I. INTRODUCTION

Data analyzed in this chapter were collected by Patricia Miller and Russell Ramsay through the use of the "Teacher Evaluation Questionnaire" developed by Patricia Miller. Usable responses were received from 314 central office administrators, 261 building-level administrators and 259 teachers. The data from the two studies were combined and analyzed through the use of the Statistical Analysis System (SAS) and the services of The University of Tennessee Computer Center. With the exceptions of the Spearman Rank Order Correlation Coefficients and the tests for significant differences between proportions, all statistical treatments were accomplished through the use of the above-mentioned computerized statistical package. The data were organized so as to approximate the design of the questionnaire and were presented under the appropriate research questions.

II. PURPOSES FOR TEACHER EVALUATION

The Spearman Rank Order Correlation technique was utilized to assess the independence of perceived school purposes and personal purposes for evaluation among central office administrators, building-level administrators and teachers. Originally the respondents were asked to rank the three most important purposes for evaluation. The purposes

for each group were ranked according to the number of times each purpose was indicated by a respondent as one of the three most important. The thirteenth category, "others," was omitted from the calculations because its content was unidentifiable.

School Purposes

Comparison of the rankings of central office and building-level administrators, as depicted in Table I, produced a Spearman Rank Correlation (ρ) of .944, significant at .05. In descending order, central office administrators viewed the five most important school purposes as:

1. "improvements of instruction,"
2. "increased job performance,"
3. "staff development and planning,"
4. "meet State Department requirements," and
5. "approval of tenure."

Building-level administrators perceived "improvements of instruction," "increased job performance," "meet State Department requirements," "approval of tenure," and "staff development and planning" as the most important school purposes. The high correlation is indicative of a lack of independence or a high similarity between the perceptions of the two groups.

A ρ value of .538 lacked significance and indicated a marked contrast between the school purposes as perceived by central office administrators and teachers. While the central office group indicated

TABLE I
 FREQUENCIES OF RANKS FOR SCHOOL AND PERSONAL PURPOSES OF TEACHER EVALUATION AS PERCEIVED BY TENNESSEE TEACHERS (T),
 BUILDING-LEVEL ADMINISTRATORS (BL) AND CENTRAL OFFICE ADMINISTRATORS (CO) DURING THE 1978-79 SCHOOL YEAR

Item Titles	CO		BL		T	
	School	Pers	School	Pers	School	Pers
1. Staff Development and Planning	114(3)	(3)168	78(5)	(3)117	28(8)	(4)92
2. Teacher Transfers	12(11)	(2)15	13(11)	(12)2	11(9)	(12)3
3. Provide Feedback to Teachers	49(6)	(4)79	32(8)	(4)85	5(11)	(3)115
4. Meet State Department Requirements	87(4)	(8)20	99(3)	(7)28	169(1)	(7)21
5. Dismissal of Teachers	29(10)	(9)13	30(9)	(9.5)9	55(6)	(10)8
6. Decrease Job Performance	174(2)	(2)209	118(2)	(2)178	68(5)	(2)130
7. Accounting to Authorities	44(8)	(10)11	52(6)	(9.5)9	108(3)	(9)9
8. Improvements of Instruction	225(1)	(1)269	158(1)	(1)211	69(4)	(1)181
9. Set Standards of Performance	46(7)	(6)54	38(7)	(5)46	35(7)	(6)57
10. Approval of Tenure	85(5)	(7)21	95(4)	(8)24	124(2)	(8)10
11. Improve Communications	32(9)	(5)59	26(10)	(6)39	10(10)	(5)71
12. Promotion of Teachers	3(12)	(11)4	3(12)	(11)4	3(12)	(11)7
13. Others	0(13)	(12.5)0	2(13)	(13)0	2(13)	(13)0

NOTE: Rankings are provided in parentheses above each frequency of response.

a strong orientation toward instructional improvement and staff development, the purpose of evaluation most often ranked by teachers was "meet State Department requirements," followed in descending order by:

1. "approval of tenure," and
2. "accounting to authorities."

Over 65 percent of the teacher-respondents perceived the meeting of State Department requirements to be one of the three most important purposes of teacher evaluation in their school district. Only 12.7 percent of the teachers perceived staff development to be one of their district's primary purposes.

Although the correlation coefficient achieved between the perceived school purposes of building-level administrators and teachers (.769) was lower than the rho between central office and building-level administrators, it was significant at .05, attesting to the high degree of harmony between the perceptions of the two groups.

Personal Purposes

Purposes for teacher evaluation personally considered to be important were highly congruent among the three groups. A Spearman Correlation Coefficient of .979 was achieved between central office and building-level administrators. Central office administrators and teachers agreed at a level represented by a coefficient of .975. The purposes identified by building-level administrators and teachers correlated at the .984 level. All three coefficients were significant at .05.

The three groups agreed that the improvement of instruction was the paramount purpose of teacher evaluation. The second most often ranked purpose was increasing job performance. Teachers felt that the provision of feedback was the third most important purpose of the evaluation process while building-level and central office administrators placed staff development and planning in this position.

Rankings for the three groups in the fourth position were exact reversals of the third ranked purposes. Both groups of administrators ranked the provision of feedback in this position while teachers awarded the fourth ranking to staff development and planning. Central office administrators and teachers ranked the improvement of communication as their fifth choice while building-level administrators viewed the setting of standards of performance as their fifth most important purpose.

III. METHODS OF EVALUATION

Significant differences among the three groups concerning the desirability of various methods of evaluation were determined through the use of the analysis of variance procedure: Where an F ratio was found to be significant at .05, Duncan's Multiple Range Test was employed to locate the difference among the three focal groups. Perceptions of the desirability of methods were indicated on a five-point scale with the following descriptions:

1. "very desirable,"
2. "desirable,"
3. "undesirable,"

4. "very undesirable," and
5. "insufficient information."

For the purposes of these calculations, the fifth category was omitted.

As indicated in Table II, no significant differences occurred among any of the three groups concerning the desirability of the six categories of teacher checklists. The use of physical characteristics in evaluative checklist received a mean rating of 2.09 from the three groups. Personal characteristics were rated at a mean of 1.67, professional qualities at 1.28, attitudes and values at 1.42, and professional training and experience at 1.46.

A relatively high degree of difference concerning the desirability of using professional training and experience was evidenced by an achieved significance level of .052. Both classifications of administrators rated this item at 1.50; however, teachers assigned it a mean rating of 1.38. The category "other" received a mean rating of 1.63.

Records of critical incidents or events were perceived differently by administrators and teachers. Central office administrators compiled a mean rating of 2.15. Building-level administrators had a mean rating of 2.06 and teachers' collective rating was 2.40. Concerning the use of parental input in teacher evaluation, a consensus among teachers, building-level and central office administrators was evidenced by ratings of 2.50, 2.51, and 2.51, respectively.

Although the critical value was not exceeded in the case of the item dealing with self-evaluation, the mean responses of administrators and teachers were different at .055. The mean teacher-response to this item was 1.51 while the average responses for central office and

TABLE II

MEAN RESPONSES OF CENTRAL OFFICE ADMINISTRATORS (CO), BUILDING-LEVEL ADMINISTRATORS (BL) AND TEACHERS (T) IN TENNESSEE CONCERNING THE DESIRABILITY OF VARIOUS METHODS OF EVALUATION AS MEASURED IN SECTION III OF THE TEACHER EVALUATION QUESTIONNAIRE IN USE DURING THE 1978-79 SCHOOL YEAR

Item Titles		*Mean Responses				**Significant Differences** (PL .05)				
		Item N	CO	Item N	BL	Item N	T	CO	BL	T
A. Teacher Checklist										
1. Physical Characteristics		277	2.04	232	2.08	213	2.15			
2. Personal Characteristics		300	1.63	252	1.66	240	1.72			
3. Professional Qualities		308	1.25	257	1.31	241	1.28			
4. Attitudes and Values		304	1.42	253	1.43	242	1.42			
5. Professional Training and Experience		294	1.50	253	1.50	237	1.38			
6. Other		47	1.72	43	1.67	22	1.50			
B. Critical Incidents or Events		272	2.15	227	2.06	150	2.40	T	T	BL, CO
C. Input From Parents		259	2.51	216	2.51	148	2.50			
D. Self-Evaluation		300	1.60	250	1.66	233	1.51			
E. Evaluation by Objectives										
1. Job Targets Identified by Teachers and Administrators		295	1.46	236	1.58	194	1.66	T, BL	CO	CO
2. Job Targets Identified by Teachers		291	1.99	224	2.08	184	1.76	T	T	CO, BL
3. Job Targets Identified by Administrators		279	2.66	220	2.68	163	2.96	T	T	CO, BL
4. Job Targets Based on a Needs Assessment		287	2.00	221	2.15	169	1.89	BL	CO, T	BL
5. Needs Assessment Based on a Job Description		273	2.07	211	2.28	145	1.92	BL	CO, T	BL
6. Other		36	2.36	36	2.33	22	2.36			
F. Competency Tests for Teachers		271	2.59	210	2.91	184	2.64	BL	CO, T	BL

TABLE II (Continued)

Significant Differences* (PL.05)										
* Mean Responses										
Item Titles		Item N	CO	Item N	BL	Item N	T	CO	BL	T
G. Student Test										
1. Standardized Tests		284	2.59	223	2.61	200	2.80	T	T	CO, BL
2. Proficiency Tests		277	2.60	222	2.63	192	2.75			
3. Teacher-made Tests		284	2.56	221	2.54	197	2.46			
4. Measures - Paper and Pencil		272	2.64	212	2.74	168	2.74			
5. Student Reports of Teacher Behavior		273	2.73	217	2.80	176	2.86			
6. Student Attitude Measures - Observations		261	2.64	215	2.68	181	2.67			
7. Others		39	2.74	30	2.73	22	2.54			
H. Classroom Observations										
1. Principal(s)		310	1.23	257	1.29	245	1.53	T	T	CO, BL
2. Supervisor(s)		307	1.59	253	1.69	221	1.87	T	T	
3. Superintendent		290	2.37	234	2.54	191	2.52	BL, T	CO	CO
4. Department Chairman or Team Leader		283	2.30	232	2.41	180	2.38			
5. Peers		277	2.65	228	2.67	179	2.58			
6. Others		49	2.55	31	2.74	30	2.40			
I. Pre-observation Conference										
1. Principal(s)		306	1.33	256	1.39	221	1.61	T	T	CO, BL
2. Supervisor(s)		297	1.71	248	1.86	197	1.98	BL, T	CO	CO
3. Superintendent		281	2.44	231	2.60	173	2.57			
4. Department Chairman or Team Leader		277	2.37	227	2.50	162	2.41			
5. Peers		266	2.71	215	2.77	152	2.62			
6. Others		42	2.62	32	3.00	32	2.53			
J. Post-observation Conference										
1. Principal(s)		307	1.26	256	1.26	231	1.47	T	T	CO, BL
2. Supervisor(s)		301	1.60	248	1.71	197	1.79	L, T	CO	CO
3. Superintendent		282	2.40	237	2.51	173	2.48			
4. Department Chairman or Team Leader		271	2.34	224	2.47	159	2.40			
5. Peers		262	2.70	212	2.74	148	2.58			
6. Others		40	2.60	27	2.89	27	2.55			

TABLE II (Continued)

Item Titles	*Mean Responses			**Significant Differences*** (PL.05)			
	Item N	CO	Item N	BL	Item N	T	BL
<u>K. Instructional Conferences</u>							
1. Planning	300	1.31	245	1.37	288	1.39	
2. Formative	303	1.36	246	1.43	218	1.41	
3. Summative	299	1.40	240	1.46	212	1.48	
<u>L. Casual Information from Others</u>	278	2.57	230	2.38	159	2.43	CO

*Response scale - 1 = very desirable, 2 = desirable, 3 = undesirable, 4 = very undesirable

**The codes which appear under each column heading represent the group(s) whose responses differed significantly from the group indicated by the column heading. A blank indicates no significant differences.

***Significant differences were determined using the Duncan's Multiple Range Test.

building-level segments were 1.60 and 1.66 respectively. Duncan's Multiple Range test indicated that the difference lay between building-level administrators and teachers.

Five of the six items dealing with evaluation by objectives displayed significant differences in excess of .05 among the various segments of respondents. Central office administrators (1.46) differed from building-level administrators (1.58) and teachers (1.66) regarding the desirability of job targets identified by teachers and administrators. Teachers (1.76) failed to concur with either building-level (2.08) or central office administrators (1.99) concerning job targets identified by teachers only. A similar situation was in evidence on job targets identified solely by administrators. Teachers rated the item a 2.96, while central office and building-level administrators were in relative agreement at 2.66 and 2.68 respectively.

Job targets resulting from a needs assessment received a rating of 2.00 from central office administrators, 2.15 from building-level administrators, and 1.89 from teachers. The rating of the building-level group differed significantly from both of the remaining segments. "Needs assessment based on a job description" drew a rating of 2.07 from the central office section, which did not differ from the teacher response of 1.92. The building-level response of 2.28 differed from both of the aforementioned groups. No significant difference is evident in the responses to the "other" category, where a mean rating of 2.35 was posted.

No significant difference was evident in the ratings given by central office administrators (2.59) and teachers (2.64) to competency

tests for teachers; however, building-level administrators (2.91) differed significantly from both.

Of the seven items dealing with student test data for teachers, only the desirability of standardized test data was perceived at a significantly different level among the three groups. At the .05 level, teachers (2.80) disagreed with both building-level (2.61) and central office administrators (2.59). Concerning the other six items, the mean responses across the three factions were:

2. "Proficiency Tests" - 2.66,
3. "Teacher-made Tests" - 2.54,
4. "Student Attitude Measures" - paper and pencil - 2.71,
5. "Student Reports of Teachers' Behaviors" - 2.80,
6. "Student Attitude Measures - observations" - 2.66, and
7. "Other" - 2.67.

Classroom observations were rated at 1.53 by teachers. This rating is significantly different from the ratings of central office (1.23) and building-level administrators (1.29) at the .05 level. Concerning observations by supervisors, teachers (1.87) again differed significantly from central office (1.59) and building-level administrators (1.69) at .05.

Central office administrators (2.37) perceived observations by the superintendent at a level significantly different from both building-level administrators (2.54) and teachers (2.52) at .05. No significant differences were achieved among the other segments concerning observations by the department chairman or team leader rated by the total group at a mean of 2.36; by peers (2.63); or by others (2.56).

Perceptions of central office (1.33) and building-level administrators (1.39) did not differ significantly with respect to the desirability of pre-observation conferences by principals. Teachers, however, viewed such conferences in a manner significantly different from both administrative groups (1.61).

A different division was noted regarding pre-observation conferences by supervisors. Central office administrators (1.71) differed significantly from building-level administrators (1.86) and teachers (1.98). No significant differences were found regarding pre-observation conferences by superintendents (2.54), department chairmen and/or team leaders (2.43), peers (2.70), or others (2.72).

Post-observation conferences by principals were regarded in the same manner by central office (1.26) and building-level administrators (1.26), while teachers (1.47) differed significantly from both groups. Such conferences by supervisors were perceived approximately the same by building-level administrators (1.71) and teachers (1.79), but central office administrators (1.60) reported perceptions which were significantly different from both groups. Post-observation conferences by superintendents (2.46), department chairmen and/or team leaders (2.40), peers (2.67) and others (2.68) were viewed without significant deviation among the three groups.

The three divisions of respondents viewed the desirability of each of the three categories of instructional conferences without significant differences. The mean rating of all respondents concerning planning conferences was 1.36. Formative conferences received a mean rating of 1.4, and summative conferences were rated at 1.45.

Central office administrators (2.57) differed significantly from building-level administrators (2.38) concerning the use of casual information from others in teacher evaluation. Neither group posted a mean rating significantly different from teachers (2.43).

Evaluation Priorities

Section IV of the Teacher Evaluation Questionnaire focuses on evaluation priorities. The first six items, IV-A through F, deal with the amount of time spent in the evaluation process, classroom observation and conferencing. These items were analyzed through the analysis of variance technique using a .05 significance level and Duncan's Multiple Range Test for significant differences.

How much time was spent in the evaluation process in one year?

Analysis of variance revealed highly significant differences among all three groups (Table III). Building-level administrators achieved a mean response of 3.28 on this item, indicating over 5.9 hours spent per teacher in the yearly evaluation process. Central office administrators expressed a mean rating of 2.26, or slightly over 3.9 hours per teacher. Teachers perceived the smallest amount of time spent in evaluation with mean assessment of 1.59, or approximately 3.25 hours.

How much time should be spent per teacher in the evaluation process in one year?

Significant differences were again achieved among all three segments of the study. Building-level administrators estimated that approximately 9.54 hours, a mean rating of 4.82, should be spent per

TABLE III

MEAN RESPONSES OF TEACHERS (T), BUILDING-LEVEL ADMINISTRATORS (BL) AND CENTRAL OFFICE ADMINISTRATORS (CO) IN TENNESSEE CONCERNING THE TIME SPENT IN TEACHER EVALUATION AND OBSERVATIONS ACCOMPANYING AN EVALUATION AS PERCEIVED DURING THE 1978-79 SCHOOL YEAR

Item Titles	Mean Responses*			Significant Differences		
	CO N=314	BL N=261	T N=243	CO	BL	T
IV-A. Yearly time spent - evaluation per teacher	2.26	3.28	1.59	BL,T	CO,T	CO,BL
IV-B. Yearly time should be spent in evaluation per teacher	4.02	4.82	3.12	BL,T	CO,T	CO,BL
IV-C. Average times observed during evaluation each year	3.16	3.96	2.63	BL,T	CO,T	CO,BL
IV-D. Average length of the observation	4.36	4.14	2.62	T	T	CO,BL
IV-E. Average times should be observed in evaluation each year	4.13	4.90	3.71	BL,T	CO,T	CO,BL
IV-F. Average length observations should be	5.05	4.75	3.84	BL,T	CO,T	CO,BL

*Mean responses are expressed in terms of response codes rather than actual time spent, length of observations or numbers of observations.

Response Codes: Yearly Time Spent in Evaluation Number of Times Observed Length of Observations in Minutes

1. 0-1.9 hours	6. 10-11.9 hours	1. 0	4. 5-6	7. 11-12	1. Less than 10	4. 30	7. 60
2. 2-3.9 hours	7. 12-13.9 hours	2. 1-2	5. 7-8	8. 13-14	2. 10	5. 40	8. 70
3. 4-5.9 hours	8. 14-15.9 hours	3. 3-4	6. 9-10	9. 15+	3. 20	6. 50	9. 71+
4. 6-7.9 hours	9. 16+ hours						
5. 8-9.9 hours							

teacher in the evaluation process. Central office administrators recorded the second highest rating of 4.02, or 7.94 hours. Teachers believed that approximately 6.14 hours should be spent in the evaluation process, as evidenced by a mean response of 3.12.

On the average, how many times is a teacher observed during an evaluation year by all those involved in the evaluation process?

Building-level administrators accumulated the highest mean response (3.96), reporting that a teacher is observed almost 5 times during an evaluation year. Central office administrators (3.15) indicated that 3 to 4 observations were accomplished. Teachers (2.63) believed that no more than 2 to 3 observations accompanied their evaluations.

What should be the average length in minutes of an observation?

No significant difference was found between the responses of central office and building-level administrators (4.14) concerning the length of observation. The central office response (4.36) translated into approximately 33 minutes per observation and the building-level response of 4.14 converted to a slightly lower 31 minutes. The teacher response of approximately 16 minutes was significantly different from the responses of both administrator groups.

On the average how many times should a teacher be observed during an evaluation year by all those involved in the evaluation process?

The mean responses of each of the three factions proved to be significantly different from those of the remaining groups. Building-level administrators (4.90) asserted that approximately 7 to 8

observations should be made during an evaluation year. Central office administrators (4.13) felt that 5 to 6 observations should occur, and teachers (3.71) indicated that 4 or 5 observations were sufficient.

What should be the average length in minutes of an observation?

The mean teacher response to this question was 3.84, or approximately 28 minutes. Building-level administrators, whose mean response was 4.75, indicated an optimum observation time of approximately 37 minutes. The mean perception of central office administrators, 5.05, was significantly higher than those of the other two groups, reflecting an optimum observation time of some 40 minutes.

Teacher Participation

With the overall rating solicited in the final item on the questionnaire, the items in section IV-G of the Teacher Evaluation Questionnaire explored various ways in which teachers may or may not participate in the development and implementation of the teacher evaluation process. Data concerning these items were analyzed using the "significant difference between proportions" procedure. For the purposes of this procedure, only the affirmative responses of the three groups were utilized.

Do teachers have an opportunity to see and review their evaluation documents?

As depicted in Table IV, teacher responses differed significantly from those of both central office and building-level administrators. While 85.7 percent of central office administrators and 88.9 percent of building-level administrators responded that teachers always had the

TABLE IV
PROPORTIONS OF TEACHERS (T), BUILDING-LEVEL ADMINISTRATORS (BL) AND CENTRAL OFFICE ADMINISTRATORS (CO) IN TENNESSEE RESPONDING AFFIRMATIVELY TO THE FIRST SEVEN ITEMS IN PART IV-G OF THE TEACHER EVALUATION QUESTIONNAIRE DURING THE 1978-79 SCHOOL YEAR

Item Titles	*Proportion Responding "yes"			Significant Difference		
	CO	BL	T	CO	BL	T
IV-G. Teachers' Opportunity to See and Review Their Evaluation Documents						
1. "Yes, always"	(269) .857	(232) .889	(144) .574	T	T	CO, BL
2. "Yes, if they wish"	(42) .134	(27) .103	(90) .359	T	T	CO, BL
IV-H. Teachers' right to make written comments on disagreements	(298) .949	(245) .939	(144) .571	T	T	BL
IV-I. Teachers' receiving a copy of the final evaluation	(217) .691	(183) .701	(97) .386	T	T	CO, BL
IV-J. Teachers informed of the evaluation process	(309) .984	(260) .996	(189) .753	T	T	CO, BL
IV-K. Teachers involved in the development of evaluation instruments	(257) .818	(186) .713	(58) .229	BL, T	CO, T	CO, BL
IV-L. Teachers involved in the development of evaluation procedures	(244) .777	(182) .697	(51) .202	BL, T	CO, T	CO, BL
IV-M. Teachers informed of the evaluation appeal process	(272) .866	(226) .865	(88) .349	T	T	CO, BL

*Numbers in parentheses are the frequencies of affirmative responses. See Appendix for supporting Z scores.

opportunity to see and review their evaluation documents, a significantly lower proportion of teachers, 57.4 percent, responded in a like manner. A significantly higher proportion of teachers (35.9 percent) disclosed that they could see and review their evaluation if they wished than either central office administrators (13.4 percent) or building-level administrators (10.3 percent).

Do teachers have the right to make written comments on disagreements concerning their evaluation?

Affirmative responses were gathered from 94.9 percent of the central office segment. Of the building-level administrators, 93.9 percent agreed that teachers had the right to disagree in writing with their evaluations. A significantly smaller proportion of teachers (57.1 percent) perceived that they were afforded this right.

Do teachers receive a copy of the final evaluation that is placed in their personnel file?

The pattern established in the preceding items was repeated here. Central office (69.11 percent) and building-level administrators (70.1 percent) agreed that teachers did receive a copy of their final evaluation; however, a significantly lower percentage of teachers (38.6) affirmed this practice.

Are teachers informed of the evaluation process?

Significant differences were again achieved between the proportion of teachers who perceived that they were informed of the evaluation process and the proportion of administrators who felt that this task had been accomplished. Approximately 99 percent of both

administrator groups believed that teachers were aware of the process by which they were evaluated, while only 75.3 percent of the teachers concurred.

Are teachers involved in the development of evaluation instruments?

A large majority of the central office administrators (81.8 percent) asserted that teachers were involved in the development of the instruments with which they were evaluated. A significantly smaller percentage of building-level administrators (71.26 percent) shared this perception. The proportion of teachers (22.9 percent) who testified to their involvement in the development of evaluation instruments was significantly lower than either of the two groups of administrators.

Are teachers involved in the development of evaluation procedures?

All three segments of respondents were significantly different in their perceptions of teacher involvement in the development of evaluation procedures. The proportion of teachers (20.2 percent) who perceived themselves to be involved in this process contrasted sharply with the percentage of building-level administrators (69.7 percent) giving a like response. This group, in turn, was significantly lower than the proportion of central office administrators (77.7 percent).

Are teachers informed of the appeal process?

Significantly fewer teachers responded affirmatively to this question than did either administrator group. An almost identical

percentage of central office and building-level administrators (86.6 percent) maintained that teachers were told about the evaluation appeal process; yet only 34.9 percent of teachers' responses were in concert.

How would you rate the overall evaluation process in your school system?

The mean teacher response to this item was slightly below "Fair" (2.21). Central office and building-level administrators tabulated an almost identical mean response of 1.67, indicating a more favorable perception of the evaluation process. As depicted in Table V, the analysis of variance procedure verified that teachers and administrators differed significantly on this question.

Teachers' overall ratings of the evaluation process as a function of their involvement in the development of the evaluation procedure.

Data from teacher-respondents were analyzed through the use of a "students t" test comparing the overall ratings of the sub-group who stated that they were involved in the development of their evaluation process with the sub-group who indicated no involvement.

The mean rating of the 51 teachers who participated in the development of their evaluation procedure (1.57) was found to be significantly higher than the mean rating of some 202 teachers who had not participated (2.26) evidenced by the calculated t value of 6.47.

Teachers' overall ratings as a function of their involvement in the development of evaluation instruments.

The mean rating given by the group of teachers who attested to their involvement in the development of their evaluation instruments

TABLE V

ANALYSIS OF VARIANCE OF THE OVERALL RATINGS GIVEN TO THE TEACHER
EVALUATION PROCESS BY TEACHERS, CENTRAL OFFICE ADMINISTRATORS,
AND BUILDING-LEVEL ADMINISTRATORS AS IMPLEMENTED DURING
THE 1978-79 SCHOOL YEAR

Source	DF	SS	MS	F
Model	2	36.98	18.49	
Error	<u>826</u>	<u>371.52</u>	<u>0.45</u>	<u>41.11*</u>
Total	828	408.50		

*Significant at .05.

(N = 88) was 1.73. Those teachers who perceived themselves not to be involved (N = 164) reported a significantly less favorable mean rating of 2.33.

Teachers' overall ratings of the evaluation process as a function of their participation in the implementation of the process.

The extent of teacher-participation in the implementation of the evaluation process was assessed by combining the teachers' responses to five of the items in section IV-G of the instrument. Responses to the item concerning the opportunity to see and review evaluation documents were collapsed so that all who responded either "yes, always" or "yes, if they wish" were given a participation score of one. If this item and the following item, dealing with the right to make written comments or disagreements about evaluation were both answered affirmatively, the teacher received a participation score of two. Other items from section IV-G which were utilized in developing the participation score were those related to reception of a copy of the final evaluation, appraisal of the evaluation process, and notification of the evaluation appeal process. An affirmative answer to each of these items increased the participation score by one so that scores could range from zero, indicating no participation in the above-mentioned areas, to five, indicating maximum participation.

Regression analysis of the participation scores of 142 teachers with their overall ratings of the process by which they were evaluated placed the y intercept at approximately 2.39. The slope of the least-squares line was determined to be -0.16 indicating that the rating given to the overall evaluation process became lower, or more favorable, as the

respondents' participation scores became higher. Both the slope and the intercept were proven to be significant at .05. Figure I provides an illustration of this relationship.

The Pearson Product Moment Correlation Coefficient, which was used to further assess the extent of the relationship, was found to be $-.35$, significant at .0001. The unusually high level of significance attests to the strength of the relationship between the two variables.

Significant differences in rankings of school purposes of evaluation within groups as a function of the method of evaluation utilized.

Within the three sub-groups of respondents--teachers, building-level administrators, and central office administrators--frequency of rankings of purposes for evaluation were broken down by the method of evaluation checked by the respondent. Where a method of evaluation item contained more than one category, such as the teacher checklist item, any respondent who checked at least one category was included in the calculations for that item. The Spearman Rank Order Correlation technique was used to compare the rankings of purposes by respondents utilizing each of the 12 methods of evaluation. Each sub-set of rankings was also compared to the rankings of the total population segment of which it was a part. A significance level of .05 was maintained for each calculation.

Teachers

A total of 259 teacher-respondents were involved in this procedure. The 12 sets of rankings depicted in Table VI correlated highly with each other and with the overall teacher ranking of purposes

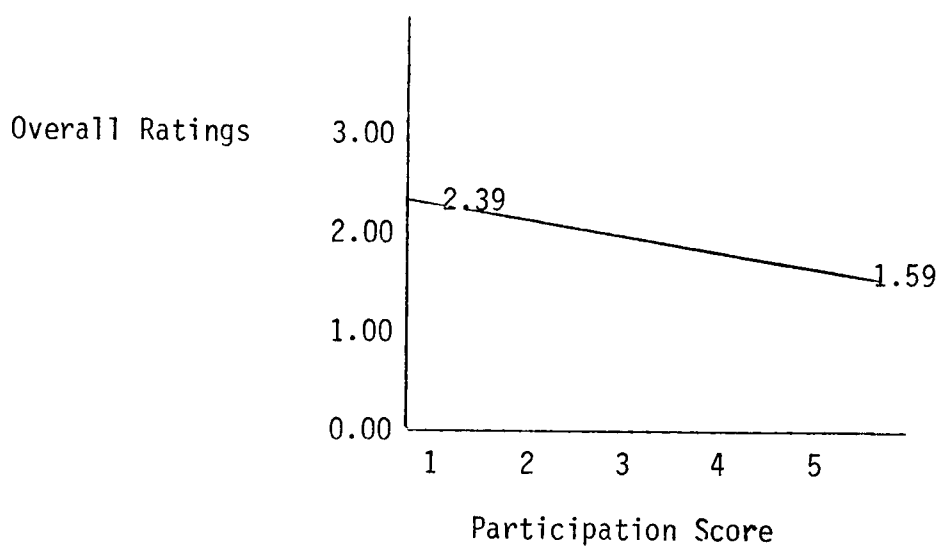


FIGURE I

GRAPHIC REPRESENTATION OF THE REGRESSION EQUATION DEVELOPED
FOR TENNESSEE TEACHERS' OVERALL RATINGS OF THEIR
EVALUATION PROCEDURES AS A FUNCTION OF THEIR
PARTICIPATION SCORES AS ASSESSED DURING
THE 1978-79 SCHOOL YEAR

TABLE VI
TENNESSEE TEACHERS' RANKINGS OF PURPOSES OF TEACHER EVALUATION ACCORDING TO THE METHOD OF
EVALUATION USED DURING THE 1978-79 SCHOOL YEAR

Purposes	Teacher Checklist	Critical Events	Input From Parents	Self-Evaluation	Evaluation by Objectives	Teacher Competency Tests	Student Test Data	Classroom Observations	Pre-Observation Conferences	Post-Observation Conferences	Instructional Conferences	Casual Info. From Others
1. Staff Development and Planning	(28)	(5)	(3)	(18)	(21)	(1)	(16)	(27)	(18)	(21)	(25)	(9)
	8	8	8	8	6	5.5	6	8	8	8	7	10
2. Teacher Transfers	(11)	(2)	(0)	(3)	(4)	(1)	(1)	(11)	(4)	(8)	(6)	(11)
	9	9	11.5	11	10.5	5.5	11.5	9	10	9	10	8
3. Provide Feedback to Teachers	(5)	(0)	(0)	(3)	(4)	(0)	(0)	(4)	(3)	(4)	(4)	(5)
	11	11	11.5	11	10.5	11	13	11	11	11	11	11
4. Meet State Dept. Requirements	(166)	(34)	(13)	(86)	(91)	(2)	(52)	(156)	(100)	(133)	(77)	(167)
	1	1	1	1	1	1.5	1	1	1	1	1	1
5. Dismissal of Teachers	(54)	(9)	(4)	(21)	(14)	(0)	(9)	(49)	(41)	(46)	(49)	(55)
	6	6	6	7	8	11	8	6	5	6	3	6
6. Increase of Job Performance	(67)	(13)	(4)	(41)	(39)	(2)	(28)	(63)	(36)	(56)	(40)	(62)
	5	5	6	4	5	1.5	4	5	6	5	6	5
7. Accounting to Authorities	(104)	(26)	(8)	(56)	(56)	(1)	(33)	(95)	(43)	(73)	(45)	(108)
	3	3	2.5	3	2	5.5	2	3	3.5	3	4	3
8. Improvement of Instruction	(69)	(19)	(8)	(40)	(46)	(1)	(25)	(65)	(43)	(59)	(44)	(69)
	4	4	2.5	5	4	5.5	5	4	3.5	4	5	4
9. Set Standards of Performance	(35)	(6)	(4)	(23)	(20)	(1)	(13)	(34)	(24)	(40)	(23)	(35)
	7	7	6	6	7	5.5	7	7	7	7	8	7
10. Approval of Tenure	(121)	(23)	(7)	(59)	(53)	(1)	(29)	(112)	(70)	(98)	(56)	(122)
	2	2	4	2	3	5.5	3	3	3	2	2	2
11. Improve Communication	(10)	(3)	(2)	(6)	(7)	(0)	(6)	(9)	(7)	(7)	(7)	(10)
	10	10	9	9	9	1	9	10	9	10	9	9
12. Promotion of Teachers	(3)	(2)	(1)	(3)	(3)	(0)	(2)	(3)	(1)	(1)	(3)	(3)
	12	12	10	11	12	1	10	13	13	12	12	12

Note: Frequencies of responses are provided in parentheses above each ranking.

for evaluation. Regardless of the method employed in their evaluation, teachers consistently perceived the three main purposes of evaluation to be meeting State Department requirements, approval of tenure, and accounting to authorities.

Four of the rankings, those for respondents using teacher checklists, critical incidents or events, classroom observations, and post-observation conferences, were identical to the overall teacher ranking. The only marked deviation from the overall ranking was among those who were evaluated using teacher competency tests. The minute number of cases involved in this method, however, caused a distortion of the rankings and made valid interpretation impossible.

Building-Level Administrators

Although one deviation of ranks among methods of evaluation was more apparent with this group than with teachers, the various Spearman Correlation Coefficients did not point to the independence of any of the rankings displayed in Table VII. One or two more building-level respondents reported the use of teacher-competency tests; however, the frequencies were insignificant, making reasonable interpretation impossible.

Building-level administrators perceived the two primary purposes of evaluation to be "improvements of instruction" and "increased job performance." The next three positions in the rankings were divided among "meet State Department requirements," "approval of tenure," and "staff development and planning." In ascending order, the least important purposes for evaluation were "promotion of teachers," "teacher transfers,"

TABLE VII
BUILDING-LEVEL ADMINISTRATORS' RANKINGS OF PURPOSES FOR TEACHER EVALUATION ACCORDING TO THE METHOD OF
EVALUATION UTILIZED DURING THE 1978-79 SCHOOL YEAR

Purposes	Teacher Checklist	Critical Events	Input From Parents	Self-Evaluation	Evaluation by Objectives	Teacher Competency Tests	Student Test Data	Classroom Observations	Pre-Observation Conferences	Post-Observation Conferences	Instructional Conferences	Casual Info. From Others
1. Staff Development and Planning	(77) 5	(21) 5	(12) 3	(53) 4	(53) 5	(1) 5.25	(31) 3	(73) 5	(58) 5	(69) 5	(64) 4	(13) 10
2. Teacher Transfers	(11) 11	(6) 10.5	(1) 11.5	(5) 11	(6) 11	(1) 5.25	(5) 11	(11) 11	(8) 11	(9) 11	(7) 11	(13) 11
3. Provide Feedback to Teachers	(31) 8	(12) 8	(4) 8	(22) 8	(17) 10	(2) 4	(14) 7	(29) 8.5	(20) 9.5	(25) 9	(20) 8	(32) 7
4. Meet State Dept. Requirements	(93) 3	(25) 4	(8) 4	(63) 3	(59) 4	(4) 2.5	(28) 5	(92) 3	(71) 3	(82) 3.5	(64) 4	(99) 3
5. Dismissal of Teachers	(27) 9	(12) 8	(2) 10	(15) 10	(20) 9	(0) 11	(7) 10	(29) 8.5	(20) 9.5	(27) 8	(19) 9	(30) 2
6. Increase of Job Performance	(114) 2	(38) 2	(13) 2	(70) 2	(78) 2	(4) 2.5	(48) 2	(113) 2	(95) 2	(106) 2	(88) 2	(116) 2
7. Accounting to Authorities	(48) 6	(12) 8	(4) 8	(27) 6	(30) 6	(1) 5.25	(16) 6	(46) 6	(29) 6.5	(38) 6	(29) 7	(52) 5
8. Improvement of Instruction	(154) 1	(52) 1	(21) 1	(99) 1	(108) 1	(5) 1	(62) 1	(159) 1	(125) 1	(142) 1	(120) 1	(158) 1
9. Set Standards of Performance	(37) 7	(14) 6	(4) 8	(26) 7	(25) 8	(0) 11	(12) 8	(35) 7	(29) 6.5	(35) 7	(31) 6	(38) 6
10. Approval of Tenure	(88) 4	(28) 3	(6) 5	(51) 5	(60) 3	(1) 5.25	(30) 4	(87) 4	(70) 4	(82) 3.5	(64) 4	(95) 4
11. Improve Communication	(26) 10	(6) 10.5	(5) 6	(16) 9	(27) 7	(0) 11	(8) 9	(26) 10	(21) 8	(23) 10	(19) 10	(26) 9
12. Promotion of Teachers	(3) 12	(1) 12.5	(1) 11.5	(1) 12	(2) 12.5	(0) 11	(2) 12	(3) 12	(2) 12	(2) 12	(2) 12	(3) 12

Note: Frequencies of response are provided in parentheses above each ranking.

"improve communications," "dismissal of teachers," and "provide feedback to teachers." Building-level respondents numbered 261.

Central Office Administrators

As indicated in Table VIII, central office administrators ranked in order "improvements of instruction," "increase job performance," "staff development and planning," "meet State Department requirements," and "approval of tenure" as the most important purposes of teacher evaluation without significant deviation with respect to methods utilized. Least important among the teacher purposes were, in ascending order, "promotion of teachers," "teacher transfers," and "dismissal of teachers." The only notable deviation from this pattern was the unusually low rating of eight awarded to "staff development and planning" when casual information from others was indicated as a method of evaluation.

Are there significant differences in the teachers' ratings of the overall evaluation process as a function of the following:

1. The number of hours perceived to have been spent per teacher in the evaluation process?
2. The number of times a teacher was observed during an evaluation period?
3. The perceived average length of the observations?

Respondents to item IV-A indicated their perceptions of the number of hours spent per teacher in the evaluation process by checking one of nine responses. Possible responses ranged from "0 to 1.0 hours"

TABLE VIII

CENTRAL OFFICE ADMINISTRATORS' RANKINGS OF PURPOSES OF TEACHER EVALUATION ACCORDING TO THE METHOD OF EVALUATION UTILIZED DURING THE 1978-79 SCHOOL YEAR

Purposes	Teacher Checklist	Critical Events	Input From Parents	Self-Evaluation	Evaluation by Objectives	Teacher Competency Tests	Student Test Data	Classroom Observations	Pre-Observation Conferences	Post-Observation Conferences	Instructional Conferences	Casual Info. from Teachers
1. Staff Development and Planning	(113) 3	(42) 3	(15) 3	(89) 3	(81) 3	(1) 8	(35) 3	(113) 3	(77) 3	(107) 3	(63) 3	(33) 3
2. Teacher Transfers	(12) 11	(6) 11	(2) 10.5	(5) 11	(5) 11	(9) 12	(7) 10	(12) 11	(9) 11	(9) 11	(9) 11	(12) 11
3. Provide Feedback to Teachers	(49) 6	(10) 9	(2) 10.5	(29) 6	(29) 7	(1) 8	(8) 8	(46) 6	(33) 6.5	(36) 7	(29) 6.5	(49) 5
4. Visit State Dept. Requirements	(87) 4	(20) 5	(9) 4.5	(56) 4	(41) 5	(2) 4.5	(16) 5	(73) 5	(55) 5	(54) 5	(51) 5	(27) 3
5. Dismissal of Teachers	(29) 10	(9) 10	(3) 8.5	(15) 10	(13) 10	(1) 8	(7) 10	(25) 10	(17) 10	(23) 10	(20) 10	(29) 10
6. Increase of Job Performance	(172) 2	(71) 2	(26) 1	(124) 2	(120) 2	(6) 1.5	(55) 2	(162) 2	(128) 2	(147) 2	(121) 2	(174) 2
7. Accounting to Authorities	(44) 8	(17) 6.5	(5) 7	(24) 8.5	(21) 9	(1) 8	(7) 10	(38) 8	(25) 8	(29) 9	(25) 8	(44) 7
8. Improvement of Instruction	(222) 1	(75) 1	(25) 2	(150) 1	(143) 1	(6) 1.5	(61) 1	(215) 1	(172) 1	(199) 1	(157) 1	(225) 1
9. Set Standards of Performance	(45) 7	(17) 6.5	(8) 6	(24) 8.5	(30) 6	(4) 8.5	(15) 6	(10) 7	(33) 6.5	(37) 6	(29) 6.5	(46) 6
10. Approval of Tenure	(84) 5	(29) 4	(9) 4.5	(50) 5	(59) 4	(2) 4.5	(21) 4	(89) 4	(59) 4	(68) 4	(59) 4	(85) 4
11. Improve Communication	(31) 9	(11) 8	(3) 8.5	(25) 7	(23) 8	(1) 8	(10) 7	(31) 9	(22) 9	(31) 8	(22) 9	(32) 9
12. Promotion of Teachers	(3) 12	(0) 12	(0) 12	(1) 12	(1) 12	(0) 12	(1) 12	(3) 12	(2) 12	(3) 12	(2) 12	(3) 12

Note: Frequencies of responses are provided in parentheses above each ranking.

to "16+ hours" in equal intervals of 1.9 hours. Regression analysis was performed to determine the existence of evidence pointing to a change in overall rating as a function of perceived hours spent in the evaluation process.

Item IV-C elicited responses concerning the number of times a teacher was observed during an evaluation period. Respondents were given 9 alternative responses from after "0" to "15+." Each response after "0" included a range of 2 observations, such as a "1 to 2" or "8 to 9." Responses were coded one through nine.

Item IV-D asked the respondent to indicate the perceived average length in minutes of an observation. There were 9 possible responses to this item ranging from "less than 10" to "more than 10" in 10 minute intervals. Responses were coded one through nine. In order to assess the impact of these three variables on the overall ratings given by teachers to their evaluation procedures, as well as the interaction of the variables, a multiple regression analysis was performed.

The analysis of variance performed in conjunction with the regression analysis resulted in an F value of 19.95, significant beyond .05. The t tests performed on the y intercept value and the slopes calculated for the three independent variables revealed that the intercept, the slope for the variable dealing with number of observations, and the slope of variable illustrating the length of an observation were significant beyond .05. The number of hours perceived to have been spent in the evaluation process calculated a t value of -1.41 which was not significant at .05. It was evident that, given the variables dealing with number and length of classroom observations, the

perceived number of hours spent in the evaluation process did not add significantly to the regression equation. A new regression equation using only the two significant variables was developed.

As in the previous analysis, the analysis of variance procedure produced a highly significant F value of 30.97. The regression analysis established a y intercept of 2.865. The slope of the portion of the regression equation dealing with the number of observations accompanying an evaluation was determined to be -0.114, and the slope of the portion dealing with the average length of an observation proved to be -0.172. The values calculated for the y intercept and the two slopes were found to be significant well beyond the established .05 level. The calculation of the Pearson Product Moment produced a coefficient of -.46. The probability of achieving a coefficient of this size by chance alone was determined to be 1 in 10,000, thus validating the existence of a highly significant relationship among the variables. In summary, it was determined that an increase in the perceived number of observations performed in conjunction with a teacher's evaluation correlated significantly with favorable perceptions of the evaluation process.

CHAPTER IV

SUMMARY, FINDINGS, CONCLUSIONS AND RECOMMENDATIONS

I. INTRODUCTION

The purpose of this chapter was to summarize the study and to report its findings. Major conclusions reached as a result of these findings were drawn. The final section listed some recommendations arising from the study.

II. SUMMARY

The Problem

In 1974 the Tennessee State Department of Education mandated the evaluation of teachers throughout the state. The lack of an adequate data base from which to assess the practice of teacher evaluation in Tennessee was the problem undertaken by this study.

The Purpose of the Study

This study had as its purpose the comparison and analysis of the descriptive data collected by Patricia Miller, concerning the perceptions of central office and building-level administrators toward teacher evaluation, and by Russell Ramsay, concerning the perceptions of teachers toward the facts of teacher evaluation explored by the same questionnaire used by Patricia Miller. Responses by randomly drawn members of each of the three segments mentioned revealed their perceptions of the purposes

of teacher evaluation, methods of evaluation, and evaluation priorities and permitted conclusions to be drawn.

The study addressed the following questions:

1. Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding:
 - a. the purposes of teacher evaluation?
 - b. the desirability of the teacher checklist, record of critical incidents, parental input, self-evaluation, evaluation by objectives, competency tests for teachers, student test data, classroom observation, and instructional conferences as methods of teacher evaluation?
 - c. the number of hours spent per teacher in the evaluation process?
 - d. the number of hours which should be spent in the teacher evaluation process?
 - e. the number of observations performed per teacher during an evaluation year?
 - f. the average length of an observation?
 - g. the number of times a teacher should be observed during an evaluation year?
 - h. the optimum average length of an observation?
 - i. the opportunity afforded the teacher to see and review his/her evaluation documents?

- j. the right of the teacher to make written comments or disagreements concerning his/her evaluations?
 - k. whether the teacher receives a copy of the final evaluation that is placed in the personnel file?
 - l. whether teachers are informed of the evaluation?
 - m. the involvement of teachers in the development of evaluation instruments?
 - n. the involvement of teachers in the development of evaluation procedures?
 - o. whether teachers are informed of the evaluation appeal process?
 - p. the ratings given to the overall evaluation processes?
2. Are there significant differences in the ratings of the overall evaluation procedures as a function of the teachers' involvement in the development of the procedure and/or their involvement in the development of evaluation instruments?
 3. Are there significant differences in the ratings of the overall evaluation procedures as a function of the extent of the teachers' participation in the implementation of the procedure?
 4. Within the three groups--teachers, building-level administrators, and central office administrators--are there significant differences in the perceived school purposes of evaluation as a function of the methods utilized in the procedure?

5. Are there significant differences in the teachers' ratings of the overall evaluation process as a function of:
 - a. the number of hours perceived to have been spent per teacher in the evaluation process?
 - b. the number of times a teacher was observed during an evaluation period?
 - c. the perceived average length of the observations?

Procedures

This study was a sequel to the descriptive studies of Patricia Miller and Russell Ramsay. Patricia Miller developed and utilized the "Teacher Evaluation Questionnaire" to gather descriptive data from some 314 central office administrators and 261 building-level administrators. Russell Ramsay utilized the same questionnaire to survey a random sample of Tennessee teachers. Usable returns were received from 259 respondents in the latter study.

The data from these two studies were synthesized to determine the existence of significant differences of the types expressed as research questions in the previous section of this chapter. Further analysis was conducted in an effort to gather evidence concerning the possible relationships between teachers' overall ratings of evaluation procedures and various indicators of participation in the development and implementation of these procedures.

In preparation for this study, an extensive review of the literature related to the theory and practice of teacher evaluation was

conducted. Textbooks, periodicals and research studies were surveyed with a particular focus on the history of teacher evaluation, purposes and criteria for evaluation, methods of evaluation, and perceptual factors affecting evaluation.

Treatment of the Data

Through the use of the Statistical Analysis System (SAS) the data from the two aforementioned studies were synthesized so that evidence concerning significant differences in the perceptions of the three groups of respondents could be gathered.

Differences in the perceptions of administrators and teachers concerning the school and personal purposes for teacher evaluation were analyzed through the Spearman Rank Order Correlation procedure. This procedure was also utilized to compare the perceived school purposes of each group as a function of the method of evaluation being employed.

The analysis of variance technique was used to determine the existence of significant differences among the three groups concerning the perceived desirability of various methods of evaluation and concerning those items within the section entitled "Evaluation Priorities" which deal with classroom observation and conferencing.

Significant differences among central office administrators, building-level administrators, and teachers with respect to the items in Section IV-G of the instrument were examined through the test for significant differences between proportions. This test, as well as the Spearman procedure, was calculated by hand by the researcher.

The "students t" test was employed to determine whether teachers who were involved in the development of evaluation instruments and/or procedures held more favorable perceptions about their evaluation processes than did those who had no involvement.

Teachers' overall ratings of the evaluation process as a function of the number of hours perceived to have been spent in the evaluation process were examined through regression analysis. This procedure was also employed with respect to teachers' overall ratings as a function of perceptions of the number of times they were observed during an evaluation period.

III. FINDINGS

Findings will be stated as they relate to the various research questions addressed within the study. Major questions are addressed in the same numerical order in which they appear throughout the study.

Question 1-a.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding the purposes of teacher evaluation?

A Spearman Correlation Coefficient of .944 between the rankings of central office and building-level administrators was significant at .05 and gave evidence that the perceived school purposes of the two groups were highly similar. The correlation coefficient between teachers and central office administrators (.538) proved not to be significant and attested to the independence of their perceptions of the purposes

of teacher evaluation. Although the coefficient between teachers and building-level administrators (.769) was significant at .05, it indicated considerably less agreement between the groups than was found between central office and building-level administrators. Both central office and building-level administrators perceived the primary school purpose for teacher evaluation to be instructional improvement. This was marked contrast to the perception of teachers that the primary purpose for their school system's engaging in teacher evaluation was to meet State Department requirements.

The trend established by these first responses continued through the rankings. Administrators perceived the most important school purposes for evaluation to be those related to instructional, personal, and professional improvement while teachers maintained that they were evaluated primarily for administrative reasons such as approval of tenure and accounting to authorities.

Although noteworthy disparities do appear in the perceived school purposes for teacher evaluation, almost no difference is apparent in the rankings of the three groups regarding purposes of evaluation personally thought to be most important. Teachers, central office administrators and building-level administrators agreed that the two primary purposes, at the personal level, were improvement of instruction and increasing job performance.

In summary, evidence points to a consensus among teachers, building-level administrators and central office administrators that teacher evaluation should be performed primarily to facilitate personal

and professional growth. A significant disparity existed among the three groups, however, as to the purposes for which evaluation was then being implemented. The two administrator groups tended to perceive that their present evaluation procedures pursued the purposes related to instructional development while teachers saw the process as being primarily administrative, meeting authoritative mandates and assisting administrative decision-making.

Question 1-b.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding the desirability of the teacher checklist, record of critical incidents, parental input, self-evaluation, evaluation by objectives, competency tests for teachers, student test data, classroom observation, and instructional conferences as methods of teacher evaluation?

Teacher Checklist. Of the six sub-categories dealing with the teacher checklist, the area of professional qualities (1.28) was rated most desirable by all respondents. In descending order, all respondents rated attitudes and values (1.42), professional training and experience (1.46), personal characteristics (1.67), and others (1.63) as being somewhat desirable. The use of physical characteristics in a teacher checklist was given a rating slightly lower than desirable (2.09).

Record of Critical Incidents or Events. Although all these groups of respondents gave this evaluative tool a rating slightly below "desirable," teachers (2.40) rated its use significantly lower than did administrators (2.10).

Input from Parents. Without significant disagreement, respondents perceived the use of parental input as less than desirable. The mean rating for this category was 2.51.

Self-Evaluation. Teachers, building-level administrators, and central office administrators agreed that self-evaluation was a desirable means of teacher evaluation (1.59).

Evaluation by Objectives. Job targets identified by the teacher and the administrator received desirable ratings from all respondents, however, central office administrators rated this type of job target significantly higher (1.46) than did either building-level administrators (1.58) or teachers (1.66). Job targets identified by the teachers only were rated as moderately desirable by teachers (1.76). This rating was significantly higher than the level of desirability expressed by central office administrators (1.99) and building-level administrators (2.08). Job targets identified by administrators only were considered somewhat less than desirable by all respondents; however, teachers thought significantly less of this type of job target (2.96) than did central office (2.66) and building-level administrators (2.68).

Building-level administrators rated the use of job targets based on a needs assessment (2.15) and needs assessment based on a job description (2.28) significantly lower than the central office administrators (2.00) or teachers (1.89, 1.92).

Competency Tests for Teachers. Although the use of competency tests in teacher evaluation did not receive a favorable rating from any segment of the respondents, the building-level administrators (2.91)

perceived the practice to be significantly less desirable than did central office administrators (2.59) or teachers (2.64).

Student Test Data. With the exception of standardized tests, which teachers (2.80) perceived to be significantly less desirable than did central office (2.59) and building-level administrators (2.61), no significant differences were encountered. Respondents agreed that proficiency tests (2.66); teacher-made tests (2.52); student attitude measures, both paper and pencil (2.71) and observations (2.66); and student reports of teacher behavior (2.80) were relatively undesirable as teacher evaluation tools.

Classroom Observations. Of the six sub-categories of potential observers listed, all respondents found observations of principals most desirable. As a group, however, teachers (1.53) gave a significantly lower rating to these observations than did central office (1.23) and building-level administrators (1.69). Central office administrators (2.37) considered classroom observations by superintendents to be significantly more desirable than did building-level administrators (2.54) or teachers (2.52). Observations by department chairman or team leader (2.36), peers (2.63), and others (2.56) were perceived with a similar degree of undesirability by all respondents.

Pre-Observation Conferences. Teachers' perceptions of the desirability of pre-observation conferences (1.61) by the principal were significantly lower than those of central office and building-level administrators (1.39). Building-level administrators' (1.86) and teachers' (1.86) perceptions of pre-observation conferences with supervisors provided ratings not dissimilar from each other, but significantly

lower than the mean response of the central office group (1.71). Respondents were in relative agreement concerning pre-observation conferences by superintendents (2.54), department chairman or team leader (2.43), peers (2.70), and others (2.72).

Post-Observation Conferences. In each of the six sub-categories the three sample factions awarded a slightly higher level of desirability to post-conferences than to pre-conferences. Teachers (1.47) continued to rate post-conferences with principals at a significantly lower level than did central office or building-level administrators, whose mean responses were an almost identical 1.26.

As in the case of pre-observation conferences, central office administrators (1.60) rated post-observation conferences with supervisors significantly higher than teachers (1.79) and building-level administrators (1.71). The remaining categories were given mean ratings which indicated a relatively low level of desirability by all respondents without significant disagreement (superintendents--2.46, department chairman or team leader--2.40, peers--2.67, others--2.68).

Instructional Conferences. Teachers, central office administrators and building-level administrators on this perceived desirability of the three types of instructional conferences. While all types of instructional conferences were perceived to be desirable, planning conferences posted the most favorable mean rating (1.36) followed by formative conferences (1.40) and summative conferences (1.45).

Casual Information from Others. The use of casual information in teacher evaluation was viewed with a relative lack of enthusiasm by

all respondents; however, central office administrators (2.57) perceived the practice to be significantly less desirable than did building-level administrators (2.38). Neither group of administrators differed significantly from the teacher group (2.43).

Question 1-c.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding the number of hours spent in the evaluative process?

A mean response of 1.59 by teachers indicated that they perceived between 0 and 3.9 hours being spent in the yearly evaluation process. This perception was the lowest of any portion of the sample.

The second lowest perception of time spent in the evaluation process was recorded by the central office group at 3.26. This response marked their perception of between 2 and 5.9 hours.

Building-level administrators, the group not often said to have primary responsibility for teacher evaluation, tabulated a mean response of 3.28. This was the highest of the three sub-groups and translated to a range of 4 to 7.9 hours. Each response was significantly different from the remaining two groups' responses, indicating a marked divergence of perceptions in this area.

Question 1-d.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding the number of hours which should be spent in the evaluation process?

All three groups of respondents again were significantly different in their perceptions. Building-level administrators (4.82) asserted that the optimum amount of time needed to evaluate a teacher was between 6 and 9.9 hours. Central office administrators (4.02) set the optimum time range at approximately 6 to 7.9 hours. Teachers (3.12) felt that from 4 to 5.9 hours was sufficient to perform the evaluation function adequately.

Question 1-e.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding the number of observations performed per teacher during an evaluation year?

A significant discrepancy was revealed among all of the factions identified in this study. Teachers (2.63) perceived that they were observed from 1 to 4 times during an evaluation year while central office administrators (3.16) believed that approximately 3 to 4 observations were performed and building-level administrators (3.96) reported that some 5 to 6 observations accompanied an evaluation.

Question 1-f.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding the average length of an observation?

Teachers (2.62) perceived each of the observations performed during an evaluation period to be approximately 16.2 minutes. This

perception proved to be significantly different from the two administrator groups, who were in relative agreement. Building-level administrators (4.14) perceived an average observation length of some 31.4 minutes and central office administrators (4.36) estimated approximately 33.6 minutes per observation.

Question 1-g.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding the number of times a teacher should be observed during an evaluation year?

Teachers (3.71) established that the number of observations conducted during an evaluation period should range between 3 and 6, central office administrators (4.13) expanded this range to between 5 and 8 observations, and building-level administrators (4.90) set a figure approaching 7 to 8 observations. Significant deviations were evident among all three segments of the sample.

Question 1-h.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding the optimum length of an observation?

Central office administrators (5.05) set the optimum length of an observation at approximately 40.5 minutes. Building-level administrators (4.75) perceived some 37.5 minutes to be a proper length. According to the teacher segment of the sample (3.84),

observations should average about 28.4 minutes. Each faction posted a mean response which was significantly different from the other two groups.

Question 1-i.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding the opportunity afforded the teacher to see and review his/her evaluation documents?

A significantly smaller proportion of teachers (.574) reported that they always had the opportunity to see and review their evaluation documents than did central office (.857) and building-level administrators (.889). Conversely, a larger proportion (.359) of teachers than either central office (.134) or building-level administrators (.103) stated that teachers could see and review their evaluation documents if they wished. In both cases the perceptions of the two administrator groups did not differ significantly.

Question 1-j.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding the right of the teacher to make written comments or disagreements concerning his/her evaluations?

Approximately 57.1 percent of the teachers surveyed perceived that they had the right to make written comments or disagreements on their evaluations. This percentage was significantly lower than the percentages of building-level (93.9) and central office administrators (94.9) who responded in a like manner.

Question 1-k.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding whether the teacher receives a copy of the final evaluation that is placed in the personnel file?

Although central office and building-level administrators (70.1 percent) agreed that teachers did receive a copy of their final evaluation, only 38.6 percent of the teachers surveyed corroborated this perception. A significant discrepancy between teachers and administrators was revealed concerning this point.

Question 1-l.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding whether teachers are informed of the evaluation process?

As in the previous four questions, the proportion of teachers (.753) who were informed of the evaluation process differed significantly from central office (.984) and building-level administrators (.966). Despite the fact that a higher number of teachers responded affirmatively to this item than to any other under section IV-G of the instrument, it is worthy of note that almost one-fourth of the teacher respondents did not perceive themselves to have been informed of this critical process.

Question 1-m.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding the involvement of teachers in the development of evaluation instruments?

Each of the sample segments had significantly different perceptions of teacher involvement in this area. While 81.8 percent of the central office administrators attested to teacher involvement, 71.3 percent of the building-level administrators concurred, and only 22.9 percent of the teachers shared this perception.

Question 1-n.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding the involvement of teachers in the development of evaluation procedures?

Approximately 77.7 percent of the central office administrators affirmed the involvement of teachers in developing evaluation procedures. A smaller 69.7 percent of the building-level administrators were in concert with their central office colleagues. Less than one-fourth of the teachers (.202) were able to respond positively to this item. The mean responses of each group were significantly deviant from each of the remaining groups.

Question 1-o.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding whether teachers are informed of the evaluation appeal process?

A significantly lower proportion of teachers (.349) confirmed their appraisal of the evaluation appeal process than did central office (.866) and building-level administrators (.866).

Question 1-p.

Are there significant differences in the perceptions of teachers, building-level administrators, and central office administrators regarding the ratings given to the overall evaluation process?

Teachers posted a mean overall rating of 2.13, or slightly below "fair." This rating differed significantly from both sets of administrators who rated their evaluation processes as 1.67, or between "good" and "fair."

Question 2

Are there significant differences in the ratings of the overall evaluation procedures as a function of the teachers' involvement in the development of the procedure and/or their involvement in the development of evaluation instruments?

The teachers who indicated that they were involved in the development of their evaluation procedure rated the overall evaluation process at a mean of 1.57 while those teachers expressing no involvement recorded a significantly lower overall rating of 2.26. Likewise, the mean overall rating of teachers who were involved in the development of their evaluation instrument (1.73) was significantly more favorable than their counterparts who had no involvement (2.33). It should be noted that only 51 teachers, or 20 percent of the total teacher respondents, expressed involvement in developing their procedure and 88 teachers, or 35 percent, verified involvement in developing their evaluation instruments.

Question 3

Are there significant differences in the ratings of the overall evaluation process as a function of the extent of the teachers' participation in the implementation of the process?

The regression analysis performed in connection with this question gave evidence of the possibility that teachers' overall ratings of their evaluation process were affected positively by their degree of participation in the implementation of the process. The lack of control for other variables which could affect ratings felt some doubt as to the extent of this relationship; however, the extremely high significance of the F ratio and the t tests of the slope and intercept cannot be ignored.

Question 4

Within the three groups--teachers, building-level administrators, and central office administrators--are there significant differences in the perceived school purposes of evaluation as a function of the methods utilized in the procedure?

Within each group, the rankings of school purposes by methods of evaluation were compared with each other and with the overall rankings for that group. No rankings were found to be independent from each other or from the overall ranking for that group; therefore, evidence points to a lack of independence of perceived school purposes as a function of the method employed in evaluating teachers.

Question 5-a.

Are there significant differences in the teachers' ratings of the overall evaluation process as a function of the number of hours perceived to have been spent per teacher in the evaluation process?

The multiple regression analysis performed pursuant to this question revealed that the number of hours perceived to have been spent in the evaluation process did not have a significant impact on teachers' overall rating of the evaluation procedure viewed in conjunction with the variables concerned with the number and length of classroom observations.

Question 5-b.

Are there significant differences in the teachers' ratings of the overall evaluation process as a function of the number of times a teacher was observed during an evaluation period?

The high degree of significance achieved in the analysis of variance procedure and the t tests of the slope and intercept of the regression equation point to the possibility of an increasingly positive rating of the overall evaluation process as a function of increased numbers of observations of the teacher to be evaluated.

Question 5-c.

Are there significant differences in the teachers' ratings of the overall evaluation process as a function of the average length of an observation?

The high level of significance achieved in the analysis of variance procedure and in the t tests of the slope and intercept of the

regression equation speak positively to the existence of an increase in teachers' overall rating of their evaluation procedure as a function of an increase in the perceived average length of an observation.

Pearson Correlation Coefficient

A Pearson Correlation Coefficient of $-.46$ was found to be significant at $.0001$. This coefficient gives further evidence of the positive impact of increased incidence and length of classroom observations on teachers' perceptions of their evaluation processes.

Summary of Findings

Based on the review of related literature and the data utilized, the findings of this study are summarized as follows:

1. The purposes for which teacher evaluation was being conducted in Tennessee were perceived in a distinctly different manner by administrators and teachers.
2. Significant differences existed among teachers, building-level administrators and central office administrators as to the desirability of the various methods of teacher evaluation explored by the "Teacher Evaluation Questionnaire."
3. There was significant disagreement among teachers, building-level administrators and central office administrators regarding:
 - a. the number of hours spent per teacher in the evaluation process,
 - b. the number of hours per teacher which should be spent in the evaluation process,

- c. the number of observations performed per teacher and the number of observations which should be performed during an evaluation year, and
 - d. the present average length of an observation and the optimum length of an observation.
4. A significant discrepancy exists between the perceptions of teachers and administrators as to the degree of teacher involvement in the development and implementation of the evaluation process as explored in section G of the "Teacher Evaluation Questionnaire."
 5. Teachers perceive their overall evaluation process to be significantly less acceptable than do administrators.
 6. Evidence pointed to a more positive perception of the overall evaluation process by teachers as a function of their participation in the development of the evaluation procedures and instruments.
 7. Evidence indicated a significant positive correlation between teachers' perceptions of their overall evaluation process and the extent of their participation in the implementation of the process.
 8. At this point in time, the method utilized in the evaluation did not influence Tennessee teachers' perceptions of the purposes for teacher evaluation.
 9. Evidence pointed to a significant positive correlation between perceptions of the overall evaluation process and the increased incidence and length of classroom observations.

IV. CONCLUSIONS

The findings of the study present evidence to suggest certain conclusions. Based on the survey of the literature and the data analyzed in this study the following conclusions were reached:

1. There was a lack of mutual agreement as to the purposes for which teacher evaluation was being conducted. Since no significant disagreement was evident in the ideal purposes for evaluation, it would seem that there is either a lack of adequate involvement of teachers in the evaluation process or that the process is being implemented by administrators in such a manner that the purposes dealing with instructional improvement are not evident to teachers.
2. The teacher checklist, the most traditional form of teacher evaluation, was the only method which was deemed mutually appropriate. This being the case, any system choosing to move to a more current model should be prepared to do extensive staff development in order to assist in the transition.
3. The lack of mutual agreement concerning the extent of administrator involvement in the evaluation process gives evidence that the evaluative behavior exhibited by administrators is construed as perfunctory.
4. Although the various evaluation processes may provide avenues for teacher involvement, a lack of involvement in the development of evaluation processes and instrumentation

contributes to a perceived lack of involvement in the implementation of the process.

5. Teacher involvement in the development and implementation of the evaluation process contributes to overall leader satisfaction with the process.
6. The more intense the involvement of the administrator in the evaluation process, particularly in the areas of classroom observation, the greater the level of overall teacher satisfaction with the process.

V. RECOMMENDATIONS

The conclusions reached as a result of this study have generated several recommendations for administrators in the field and for further research in the areas of teacher evaluation.

1. Central office and building-level administrators must analyze their own evaluation practices in an effort to determine what factors are responsible for the present teacher perception of evaluation primarily as a perfunctory administrative task.
2. If teacher evaluation is to be a positive force for instructional improvement in Tennessee, administrators and teachers must achieve a more positive perception of the process through:
 - a. cooperative assessment of present evaluation practices;

- b. effective interaction among those involved in the evaluation process concerning evaluation policies, procedures, and instrumentation;
 - c. increased teacher participation in the development of teacher evaluation instruments and procedures; and
 - d. increased teacher involvement in the implementation of the evaluation process.
- 3. The State Department of Education should examine evaluation practices of other school systems within and outside the State of Tennessee to determine the state of the art and to provide current information to the various school systems.
- 4. The following questions arising from this study have implications for further research:
 - a. What factors account for the apparent division of perceptions of purposes for teacher evaluation between administrators and teachers?
 - b. How may the value of teacher-participation in the evaluation process be more clearly defined?
 - c. What methods of evaluation allow for the greatest teacher involvement and, at the same time, provide the greatest potential for instructional improvement?

LIST OF REFERENCES

LIST OF REFERENCES

1. Victor Vroom, Work and Motivation (New York: Wiley Publishing Company, 1964).
2. Superintendents' Study Council Task Force, Developing Policies, Criteria, and Procedures for Evaluating Professional Personnel (Tennessee State Department of Education, 1974), pp. 1-2.
3. All references made in this study to the Miller and Ramsay studies or to the Teacher Evaluation Questionnaire refer to Patricia Miller, "A Survey of the Design and Implementation of Teacher Evaluation Systems in Tennessee Public Schools," (Ed.D. Dissertation, The University of Tennessee, Knoxville, June, 1980); Russell Ramsay, "Teacher Perceptions of Teacher Evaluation Systems in Tennessee Public Schools," (Ed.D. Dissertation, The University of Tennessee, Knoxville, August, 1980).
4. Tennessee State Department of Education, Rules, Regulations, and Minimum Standards (Nashville: State Department of Education, 1974), Part II, Section D, pp. 72-73.
5. Select Committee for the Study of Teacher Training, Summary Report of Hearings (Fall, 1979).
6. Superintendents' Study Council Task Force, pp. 1-2.
7. Dale L. Bolton, Selection and Evaluation of Teachers (Berkeley, California: McCutchan Publishing Corp., 1973), p. 22.
8. W. James Popham, Educational Evaluation (Englewood Cliffs, New Jersey: Prentice-Hall, 1975), p. 1.
9. Glen G. Eye, "Supervisory Skills: The Evaluation of the Art," Journal of Educational Research (Summer, 1976), p. 14.
10. John Whithall and Fred H. Wood, "Taking the Threat Out of Classroom Observation and Feedback," Journal of Teacher Education, XXX (January-February, 1979), p. 55.
11. Eye, p. 14.
12. Eye, p. 15.
13. Hazel Davis, "Evaluation of Current Practices in Evaluating Teacher Competence," in Bruce J. Biddle and William J. Ellena (ed.) Contemporary Research on Teacher Effectiveness (New York: Holt, Rinehart and Winston, 1964), p. 43.

14. B. Othaniel Smith, "Teaching: Conditions of its Evaluation," The Evaluation of Teaching (Washington, D. C.: Phi Lamda Theta, 1967), p. 66.
15. Dwight E. Beecher, The Evaluation of Teaching: Backgrounds and Concepts (New York: Syracuse University Press, 1949), p. 5.
16. Beecher, p. 5.
17. Davis, p. 44.
18. A. S. Barr, Characteristic Differences in Good and Poor Teachers (Illinois: Public School Publishing Company, 1929), p. 42.
19. Beecher, p. 7.
20. F. B. Knight, Qualities Relating to Success in Teaching (New York: Bureau of Publications, Teachers College, Columbia University, 1922), p. 45.
21. Barr, pp. 49-50.
22. Beecher, p. 19.
23. Beecher, p. 20.
24. Beecher, p. 21.
25. Bruce J. Biddle and William J. Ellena, ed., Contemporary Research on Teacher Effectiveness (New York: Holt, Rinehart, and Winston, 1964), pp. 52-54.
26. James Lewis, Jr., Appraising Teacher Performance (West Nyack, New York: Parker Publishing Co., 1973), p. 11.
27. Popham, p. 3.
28. Gary D. Borich, The Appraisal of Teaching: Concepts and Process (Reading, Massachusetts: Addison-Wesley, 1977), p. vii.
29. Ronald T. Hyman, School Administrator's Handbook of Teacher Supervision and Evaluation Methods (New Jersey: Prentice-Hall, 1975), p. 8.
30. Madeline Hunter, "Appraising Teacher Performance: One Approach," National Elementary Principal, LII (February, 1973), pp. 60-63.
31. Henry Mintzberg, The Structuring of Organizations (New Jersey: Prentice-Hall, 1979), pp. 151-152.

32. Benjy Levin, "Teacher Evaluation - A Review of Research," Educational Leadership, 37 (December, 1979), p. 240.
33. Thomas J. Sergiovanni and Robert J. Starratt, Supervision: Human Perspectives (New York: McGraw-Hill, 1979), pp. 305-309.
34. John Frank, Jr., School Principal's Handbook of Evaluation Guidelines (New York: Parker Publishing Company, 1979), p. 14.
35. National School Public Relations Association, p. 42.
36. Gladys Styles and Carol Camp, "The Supervision of Teacher Evaluation: A Brief Overview," Journal of Teacher Education (March - April, 1979), p. 17.
37. George G. Redfern, How to Appraise Teaching Performance (Columbus, Ohio: School Management Institute, 1963), p. 25.
38. William B. Castetter, The Personnel Function in Educational Administration (New York: Macmillan, 1976), p. 232.
39. Wayne Cascio, Applied Psychology in Personnel Management (Reston, Virginia: Reston Publishing Company, 1978), p. 316.
40. Lyman W. Porter, Edward E. Lawler, III, and J. Richard Hackman, Behavior in Organizations (New York: McGraw-Hill, 1975), p. 317.
41. Bolton, p. 21.
42. "Negotiating the Teacher Evaluation Issue," ERS Negotiation Aid for School Management (Arlington, Virginia: Educational Research Service, 1979), p. 2.
43. Robert Howsam, "Teacher Evaluation: Facts and Folklore," National Elementary Principal, LIII (November, 1963), p. 13.
44. Mel J. Zelenak and Bill C. Snider, "Teachers Don't Resent Evaluation if it's for the Improvement of Instruction," Phi Delta Kappa, 55 (April, 1974), p. 571.
45. Henry J. McNally, "What Makes a Good Evaluation Program?" National Elementary Principal, 52 (February, 1973), pp. 2-9.
46. John McNeil and W. James Popham, "The Assessment of Teacher Competence," in Robert Travers, editor, Second Handbook of Research on Teaching (Chicago: Rand-McNally, 1973), p. 221.
47. Styles and Camp, p. 22.

48. Peter W. Arasian, "Performance-Based Teacher Education: Evaluation Papers," (New Orleans: Multi-State Consortium on Performance-Based Teacher Education, February 26, 1973), ED 102200.
49. S. K. Stemnock, Evaluating Teaching Performance, Educational Research Service, Circular 33 (Washington, D. C.: National Education Association, 1969).
50. Bolton, p. 102.
51. David G. Ryans, "Teacher Behavior Can Be Evaluated," The Evaluation of Teaching (Washington, D. C.: Pi Lambda Theta, 1967), p. 57.
52. Smith, p. 84.
53. Linda Daniels and Margaret Phelps, "Teacher Evaluation in Tennessee," Tennessee Education, 9 (February, 1979), p. 17.
54. Rita Dunn and Kenneth J. Dunn, Administrator's Guide to New Programs for Faculty Management and Evaluation (West Nyack, New York: Parker Publishing Company, 1977), p. 163.
55. Howsam, p. 15.
56. Nicholas A. Fattu, "Research on Teacher Evaluation," National Elementary Principal, LII (February, 1973), p. 58.
57. John McNeil and W. James Popham, "The Assessment of Teacher Competence," Second Handbook of Research on Teaching, Robert Travers, editor (Chicago: Rand-McNally, 1973), p. 221.
58. W. James Popham, "Teaching Performance Tests," National Elementary Principal, LII (February, 1973), p. 47.
59. McNeil and Popham, p. 239.
60. J. Merrell Hansen, "The Evaluation of Teaching: No Guppies or Goldfish in My Classroom," NASSP Bulletin, 62 (March, 1978), p. 12.
61. Borich, p. ix.
62. Gordon Lawrence, "Delineating and Measuring Professional Competencies," Educational Leadership, XXXI (January, 1974), pp. 301-302.
63. Bolton, p. 102.
64. Sergiovanni and Starratt, p. 308.
65. McNally, p. 23.

66. Levin, p. 240.
67. McNally, p. 26.
68. Hyman, p. 16.
69. Castetter, p. 232.
70. C. R. Ingils, "Let's Do Away With Teacher Evaluation," The Clearing House 44 (November, 1970), p. 456.
71. James R. Lewis, Appraising Teacher Performance (West Nyack, New York: Parker Publishing Company, 1973), p. 25.
72. A. S. Barr, et al., Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness (Madison, Wisconsin: Dembar Publications, 1961), p. 8.
73. Barr, et al., p. 9.
74. Bolton, p. 36.
75. James Trent, Arthur Cohen, and Clare Rose, "Evaluation of Teaching in Research on Teaching in Higher Education," Second Handbook of Research on Teaching, Robert M. W. Travers, editor (Chicago: Rand-McNally, 1973), p. 1042.
76. McNeil and Popham, p. 232.
77. Cascio, p. 321.
78. Cascio, p. 329.
79. Donald Thomas, "The Principal and Teacher Evaluation," NASSP Bulletin, LXIII (December, 1974), p. 4.
80. Richard P. Mannatt, et al., "Evaluating Teacher Performance With Disproved Rating Scales," NASSP Bulletin, 60 (No. 401, 1976), pp. 21-25.
81. Francis Griffith, A Handbook for the Observation of Teaching and Learning (Midland, Michigan: Pendell Publishing Company, 1973), p. 54.
82. National School Public Relations Association, p. 18.
83. Lewis, p. 33.
84. Kimball Wiles and John T. Lovell, Supervision for Better Schools (New Jersey: Prentice-Hall, 1975), p. 242.

85. Sergiovanni and Starratt, p. 308.
86. Davis, p. 48.
87. Cascio, pp. 326-327.
88. Levin, p. 242.
89. Bolton, p. 38.
90. McNeil and Popham, p. 231.
91. Levin, p. 243.
92. Robert Wolf, "How Teachers Feel Toward Evaluation," in School Evaluation, Earnest Haus, editor (Berkeley, California: McCutchan, 1973).
93. William J. Reddin, Effective Management By Objectives (New York: McGraw-Hill, 1971), p. 23.
94. George S. Odiorne, Management by Objectives: A System of Managerial Leadership (New York: Pitman Publishing Company, 1965), p. 19.
95. Reddin, p. 12.
96. Lewis, p. 33.
97. James R. Lewis, School Management by Objectives (New York: Parker Publishing Company, 1974), pp. 54-55.
98. Bolton, p. 138.
99. Redfern, pp. 1-99.
100. National School Public Relations Association, p. 17.
101. Frank Craig and Margaret Burns, "Does Management by Objectives Work in Education?" Educational Leadership, 36 (March, 1979), p. 415.
102. Bolton, p. 37.
103. Barr, et al., p. 8.
104. W. J. Popham, "The Performance Test: A New Approach to the Assessment of Teaching Proficiency," Journal of Teacher Education, 19 (1968), p. 218.

105. Gene V. Glass, "Teacher Effectiveness," in Evaluating Educational Performance, Herbert J. Wolberg, editor (Berkeley, California: McCutchan, 1974), p. 11.
106. Jason Millman, "Program Assessment, Criterion Referenced Tests, and Things Like That," Educational Horizons, 52, No. 4 (1974), pp. 188-189.
107. Donald L. Haefele, "How to Evaluate Thee, Teacher--Let Me Count the Ways," Phi Delta Kappan, 61 (January, 1980), p. 349.
108. Haefele, p. 349.
109. W. James Popham, "Educational Measurement for the Improvement of Instruction," Phi Delta Kappan, 61 (April, 1980), p. 532.
110. Levin, p. 242.
111. Shirley B. Stow, "Using Effectiveness Research in Teacher Evaluation," Educational Leadership, 37 (October, 1979), p. 58.
112. Glen F. Ovard, "Teacher Effectiveness and Accountability," NASSP Bulletin, 59 (January, 1975), pp. 91-93.
113. Harry E. Eastridge, "Student Evaluation and Teacher Performance," NASSP Bulletin, 60 (September, 1976), pp. 49-54.
114. Levin, p. 241.
115. Haefele, p. 351.
116. J. Jackson Barnette and John C. Thompson, "A Descriptive Assessment of the Effect of Evaluation on Instruction," Studies in Educational Evaluation, No. 5 (1979), p. 86.
117. McNeil and Popham, p. 232.
118. Gerald Becker, et al., "Methods of Evaluating Teachers," Elementary Principals and Their Schools (Eugene, Oregon: Center for the Advanced Study of Educational Administration, 1971), pp. 42-44.
119. John R. Hoyle, "Teacher Versus Administrator: A Growing Crisis," Planning and Changing, 9 (Winter, 1978), pp. 23-24.
120. Sergiovanni and Starratt, p. 306.
121. National Education Association, Research Division, "Evaluation of Classroom Teachers," Research Report - R 14 (Washington, D. C.: National Education Association, 1964), p. 19.

122. Stemnock, p. 18.
123. Haefele, p. 350.
124. David McAfee, "Evaluation of the Teacher: Do Teachers and Supervisors Agree?" High School Journal, 58 (May, 1976), p. 337.
125. National Education Association, p. 23.
126. Charles Reavis, "Clinical Supervision: A Review of Research," Educational Leadership, 35, No. 7 (1978), pp. 580-584.
127. John Robinson, "The Observation Report: A Help or a Nuisance?" NASSP Bulletin, 62 (1978), pp. 22-26.
128. Barnette and Thompson, p. 81.
129. Terrence E. Deal and Lynn D. Celotti, "How Much Influence Do (and Can) Educational Administrators Have on Classrooms?" Phi Delta Kappan, 61 (March, 1980), p. 473.
130. John R. P. French, Jr., Emmanuel Kay, and Herbert H. Meyer, "Participation and the Appraisal System," Human Relations, 19 (February, 1966), pp. 3-19.
131. Popham, "The Performance Tests: A New Approach to the Assessment of Teaching Performance," p. 220.
132. Wolf, p. 14.
133. Larry E. Wicks, "Opinions Differ: Teacher Evaluation," Today's Education, 62 (March, 1973), pp. 42-43.
134. Eric Berger, "The Evaluation of Teachers," NASSP Bulletin, 58 (March, 1974), p. 147.
135. John H. Bushman, "Are Teachers Playing 'Statue' in the Classroom?" NASSP Bulletin, 58 (December, 1974), p. 27.
136. Charles Jerrold Salek, "Helping Teachers versus Evaluating Teachers," NASSP Bulletin, 60 (September, 1976), p. 35.
137. Lawrence E. Kult, "Improving Teacher Evaluation by Principals," Clearing House, 52 (September, 1978), p. 21.
138. W. James Popham, "Pitfalls and Pratfalls of Teacher Evaluation," Educational Leadership, 32 (November, 1974), p. 144.

139. E. M. Bridges, "Preferences of Principals for Instrumental and Expressive Characteristics of Teachers Related to System Type." Paper presented at the annual meeting of the American Educational Research Association, Chicago (February, 1968).
140. Robert G. Owens, Organizational Behavior in Schools (New Jersey: Prentice-Hall, 1970), p. 73.
141. Bolton, p. 102.
142. Lewis, Teacher Performance, p. 33.
143. Porter, Lawler, and Hackman, pp. 316-319.
144. Dean Tjosvold, "Cooperation and Conflict Between Administrators and Teachers," Journal of Research and Development in Education, 12, No. 1 (1978), p. 139.
145. Cascio, p. 319.
146. Warren E. Tobin, "A Principal Looks at Evaluation," National Elementary Principal, 52 (February, 1973), p. 95.
147. Frank, p. 8.
148. McNally, p. 29.
149. French, Kay, and Meyer, pp. 11-18.
150. Lorraine Poliakoff, "Recent Trends in Evaluating School Personnel," National Elementary Principal, 52 (February, 1973), pp. 42-43.
151. Whithall and Wood, pp. 55-58.

APPENDIX

Cd 1 Code _____
(1) (2-6)

TEACHER EVALUATION QUESTIONNAIRE

This questionnaire is designed to gather information regarding your attitudes and perceptions toward teacher evaluation.

I. RESPONDENT DATA

- A. Check your position.
- | | | |
|------------------------|------------------------------|-----|
| 1 _____ Superintendent | 4 _____ Principal | (7) |
| 2 _____ Supervisor | 5 _____ Asst. Principal | |
| 3 _____ Other, _____ | 6 _____ Curriculum Principal | |
| (specify) | 7 _____ Other, _____ | |
| | (specify) | |
- B. Check the category which best describes the school level(s) in which you have responsibility:
- | | | | |
|--------------------|----------------------------|----------------------|-----|
| 1 _____ Primary | 4 _____ Elementary-Middle | 7 _____ Senior High | (8) |
| 2 _____ Elementary | 5 _____ Junior High | 8 _____ All levels | |
| 3 _____ Middle | 6 _____ Junior-Senior High | 9 _____ Other, _____ | |
| | | (specify) | |
- C. Check the size of the system in which you work:
- | | |
|-----------------------------------|-----|
| 1 _____ Less than 5,000 students | (9) |
| 2 _____ 5,000-10,000 students | |
| 3 _____ More than 10,000 students | |
- D. Length of time you have worked in the system (not counting current school year): _____ (10-11)
- Length of time you have worked in your present position (not counting current school year): _____ (12-13)

II. SCHOOL SYSTEM INFORMATION

- A. On the left indicate what you perceive to be your school system's three most important purposes of teacher evaluation.
- B. On the right indicate what you personally consider the three most important purposes of teacher evaluation.
- (1 - primary importance; 2 - second in importance; 3 - third in importance)

	School	Myself	
(14)		1. Staff development and planning	(27)
(15)		2. Teacher transfers	(28)
(16)		3. Provide feedback to teachers	(29)
(17)		4. Meet State Department requirements	(30)
(18) Number		5. Dismissal of teachers	(31)
(19)		6. Increase job performance	(32)
(20) only		7. Accounting to authorities	(33)
(21)		8. Improvements of instruction	(34)
(22) three		9. Set standards of performance	(35)
(23)		10. Approval for tenure	(36)
(24)		11. Improve communication	(37)
(25)		12. Promotion of teachers	(38)
(26)		13. Others, _____	(39)
		specify	

- C. Rank these functions of a principal in their order of importance. (1) = most important to (8) = least important
- | | |
|---|---|
| (40) 1 _____ School Community Relations | (41) 5 _____ School Business Management |
| (41) 2 _____ Staff Development | (42) 6 _____ Curriculum Development |
| (42) 3 _____ Teacher Evaluation | (43) 7 _____ School Plant Management |
| (43) 4 _____ Student Personnel Work | (44) 8 _____ Transportation |
- D. Who is involved in the evaluation of teachers in your school system? (check all appropriate responses)
- | | | |
|--------------------------------|---------------------------------------|------------------------------------|
| (48) 1 _____ Principal | (52) 5 _____ Asst. Principal | (56) 9 _____ Teachers |
| (49) 2 _____ Personnel officer | (53) 6 _____ Curriculum Principal | (57) 10 _____ School Board Members |
| (50) 3 _____ Supervisor | (54) 7 _____ Department Head | (58) 11 _____ Other _____ |
| (51) 4 _____ Superintendent | (55) 8 _____ Assistant Superintendent | |
- E. From the above select the number of the one person with primary responsibility for evaluation of teachers in your school system. _____ (59-60)
- F. Who do you think should be responsible for teacher evaluation in a school system? (check all appropriate responses)
- | | | |
|--------------------------------|---------------------------------------|------------------------------------|
| (61) 1 _____ Principal | (65) 5 _____ Assistant Principal | (69) 9 _____ Teachers |
| (62) 2 _____ Personnel Officer | (66) 6 _____ Curriculum Principal | (70) 10 _____ School Board Members |
| (63) 3 _____ Supervisor | (67) 7 _____ Department Head | (71) 11 _____ Other _____ |
| (64) 4 _____ Superintendent | (68) 8 _____ Assistant Superintendent | |
- G. From the above select the number of the one person you think should have primary responsibility for teacher evaluation: _____ (72-73)
- Cd 2 Code _____
(1) (2-6)

III. METHODS OF EVALUATION

This part of the questionnaire is divided into two sections - left and right

- A. On the left check those items which are presently a part of the teacher evaluation system in your school or system.
- B. On the right indicate your attitudes regarding the desirability of including this item in a teacher evaluation system. Circle a number for each item. If you do not have an opinion or insufficient information - circle 5.

Check those which you are presently using in your teacher evaluation system:		Attitude toward including this method-even if it is not a part of your present system.					
		very desirable	desirable	undesirable	very undesirable	insufficient information	
(7) ☐	A. Teacher Checklist						
(8) ☐	1. Physical Characteristics	1	2	3	4	5	(16)
(9) ☐	2. Personal Characteristics	1	2	3	4	5	(17)
(10) ☐	3. Professional Qualities	1	2	3	4	5	(18)
(11) ☐	4. Attitudes and Values	1	2	3	4	5	(19)
(12) ☐	5. Professional Training & Experience	1	2	3	4	5	(20)
	6. Other, _____	1	2	3	4	5	(21)
(13) ☐	B. Record of Critical Incidents or Events	1	2	3	4	5	(22)
(14) ☐	C. Input From Parents	1	2	3	4	5	(23)
(15) ☐	D. Self-Evaluation	1	2	3	4	5	(24)

Altitude toward including this method even if not in your system:

Check those which you are presently using in your teacher evaluation system:

↓

		very desirable	desirable	undesirable	very undesirable	insufficient information	
	E. Evaluation by Objectives:						
(25) ☐	1. Job targets identified by teachers & administrators	1	2	3	4	5	(7)
(26) ☐	2. Job targets identified by teachers being evaluated	1	2	3	4	5	(8)
(27) ☐	3. Job targets identified by administrators only	1	2	3	4	5	(9)
(28) ☐	4. Job targets based on a needs assessment	1	2	3	4	5	(10)
(29) ☐	5. Needs assessment based on a job description	1	2	3	4	5	(11)
(30) ☐	6. Other,	1	2	3	4	5	(12)
(31) ☐	F. Competency Tests for Teachers	1	2	3	4	5	(13)
	G. Student Test Data for Teachers:						
(32) ☐	1. Standardized Tests	1	2	3	4	5	(14)
(33) ☐	2. Proficiency Tests	1	2	3	4	5	(15)
(34) ☐	3. Teacher-made Tests	1	2	3	4	5	(16)
(35) ☐	4. Student Attitudes Measures - paper and pencil	1	2	3	4	5	(17)
(36) ☐	5. Student Reports of Teachers Behaviors	1	2	3	4	5	(18)
(37) ☐	6. Student Attitude Measures - observations	1	2	3	4	5	(19)
(38) ☐	7. Other,	1	2	3	4	5	(20)
	H. Classroom Observations by:						
(39) ☐	1. Principal(s)	1	2	3	4	5	(21)
(40) ☐	2. Supervisor(s)	1	2	3	4	5	(22)
(41) ☐	3. Superintendent	1	2	3	4	5	(23)
(42) ☐	4. Department Chairman and/or team leader	1	2	3	4	5	(24)
(43) ☐	5. Peers	1	2	3	4	5	(25)
(44) ☐	6. Others,	1	2	3	4	5	(26)
	I. Pre-Observation Conference by:						
(45) ☐	1. Principal(s)	1	2	3	4	5	(27)
(46) ☐	2. Supervisor(s)	1	2	3	4	5	(28)
(47) ☐	3. Superintendent	1	2	3	4	5	(29)
(48) ☐	4. Department Chairman and/or team leader	1	2	3	4	5	(30)
(49) ☐	5. Peers	1	2	3	4	5	(31)
(50) ☐	6. Others,	1	2	3	4	5	(32)
	J. Post-Observation Conference by:						
(51) ☐	1. Principal(s)	1	2	3	4	5	(33)
(52) ☐	2. Supervisor(s)	1	2	3	4	5	(34)
(53) ☐	3. Superintendent	1	2	3	4	5	(35)
(54) ☐	4. Department Chairman and/or team leader	1	2	3	4	5	(36)
(55) ☐	5. Peers	1	2	3	4	5	(37)
(56) ☐	6. Others,	1	2	3	4	5	(38)
	K. Instructional Conferences:						
(57) ☐	1. Planning - early in the year	1	2	3	4	5	(39)
(58) ☐	2. Formative - throughout the year	1	2	3	4	5	(40)
(59) ☐	3. Summative - end of the year	1	2	3	4	5	(41)
(60) ☐	L. Casual Information from Others: (phone calls, notes, comments)	1	2	3	4	5	(42)

IV. EVALUATION PRIORITIES

A. How much time do you spend per teacher in the evaluation process in one year including pre- and post-conferences, observations, etc.

B. How much time should be spent per teacher in the evaluation process in one year including pre- and post-conferences, observations, etc.

(43)

1	0-1.9 hours	
2	2-3.9 hours	
3	4-5.9 hours	
4	6-7.9 hours	
5	8-9.9 hours	
6	10-11.9 hours	
7	12-13.9 hours	
8	14-15.9 hours	
9	16+ hours	

- C. On the average how many times is a teacher observed during an evaluation year by all those involved in the evaluation process? (45)
- | | | |
|--------|---------|----------|
| 1. 0 | 4. 5-6 | 7. 11-12 |
| 2. 1-2 | 5. 7-8 | 8. 13-14 |
| 3. 3-4 | 6. 9-10 | 9. 15+ |
- D. What is the average length in minutes of each observation? (46)
- | | | |
|-----------------|-------|-----------------|
| 1. less than 10 | 4. 30 | 7. 60 |
| 2. 10 | 5. 40 | 8. 70 |
| 3. 20 | 6. 50 | 9. more than 70 |
- E. On the average how many times should a teacher be observed during an evaluation year by all those involved in the evaluation process? (47)
- | | | |
|--------|---------|----------|
| 1. 0 | 4. 5-6 | 7. 11-12 |
| 2. 1-2 | 5. 7-8 | 8. 13-14 |
| 3. 3-4 | 6. 9-10 | 9. 15+ |
- F. What should be the average length in minutes of an observation? (48)
- | | | |
|-----------------|-------|-----------------|
| 1. less than 10 | 4. 30 | 7. 60 |
| 2. 10 | 5. 40 | 8. 70 |
| 3. 20 | 6. 50 | 9. more than 70 |
- G. Do teachers have an opportunity to see and review their evaluation documents? (49)
- | | |
|---|--|
| 1. ☐ yes, always | 3. ☐ not usually |
| 2. ☐ yes, if they wish | 4. ☐ no, confidential |
- Do teachers have the right to make written comments on disagreements concerning their evaluation? (50)
- | | | |
|---------------------------------|--------------------------------|--|
| 1. ☐ yes | 2. ☐ no | 3. ☐ don't know |
|---------------------------------|--------------------------------|--|
- Do teachers receive a copy of the final evaluation that is placed in their personnel file? (51)
- | | | |
|---------------------------------|--------------------------------|--|
| 1. ☐ yes | 2. ☐ no | 3. ☐ don't know |
|---------------------------------|--------------------------------|--|
- Are teachers informed of the evaluation process? (52)
- | | | |
|---------------------------------|--------------------------------|--|
| 1. ☐ yes | 2. ☐ no | 3. ☐ don't know |
|---------------------------------|--------------------------------|--|
- Are teachers involved in the development of evaluation instruments? (53)
- | | | |
|---------------------------------|--------------------------------|--|
| 1. ☐ yes | 2. ☐ no | 3. ☐ don't know |
|---------------------------------|--------------------------------|--|
- Are teachers involved in the development of evaluation procedures? (54)
- | | | |
|---------------------------------|--------------------------------|--|
| 1. ☐ yes | 2. ☐ no | 3. ☐ don't know |
|---------------------------------|--------------------------------|--|
- Are teachers informed of the evaluation appeal process? (55)
- | | | |
|---------------------------------|--------------------------------|--|
| 1. ☐ yes | 2. ☐ no | 3. ☐ don't know |
|---------------------------------|--------------------------------|--|
- How would you rate the overall evaluation process in your school system? (56)
- | | | |
|----------------------------------|----------------------------------|----------------------------------|
| 1. ☐ good | 2. ☐ fair | 3. ☐ poor |
|----------------------------------|----------------------------------|----------------------------------|

Thank you for participating in this study.

FOLD

Place
Stamp
Here

Patricia Miller
203 Henson Hall
The University of Tennessee
Knoxville, TN 37916

TABLE IX-A

ANALYSIS OF VARIANCE DATA SUPPORTING TABLE II, PAGE 56

Item		DF	SS	MS	F
Checklist Phys. Char.	Model	2	1.40	.70	1.42
	Error	719	353.28	.49	
	Total	721	354.68		
Pers. Char.	Model	9	.95	.47	1.16
	Error	789	323.05	.41	
	Total	791	324.00		
Prof. Qual.	Model	2	.60	.30	1.06
	Error	803	225.59	.28	
	Total	805	226.19		
Att. & Val.	Model	2	.00	.00	.00
	Error	796	297.17	.37	
	Total	798	297.17		
Pro. Train. and Exper.	Model	2	2.20	1.10	2.97
	Error	781	288.80	0.37	
	Total	783	291.00		
Crit. Even.	Model	2	10.57	5.28	8.59
	Error	646	397.34	0.62	
	Total	648	407.91		
Input Par.	Model	2	0.01	.01	.01
	Error	620	377.70	.61	
	Total	622	377.71		
Self-Eval.	Model	2	2.68	1.34	2.90
	Error	780	360.17	0.46	
	Total	782	362.85		
I.D. Teach and Admin.	Model	2	4.94	2.47	6.07
	Error	722	294.15	0.41	
	Total	724	299.09		
I.D. Teach Be Eval.	Model	2	10.59	5.30	8.94
	Error	696	412.28	0.59	
	Total	698	422.87		

TABLE IX-A (Continued)

Item		DF	SS	MS	F
I.D. Admin. Only	Model	2	10.35	5.18	8.90
	Error	659	383.44	0.58	
	Total	661	393.72		
J.T. Based Need Assess	Model	2	6.73	3.37	5.46
	Error	674	415.93	0.62	
	Total	676	422.66		
Need Assess on Job Desc.	Model	2	11.45	5.72	10.20
	Error	626	351.20	0.56	
	Total	628	362.65		
Comp. Test Teach	Model	2	13.37	6.69	8.85
	Error	662	500.03	0.76	
	Total	664	513.40		
Stud. Test Data Stan. Test	Model	2	5.71	2.86	3.84
	Error	704	523.45	0.74	
	Total	706	529.16		
Prof. Tests	Model	2	2.61	1.31	1.68
	Error	688	536.03	0.78	
	Total	690	538.64		
Teacher Tests	Model	2	1.26	.63	.76
	Error	699	579.82	.83	
	Total	701	581.08		
Paper/Pen	Model	2	1.87	.94	1.56
	Error	649	389.21	.60	
	Total	651	391.08		
Stud. Rpt. Teacher Be.	Model	2	1.98	.99	1.67
	Error	663	391.75	.59	
	Total	665	393.73		
Stud. Att. Mes/Obs	Model	2	.17	.09	.15
	Error	654	388.49	.59	
	Total	656	388.66		
Class Ofs. Prin.	Model	2	13.41	6.71	25.78
	Error	809	210.45	.26	
	Total	811	223.86		

TABLE IX-A (Continued)

Item		DF	SS	MS	F
Supv.	Model	2	11.46	5.73	12.36
	Error	778	360.83	.46	
	Total	780	372.29		
Superin.	Model	2	4.74	2.37	3.82
	Error	712	441.24	.62	
	Total	714	445.98		
Dept. Chr./ Team Leader	Model	2	1.79	.89	1.36
	Error	692	453.72	.66	
	Total	694	455.51		
Peers	Model	2	.85	.42	.70
	Error	681	415.23	.41	
	Total	683	416.08		
Pre-Obs. Conf.-Prin.	Model	2	10.61	5.30	15.47
	Error	780	267.35	.34	
	Total	782	277.96		
Supervisor	Model	2	8.92	4.46	8.95
	Error	739	368.36	.50	
	Total	741	377.28		
Superin.	Model	2	3.54	1.77	2.71
	Error	682	445.10	.65	
	Total	684	448.64		
Dept. Chr./ Team Leader	Model	2	2.31	1.15	1.86
	Error	663	412.73	.62	
	Total	665	415.04		
Peers	Model	2	1.81	.91	1.72
	Error	630	332.28	.53	
	Total	632	334.09		
Superin.	Model	2	1.66	.83	1.23
	Error	689	466.12	.68	
	Total	691	467.78		
Dept. Chr./ Team Leader	Model	2	1.93	.97	1.44
	Error	651	437.11	.67	
	Total	653	439.04		

TABLE IX-A (Continued)

Item		DF	SS	MS	F
Peers	Model	2	2.30	1.15	2.07
	Error	619	343.94	.56	
	Total	621	346.24		
Instr. Conf.	Model	2	1.03	.52	1.88
	Error	770	211.84	.28	
	Total	772	212.87		
Form/All Year	Model	2	.84	.41	
	Error	764	218.67	.29	
	Total	766	219.51		

TABLE X-A
ANALYSIS OF VARIANCE DATA SUPPORTING TABLE III, PAGE 63

Item	Source	DF	SS	MS	F
IV-A	Model	2	367.85	183.92	61.11*
	Error	815	2,452.75	3.01	
	Total	817	2,820.60		
IV-B	Model	2	358.72	179.36	38.33*
	Error	808	3,781.10	4.68	
	Total	810	4,139.82		
IV-C	Model	2	233.11	116.56	41.36*
	Error	822	2,316.45	2.82	
	Total	824	2,549.56		
IV-D	Model	2	457.10	228.56	87.69*
	Error	810	2,111.05	2.61	
	Total	812	2,568.15		
IV-E	Model	2	189.64	94.82	28.60*
	Error	821	2,721.85	3.31	
	Total	823	2,911.49		
IV-F	Model	2	211.37	105.68	48.52*
	Error	822	1,790.51	2.18	
	Total	824	2,001.88		

TABLE XI-A
SIGNIFICANT Z SCORES SUPPORTING TABLE IV, PAGE 66

Item	Number	CO - BL	CO - T	BL - T
IV	G. (1)		7.64	8.07
IV	G. (2)		6.27	7.46
IV	H.			9.68
IV	I.		7.26	7.15
IV	J.		8.25	8.38
IV	K.	3.00	14.02	11.00
IV	L.	2.16	13.69	10.31
IV	M.		12.61	12.02

Note: Z values of 1.65 or greater are significant at the .05 level. A dash (-) indicates that the score lacked significance.

VITA

Samuel Lawson Houston was born on September 1, 1946 in Birmingham, Alabama. He received his elementary and secondary education in the Fort Payne, Alabama schools. His Bachelor of Science degree is from The University of Alabama, Tuscaloosa with a major in Business Administration. He received secondary teacher certification in English and Social Studies in 1972 after teaching for three years at Williams Avenue Junior High School in Fort Payne.

The writer returned to The University of Alabama in 1972 to pursue a Master of Science degree in Educational Administration, which was awarded in August, 1972.

Upon completing the Masters program, he was employed at Oak Park Middle School in Decatur, Alabama where he taught English. In March, 1974 the writer was appointed Principal of Westlawn Elementary School. He served in this capacity until June, 1979 when he became a full-time doctoral student at The University of Tennessee, Knoxville.

He was admitted to candidacy for the Doctor of Education degree in 1980. During his residency he served as a graduate assistant in the Department of Educational Administration and Supervision. He received the Doctor of Education degree with a major in that department in March, 1981.

The author is a member of Phi Delta Kappa, Phi Kappa Phi, National Association of Elementary School Principals, Alabama Association of Elementary School Administrators, Alabama Council for School Administration and Supervision.