1989

Teacher Perceptions of Evaluation as an Agent for Teacher Growth and Improvement of Instruction

Beverly Ann Hobson
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

Part of the [Junior High, Intermediate, Middle School Education and Teaching Commons](#), and the [Other Teacher Education and Professional Development Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Hobson, Beverly Ann, "Teacher Perceptions of Evaluation as an Agent for Teacher Growth and Improvement of Instruction" (1989). *Dissertations and Theses*. Paper 1221.
<https://doi.org/10.15760/etd.1220>

This Dissertation is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

TEACHER PERCEPTIONS OF EVALUATION AS AN AGENT FOR
TEACHER GROWTH AND IMPROVEMENT OF INSTRUCTION

by

BEVERLY ANN HOBSON

A dissertation submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF EDUCATION
in
EDUCATIONAL LEADERSHIP:
EDUCATIONAL ADMINISTRATION AND SUPERVISION

Portland State University
© 1989

TO THE OFFICE OF GRADUATE STUDIES:

The members of the Committee approve the dissertation
of Beverly Ann Hobson presented June 19, 1989.

John D. Lind, Chair

Judith E. Allen

Kenneth D. Peterson

Maxine L. Thomas

Jack C. Finley

APPROVED:

Robert B. Everhart, Dean, School of Education

C. William Savery, Interim Vice Provost for
Graduate Studies and Research

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to the members of my dissertation committee for their assistance in the development of this study. I especially want to thank Jack Lind, my chairperson, for his guidance and direction and Ken Peterson for his insight and expertise in the field of teacher evaluation.

I also want to acknowledge my friends and colleagues for their continuing interest and encouragement. The support I received from my principal, Doug Smith, has been invaluable. Special thanks go to Ray Brown and Doug Anderson for their assistance in processing the data and to Pati Sluys for her diligent work in the typing of this dissertation.

I am most grateful for the confidence and support of my parents who taught me the value of pursuing a goal through hard work and perseverance. My greatest appreciation is extended to my children, Mark and Kara, for their love and patience throughout my doctoral studies.

TABLE OF CONTENTS

	PAGE
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	.viii
CHAPTER	
I INTRODUCTION	1
Statement of the Problem	3
Purpose of the Study	9
Significance of the Study.	9
Research Questions	12
Research Hypotheses.	12
Assumptions.	13
Limitations of the Study	13
Definition of Terms.	14
Organization of the Document	16
II REVIEW OF THE LITERATURE	18
Introduction	18
Purposes of Teacher Evaluation	19
Summative Evaluation	
Formative Evaluation	
Problems and Conflicts	
Teacher Attributes	24
Evaluator Attributes	26

CHAPTER

PAGE

	Evaluation Procedures.	33
	Performance Standards	
	Sources of Teacher Performance Data	
	Frequency of Observations	
	Evaluation Feedback.	53
	The Context of Teacher Evaluation.	55
	Amount of Time Spent on Teacher Evaluation	
	Professional Development Resources	
	Factors Regulating Teacher Evaluation	
	Evaluation Issues Regarding Elementary and Secondary Teachers	67
	Teacher Attributes	
	Evaluator Attributes	
	Evaluation Procedures	
	Professional Development Resources	
	A Summary of the Literature Review	71
III	METHODOLOGY.	74
	Method of Research	74
	Statistical Hypotheses	75
	Description of the School District	76
	Description of Subjects.	79
	The Instrument	81
	Data Collection.	85
	Data Analysis.	89
IV	PRESENTATION AND ANALYSIS OF DATA.	91
	Introduction.	91
	Description of Respondents.	92

CHAPTER

PAGE

Teacher Perceptions Regarding Evaluation	94
Frequency Distribution Summary of District Results Item Means and Standard Deviations for Total Teacher Sample	
The Relationship Between Attributes of Teacher Evaluation and Outcomes. . . .	107
Differences Between Elementary and Secondary Teachers Regarding Evaluation	111
Teacher Attributes Evaluator Attributes Evaluation Procedures Evaluation Feedback Evaluation Context	
Differences Between Classroom and Nonclassroom Teachers Regarding Evaluation	121
Teacher Attributes Evaluator Attributes Evaluation Procedures Evaluation Feedback Evaluation Context	
Summary of Results	129
V SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS. . .	135
Introduction	135
Summary.	135
Results	
Conclusions and Implications	143
Limitations of the Study	145
Recommendations.	146
Future Study	

REFERENCES	153
----------------------	-----

APPENDICES

A	DISTRIBUTION OF TEACHERS INCLUDED IN THE SAMPLE.	163
B	CORRESPONDENCE TO TEACHERS	166

LIST OF TABLES

TABLE		PAGE
I	Distribution of Respondents by Level of Instruction.	93
II	Distribution of Respondents by Role	94
III	Frequency Distribution Summary of District Results	95
IV	Item Means and Standard Deviations for Total Teacher Sample.	103
V	Descriptive Data on the Five Subscales.	106
VI	Correlations Between Individual Questionnaire Items and Outcome Ratings	109
VII	Correlations Between the Five Subscales and Outcome Ratings	112
VIII	Teacher Perceptions of Self Regarding Evaluation.	115
IX	Teacher Perceptions of the Evaluator.	117
X	Teacher Perceptions of Evaluation Procedures.	118
XI	Teacher Perceptions of Evaluation Feedback.	120
XII	Teacher Perceptions Regarding the Context of Evaluation	121
XIII	Classroom and Nonclassroom Teacher Perceptions of Self Regarding Evaluation.	123
XIV	Classroom and Nonclassroom Teacher Perceptions of the Evaluator.	125
XV	Classroom and Nonclassroom Teacher Perceptions of Evaluation Procedures.	126

TABLE		PAGE
XVI	Classroom and Nonclassroom Teacher Perceptions of Evaluation Feedback.	128
XVII	Classroom and Nonclassroom Teacher Perceptions Regarding the Context of Evaluation	130

AN ABSTRACT OF THE DISSERTATION OF Beverly Ann Hobson for
the Doctor of Education in Educational Leadership presented
June 19, 1989.

Title: Teacher Perceptions of Evaluation as an Agent for
Teacher Growth and Improvement of Instruction.

APPROVED BY THE MEMBERS OF THE DISSERTATION COMMITTEE:

[REDACTED]

John D. Lind, Chair

[REDACTED]

Judith E. Allen

[REDACTED]

Kenneth D. Peterson

[REDACTED]

Maxine L. Thomas

[REDACTED]

Jack C. Finley

The purpose of this study was to measure the perceptions of teachers regarding their most recent evaluation

experience and to determine whether teachers perceive any significant relationships between attributes of teacher evaluation and its quality and impact on teacher growth. The study also considered whether there are significant differences between elementary and secondary teachers, as well as between classroom and nonclassroom teachers, in their perceptions of evaluation.

The suburban school district in this study serves approximately 22,000 students. The subjects included 402 elementary and secondary teachers who were randomly selected from 1,081 permanent teachers.

The Teacher Evaluation Profile (TEP) instrument was used to measure teachers' perceptions of their most recent evaluation experience. The 44 items on the questionnaire were examined based on the following five categories of evaluation: (a) teacher attributes; (b) evaluator attributes; (c) evaluation procedures; (d) evaluation feedback; and (e) evaluation context. A total of 284 respondents, or 71 percent of the teachers surveyed, returned a completed questionnaire.

Data were reported in terms of frequency distributions, means, and standard deviations. Data analysis consisted of correlational analyses and an analysis of variance.

The results of this study suggested that teachers judge the quality of their evaluation based on the

attributes of the person who evaluates them and the feedback they receive. The quality of evaluation appears to be determined by the following attributes of effective feedback: the merit of the ideas and suggestions contained in the feedback, the depth of information provided, the specificity of information provided, and the amount of information received. Teachers appreciate an evaluator who gives useful suggestions for improvement, has a persuasive rationale for suggestions, and is a credible source of feedback.

None of the attributes on the TEP had a significant relationship to the overall impact of evaluation on teacher growth. The results indicated that significant differences exist between elementary and secondary teachers, as well as between classroom and nonclassroom teachers, in their perceptions of evaluation.

Recommendations were made for establishing a teacher evaluation system that is supportive of professional growth.

CHAPTER I

INTRODUCTION

Effective teacher evaluation that focuses on the professional development of teachers has the potential to improve the quality of instruction and contribute to school improvement (Lewis, 1982). Increasingly, educational policymakers consider more effective teachers as the key to better education (Wise, Darling-Hammond, McLaughlin, & Bernstein, 1984). As a result, the majority of school districts look to teacher evaluation as a way of improving teacher performance (Bolton, 1973; Educational Research Service [ERS], 1978; Ellis, 1984).

Performance of the nation's public elementary and secondary schools has been a concern of both professional educators and the general public. During the past decade public education has been swept by a series of reform efforts intended to increase accountability, efficiency, and effectiveness. Interest in the evaluation and supervision of teachers has heightened, and the public wants assurance and evidence that teachers are competent professionals (Drake, 1984; Duke, 1985; Natriello, 1983). The movement toward increased accountability in education has led school districts to reassess their teacher evaluation systems (Wise

et al., 1984). Soundly based evaluation of teachers can help answer public demands for greater accountability (Natriello & Dornbusch, 1981).

In the report, A Nation at Risk, the National Commission on Excellence in Education [NCEE] (1983) recommended teacher evaluation for several issues related to teaching.

Salary, promotion, tenure, and retention decisions should be tied to an effective evaluation system that includes peer review so that superior teachers can be rewarded, average ones encouraged, and poor ones either improved or terminated. (p. 25)

Since 1969, the Gallup organization has conducted an annual poll of the public's attitudes toward its schools. The results of these surveys have consistently expressed public school parents' concerns about the quality of teaching in their local schools (Bridges & Graves, 1984). Improving the quality of teachers was the most frequent response to the 1979 Gallup Poll's question on what public schools had to do to earn an "A" in performance (Gallup, 1979). In view of the interest on educational reform since the release of the report A Nation at Risk, respondents to the 1987 Phi Delta Kappa Gallup Poll were asked whether, in the past five years, the public schools in their community have improved, gotten worse, or stayed about the same. The findings showed that 25 percent of the parents and public believed that public schools in their community have

improved, and 22 percent believed that they have gotten worse (Gallup & Clark, 1987).

The National Education Association (NEA) favors the evaluation of teachers as a means for improved performance and personal growth but not for accountability or control (McNeil, 1981). The NEA position is that evaluation should be based upon written criteria and procedures mutually developed by and agreed upon by the teacher association, administration, and governing board.

STATEMENT OF THE PROBLEM

There is widespread consensus regarding the state of personnel evaluation practices in education and the need for substantial reforms (Darling-Hammond, Wise, & Pease, 1983; Scriven, 1983). Reports that challenged educators to bring about major reforms also charged that personnel evaluation practices are inadequate (NCEE, 1983; Goodlad, 1984).

Natriello, Hoag, Deal, and Dornbusch (1977) reviewed the literature on teacher evaluation and concluded that effective evaluation systems do not exist in schools today. Similarly, Scriven (1981) found no exemplary teacher evaluation system where current practices match available knowledge. According to the Rand Report, this widespread concern with teacher evaluation practices exists because most school districts do not spend enough time developing effective evaluation systems (Wise et al., 1984).

Stiggins and Duke (1986) contended that teacher evaluation has the potential to help most teachers improve, yet in actual practice, it does little to improve the quality of performance and instruction. Research findings suggest that teachers seldom derive any professional improvement from participating in the evaluation process (Stiggins & Bridgeford, 1985).

There is general agreement that teacher evaluation is often pro forma, meaningless, and ineffective. Scriven (1981) described teacher evaluation as a disaster with shoddy practices and unclear principles. In many schools there is no systematic supervision of teachers by school administrators, and in most other schools, this process is infrequent and ritualistic at best (Guthrie & Willower, 1973; Murphy, 1987). Guthrie and Willower (1973) concluded that the present system of classroom observation is essentially impotent as a method of improving instruction. Supervision usually turns out to be little more than a paper audit for organizational record keeping (Stiggins & Bridgeford, 1985; Wise et al., 1984).

McLaughlin and Pfeifer (1986) found that teachers and administrators typically see teacher evaluation as threatening and irrelevant. The Rand Report went on to state that other problems include teacher resistance or apathy toward evaluation, the lack of uniformity and consistency of evaluation within a school district,

inadequate inservice and support for evaluators, and difficulties in the evaluation of secondary school staff and specialists (Wise et al., 1984).

Evaluation of teachers by principals is the most common form of evaluation and must be improved if teacher evaluation is to become more effective (Natriello et al., 1977). Kauchak, Peterson, and Driscoll (1985) found that teachers view the principal's supervisory visit as perfunctory with little or no impact on their actual teaching practice. Teachers feel that the visits are designed more to maintain the status quo than to improve instruction or evaluate. Furthermore, teachers perceive the principal as lacking supervisory and instructional competence.

A Rand Study survey by Wise et al. (1984) concluded that "principals lacked sufficient resolve and competence to evaluate accurately" (p. 22) and that the conflict between the principal as instructional leader and evaluator is still unresolved. Respondents stated that "principals considered evaluation a necessary evil or a time-consuming chore" (p. 22). Anderson (1982) observed that some supervisors are easily diverted from working with teachers because they are uncomfortable in their supervisory role or because managerial duties seem more important.

Educational administrators are aware of the problems with teacher evaluation and are seeking refinements in the process. A survey by the American Association of School Administrators (Lewis, 1982) specified the following personnel evaluation needs:

- o Better definitions of effective teaching. Although many evaluation procedures attempt to define effective teaching and teachers, the emphasis seems to be on observation of teacher behavior with little emphasis on how the behavior accommodates learning styles and produces outcomes.
- o More trust in the process. "How to evaluate people and get them to feel good about it" is how one superintendent expressed his concern. In many places, the "spirit" of evaluation procedures has been so strictured by teacher contract agreements that it is almost "pro forma."
- o Proof of the link between evaluation and instructional improvement. Until there is some specific indication that the process is worth the trouble, some say it will remain "pro forma." A major issue facing school districts is the use of pupil achievement, measured primarily by standardized tests, as an indicator of teacher performance.
- o More specifics on evaluation techniques. Conferences, personal goal-setting, classroom observations--these are common to evaluation, but administrators want to know how to do them better.
- o More sensitivity to the needs of the evaluator, primarily the principal. Many participants feel they have neither the skills nor time for successful evaluations. They are also under pressure "to cease accepting marginal services," as one administrator expressed it, "and to become more assertive." What kind of training should they have and how should evaluators themselves be evaluated? (p. 11)

The two major roles of teacher evaluation, summative and formative, often serve conflicting purposes.

Evaluations that address the issue of accountability provide summative information about the value and quality of a teacher's performance. Summative evaluations support personnel decisions such as retention, dismissal, promotion, tenure, assignments, and salary increases. Formative evaluations, also referred to as growth-oriented evaluations in this study, promote the professional development of teachers, increase the effectiveness of an individual's performance, and provide information on teacher strengths and weaknesses so that appropriate resources can be made available.

Barber (1987) stated that the major reason teacher evaluation systems fail is due to the indiscriminate mixing of formative and summative purposes. The purposes of an evaluation system must be clearly understood and reflected in the procedures and processes (McGreal, 1983). Improved performance may be difficult to achieve in an atmosphere where job decisions are directly related to the results of evaluation (Beckham, 1981).

Few models exist to guide districts interested in linking staff development and teacher evaluation processes (Wise et al., 1984). Pfeifer (1986) contended that the lack of time and necessary resources, the lack of trust between teachers and administrators, and the lack of acceptable models linking staff development with evaluation make it difficult for school districts to promote the instructional

improvement of teachers while simultaneously holding them accountable for performance standards.

The lack of a clear definition of the teaching task causes many teacher evaluation systems to fail (Barber, 1987). Drake (1984) and Popham (1986) voiced concern about a defensible technology for teacher evaluation and the failure to reach consensus on what characterizes an effective teacher or constitutes effective teaching. Drake and Popham each concluded that teacher effectiveness research has not clearly demonstrated that a good teacher possesses a particular trait or set of traits. They found no evidence that good teaching can be evaluated by a study of certain skills or by the existence of certain classroom conditions.

Good and Brophy (1984) found a great deal of uncertainty about the connections between teaching behaviors and student learning. Gudridge (1980) reported little evidence that any one method of teaching is superior to another. Moreover, researchers have not discovered any teaching strategies which will consistently work with all students. What appears to separate the competent teacher from the less competent teacher is knowing which technique is appropriate for a given student in a specific situation.

PURPOSE OF THE STUDY

The purpose of this study is to provide descriptive data and address the following significant issues associated with growth-oriented evaluation:

1. The perceptions of teachers regarding their evaluation experience;
2. The relationship between specific attributes of teacher evaluation and its overall quality as perceived by teachers;
3. The relationship between specific attributes of teacher evaluation and its perceived impact by teachers;
4. The effect of level--elementary or secondary--on a teacher's perception of his or her evaluation experience;
5. The effect of role--classroom or nonclassroom--on a teacher's perception of his or her evaluation experience.

SIGNIFICANCE OF THE STUDY

This study will examine growth-oriented teacher evaluation because of its potential impact as a school-based method for improving teachers' skills and contributing to school improvement. In order to maximize the chances that teachers will grow professionally as a result of their experiences with evaluation, this study will explore the critical attributes that affect the quality of teacher evaluation. The key ingredients in an effective teacher

evaluation experience that lead to changes in attitudes and teaching practices will also be identified.

Teacher evaluation has assumed increasing importance as the demand for accountability in education has shifted to specific concerns about the quality of classroom teaching and teachers. These concerns have led to renewed interest in developing teacher evaluation systems that emphasize professional improvement and growth for all teachers, not just those having difficulty. An important characteristic of a professional teacher is the continual process of experimenting with new behaviors and improving teaching skills (Snyder & Anderson, 1986).

Stiggins and Duke (1986) found that teacher evaluation can lead to improved performance, personal growth, and professional esteem. They also emphasized that teacher evaluation should encourage tenured teachers to maximize their performance and share their strengths with other teachers. Research on school effectiveness demonstrates that characteristics of effective schools include norms of collegiality and continuous improvement (Little, 1981).

In order for an evaluation system to facilitate the development of professionals to their potential, Hackman and Oldham (1980) stated three conditions that must be present for high internal work motivation to develop and persist: knowledge of the actual results of one's work activities, responsibility for the results, and meaningful work

according to one's system of values. Herzberg (1968) found that the growth or motivator factors that are intrinsic to the job are achievement, recognition for this achievement, the work itself, responsibility, and growth or advancement.

A teacher evaluation model should be based on the principle of continued improvement of knowledge and skills because growth itself is a basic need of those in professions (Hackman & Oldham, 1980). Social science researchers have suggested that if the work environment provides opportunities for teachers to feel personally responsible for their work, they will develop a greater sense of personal competence, job satisfaction, and motivation (Hackman & Lawler, 1971). The most powerful incentives for teachers are those related to the achievement and development of students; teachers value reaching their students and knowing they have learned (Lortie, 1975). In fact, Lortie found that teachers look to their students, rather than to outside sources, for indicators of teaching performance, professional satisfaction, and encouragement.

This study should prove beneficial to the participating school district. In the spring of 1988, the assistant director of certificated personnel reviewed the current research and literature on teacher evaluation and examined the data from this study. Based on the findings, this district is not considering any immediate changes in its evaluation program. However, the results from this

study will provide baseline data that reflect the current attitudes and perceptions of the participating district's teachers toward the evaluation process. This information should prove useful for any future study by this district that focuses on how the evaluation process can be revised and made more relevant and useful to teachers.

RESEARCH QUESTIONS

Four research questions give direction to this study. The questions focus on teachers' perceptions of the evaluation process:

1. Is there a relationship between specific attributes of teacher evaluation and its perceived quality by teachers?
2. Is there a relationship between specific attributes of teacher evaluation and its perceived impact by teachers?
3. Is there a statistically significant difference between permanent elementary and secondary teachers in the perception of their evaluation experience?
4. Is there a statistically significant difference between permanent classroom and nonclassroom teachers in the perception of their evaluation experience?

RESEARCH HYPOTHESES

Out of the research questions come the following research hypotheses that are to be tested in this study:

Hypothesis 1. There are significant relationships between teachers' perceptions of the quality of their evaluation experience and the specific attributes of the evaluation process.

Hypothesis 2. There are significant relationships between teachers' perceptions of the impact of their evaluation experience and the specific attributes of the evaluation process.

Hypothesis 3. There are significant differences between permanent elementary and secondary teachers in the perception of their evaluation experience.

Hypothesis 4. There are significant differences between permanent classroom and nonclassroom teachers in the perception of their evaluation experience.

ASSUMPTIONS

The following assumptions were made in this study:

1. Teachers will give accurate perceptions on the survey instrument.
2. The sampling procedures used to identify respondents are appropriate, and the participants provide a representative sample to test the hypotheses.
3. The time during which the sample was taken will not affect the responses of permanent teachers.

LIMITATIONS OF THE STUDY

The following limitations may place restrictions on the conclusions of this study and their application to other situations:

1. A unique population of teachers participated in this study. They were permanent elementary and secondary

teachers from a large suburban school district located near a metropolitan city in Oregon.

2. The data collected are specific to the teachers in this school district. The findings of this study apply only to the population of permanent teachers in the school district sampled.

DEFINITION OF TERMS

For the purposes of this study, the following terms are defined:

Accountability. Accountability refers to evidence related to the value and quality of a person's performance that supports personnel decisions in regard to dismissal, demotion, promotion, and pay increases (Pfeifer, 1986).

Coaching. Coaching is providing expertise on how to improve and enhance a teaching technique through such processes as conducting conferences, arranging classroom demonstrations and visitations, and providing inservice classes (Sweeney & Lindsey, 1987).

Feedback. Feedback is the process of giving information for the purpose of changing the behavior of those receiving the information (McLaughlin & Pfeifer, 1986).

Formal Observations. Formal refers to prearranged observations that are preceded and followed by a conference with the evaluator (Duke & Stiggins, 1986).

Formative Evaluation. The goal of formative evaluation is to identify a teacher's strengths and weaknesses and plan appropriate professional development activities (Millman, 1981; Stiggins & Bridgeford, 1985).

Informal Observations. Informal refers to unannounced drop-in visits by the evaluator (Duke & Stiggins, 1986).

Level of Instruction. Level of instruction refers to being either an elementary teacher of K-6 grade students or a secondary teacher of 7-12 grade students.

Permanent Teachers. Permanent teachers are defined as those teachers who have successfully completed three years of probationary status and have successfully met all of the participating district's Standards of Competent Performance for years one, two, and three.

Role of the Teacher. The role of a teacher is defined as either being a classroom or a nonclassroom teacher. A classroom teacher is involved in the direct instruction of students. A nonclassroom teacher is in a support position and may or may not provide direct instruction. Nonclassroom teachers include media specialists, counselors, psychologists, program specialists, and teachers on special assignment.

Summative Evaluation. Summative evaluation provides a base for administrative decision making in regard to hiring and firing, promotion and tenure, assignments, and salary increases (Millman, 1981; Stiggins & Bridgeford, 1985).

Supervision. Supervision includes the responsibilities and activities designed to promote instructional improvement in schools (Sergiovanni, 1987).

Teacher Evaluation. The teacher evaluation process usually consists of a goal-setting plan, classroom observation, and conferences between teacher and supervisor before and after the observation. It may also include informal observations (Northwest Regional Educational Laboratory, 1988).

ORGANIZATION OF THE DOCUMENT

This dissertation is divided into five chapters. Chapter I provides an introduction and overview of the study. In addition, the statement of the problem, purpose of the study, significance of the study, research questions, research hypotheses, assumptions, limitations of the study, and definition of terms are discussed. Chapter II includes a review of the literature on teacher evaluation and a discussion of the evaluation issues regarding elementary and secondary teachers. Chapter III offers an examination of the methodology and procedures used to investigate the research problem and a description of the school district, subjects, and survey instrument. Chapter IV reports the analyses of the findings. Chapter V summarizes the findings of this study, presents the conclusions and implications,

discusses the limitations of the study, makes recommendations, and suggests areas for further research.

CHAPTER II

REVIEW OF THE LITERATURE

INTRODUCTION

In the current reform movement, growing public pressure for accountability has focused attention on the quality of teachers and the need for closer supervision (Duke, 1985; Popham, 1986; Weber, 1987). This concern for the quality of education and teachers has resulted in the reassessment of teacher evaluation practices by many school districts (Wise, Darling-Hammond, McLaughlin, & Bernstein, 1984; Wolf, 1973). As a result, pressure has been placed on school administrators to assess and upgrade the competency of their teaching staff (Ellis, 1984).

Duke and Stiggins (1986) described teacher evaluation as a highly personalized partnership between teachers and administrators. Based on their research and work in schools, they identified five keys to successful teacher evaluation: the teacher, the evaluator(s), teacher performance data, the nature and quality of the feedback, and the evaluation context. Duke and Stiggins recommended that these elements be considered separately as dimensions of the evaluation process and used as guidelines by teachers

and administrators in promoting teacher growth and school effectiveness.

The purpose of Chapter II is to review the literature on teacher evaluation with an emphasis on the purposes of evaluation and those attributes that describe the teacher, evaluator, procedures, feedback, and context that contribute to the professional growth of teachers. In addition, evaluation issues regarding elementary and secondary teachers will be discussed.

PURPOSES OF TEACHER EVALUATION

Bolton (cited in Barber, 1987) stated that "the general goal of teacher evaluation is to safeguard and improve the quality of instruction received by students" (p. 9). Bolton identified the purposes of teacher evaluation as improving instruction, rewarding superior performance, modifying assignments, protecting individuals and the organization, validating the selection process, satisfying policy and law, improving public decisions and operational decisions, providing a basis for career planning, and contributing to morale and cooperation through communication and trust. However, these purposes do not distinguish the two major categories of evaluation which are formative and summative.

Millman (1981) noted that the two major roles of teacher evaluation, summative and formative, are important

and useful but serve different purposes. They differ in terms of the decisions made, interests served, evaluation instruments used, impact on teachers, importance of the decisions, potential limitations, and benefits (Stiggins, 1986). Since a variety of decisions may be made on the basis of teacher evaluation, it is essential that all participants understand the purposes, procedures, and roles (Bolton, 1973). Summative evaluation judges overall teaching performance and leads to personnel management decisions relative to hiring, dismissal, promotion, tenure, assignments, and salary increases (Beckham, 1981; ERS, 1978; Lewis, 1982; Popham, 1988). The goal of formative evaluation is to improve teacher performance by providing specific information on a teacher's strengths and weaknesses so that appropriate professional growth activities and resources can be planned (ERS, 1978; Lewis, 1982).

Summative Evaluation

A goal of virtually every teacher evaluation system is to hold teachers accountable for demonstrating minimum levels of competency (ERS, 1978; Wise et al., 1984). Most teacher evaluation systems are summative and promote educational accountability by judging teacher effectiveness (Stiggins & Bridgeford, 1985). Accountability evaluation systems attempt to serve the interests of the district and community by protecting students from incompetent teachers

(Stiggins, 1986). However, this form of evaluation tends to affect only those teachers who are having difficulty.

Accountability evaluation systems should provide objective, standardized, and reliable information about teacher performance (Darling-Hammond, Wise, & Pease, 1983). The school district must specify the criteria, behavioral basis for ratings, and procedures (Wise et al., 1984). When the purpose of evaluation is to ensure that teachers have met the minimum levels of performance necessary for personnel management decisions, the performance standards and criteria need to be legally defensible and uniform in application for all teachers.

Bell (1986) noted that in many states comprehensive school reform legislation has made heavy demands on principals. The most controversial of these responsibilities involves evaluation for the promotion of teachers participating in career ladder, merit pay, and master and mentor teacher programs. Wise et al. (1984) recommended that decisions involving pay and promotion receive the same rigor as required for dismissal purposes and use procedures that are perceived as reliable and valid.

Formative Evaluation

Formative teacher evaluation is a valuable tool for improving instructional effectiveness and encouraging the professional development of teachers (Ellis, 1984; Popham,

1986). It is based on the principle that all teachers can improve some dimension of their performance, knowledge, and skills (Bacharach, Conley, & Shedd, 1987; Stiggins, 1986). Formative evaluation is designed to help the competent teacher attain new levels of professional excellence and offers a rich source of performance information on which to base professional development. It has the potential to improve instruction and individual development but often assumes a secondary role because it demands more time and effort than many evaluators can afford (Stiggins & Bridgeford, 1985). For overall school improvement and teaching reform, the focus of evaluation should be on teachers' professional growth and development (Ellis, 1984; Stiggins & Bridgeford, 1985).

Stiggins and Bridgeford (1985) recommended an emphasis on formative evaluation and an environment where it can be successful. They stated that for formative evaluation to work effectively, it needs to be the primary purpose of evaluation and include appropriate procedures for collecting information. Formative evaluation requires an openness to change and a commitment to improvement on the part of teachers and supervisors.

Evaluation for the improvement of instruction must be flexible and individualized for the teacher and the specific teaching context (Wise et al., 1984). Redfern (cited in Gudridge, 1980) noted that teachers and their evaluators

need to work cooperatively and agree on priorities, identify specific objectives, share decisions on what needs to be accomplished, and establish a time line. There also needs to be a cooperative selection of appropriate inservice, the provision of resources to support changing behavior, and the selection of individualized performance criteria for the accomplishment of objectives and new skills.

Problems and Conflicts

Most teacher evaluation systems have attempted to address two goals: to support personnel management decisions for accountability purposes and to improve instruction by promoting the professional development of teachers (Drake, 1984; Duke & Stiggins, 1986). Wood and Pohland (1979) suggested that helping teachers improve their teaching performance is fundamentally different from summative evaluation. Wise et al. (1984) concluded that a single evaluation process can serve only one goal satisfactorily and cannot meet both the goal of judging and serving as a stimulus for teacher improvement.

Popham (1988) stated that combining formative and summative evaluation has rendered both dysfunctional because they serve contradictory functions. He emphasized that the individuals who carry out formative teacher evaluations must be different from those who carry out summative teacher evaluations. The procedures employed and records gathered

should also be kept separate. In contrast, Hunter (1988) took the position that summative and formative evaluation are compatible and that to supervise or evaluate teachers one needs to be highly skilled in both to determine whether the teacher's decisions and behaviors were appropriate.

TEACHER ATTRIBUTES

Duke and Stiggins (1986) noted the following teacher attributes which may influence the evaluation process:

1. A teacher's instructional competence in regard to understanding the elements of instruction and the delivery of instructional services;
2. A teacher's personal expectations of himself or herself;
3. A teacher's openness to constructive suggestions that might enhance his or her effectiveness;
4. A teacher's orientation to change and willingness to learn and try new techniques and ideas;
5. A teacher's knowledge of the subject matter to be taught and understanding of how that content is reflected in the district's curriculum plan;
6. A teacher's general professional experiences that can influence his or her responsiveness to evaluation.

Berman and McLaughlin (1978) found that the attitude of teachers regarding their level of professional competence and ability to help even the most difficult of students may

have a positive effect on what happens in the classroom. They concluded that professionalism is a primary motivator for change. Teachers will spend the time and energy to learn new skills if they believe that they will become better teachers and that their students will improve.

Career length, another teacher attribute, is related to teacher performance. Berman and McLaughlin (1978) found evidence that many teachers become less effective as their length of teaching experience increases. The average teacher is most productive from approximately the third to sixth year of teaching and needs professional development activities after five to seven years to encourage growth. Berman and McLaughlin's research also indicated that teachers with many years of experience are less likely to change their practices. As a result of the teaching work force becoming relatively stable, Drake (1984) observed that school reform and change must be accomplished by working with existing personnel.

Popham (1986) stated that teachers tend to be partisan when judging their own instructional abilities. Barber (1987) also felt that teachers lack objectivity, accuracy, and reliability when evaluating themselves. He commented that teachers tend to regard themselves as proficient. Similarly, Stiggins and Duke (1986) found that teachers tend to rate their professional expectations of self, knowledge of subject matter, orientation to change, and technical

knowledge of teaching as very high. Stark and Lowther (1984) concluded that teachers feel relatively strong about their accomplishments. Teachers tend to view themselves as competent in the areas of planning, problem solving, communicating, and working with others.

EVALUATOR ATTRIBUTES

Duke and Stiggins (1986) reported that one component of effective teacher evaluation is the person who observes and evaluates. The attributes of the evaluator that affect the quality of the teacher evaluation process include credibility, persuasiveness, patience, trustworthiness, track record, and the ability to model needed improvements. The evaluator gains credibility by sharing knowledge that is relevant to the teacher and appropriate to the content area(s), the grade level, or the particular group of students being observed. Evaluators must be able to present clear, convincing reasons why change is needed. Patience is critical in the evaluation process, and explaining why change is needed takes time. Trust is also crucial to the relationship of the supervisor and teacher if the goal is teacher growth. In addition, every supervisor develops a track record for sound advice and his or her ability to model teaching skills.

McLaughlin and Pfeifer (1986) found that the credibility of the evaluator as a source of feedback is an

important feature of an effective evaluation system. For a teacher to recognize a problem, acknowledge a needed change, or act on a prescription for change, the teacher must respect the judgment of the evaluator and perceive that the feedback comes from someone with expertise.

Duke and Stiggins (1986) found trust a critical factor in the supervisor's ability to change teacher behavior. They proposed that the factors most likely related to trust include the supervisor's intentions regarding the purpose of evaluation, the maintaining of confidentiality, the consistency with which evaluation rules and regulations are applied, the honesty and sincerity of interpersonal communications, and the extent of collaboration and teacher participation in the supervisory process.

Lyman (1987) also stated that the challenge of effective teacher supervision is to build trust and encourage collaboration between teacher and supervisor. Wolf (1973) noted that the school administrator is in the best position to create an open and nonthreatening atmosphere that is supportive of teacher evaluation. Higher levels of trust are developed when teachers see consistency between what the supervisor says and does (Lyman, 1987).

Lyman (1987) identified specific supervisory behaviors that promote trust and contribute to the professional growth of teachers. He indicated that providing information about procedures, schedules, and expectations of the supervisory

process enhance trust. Supervisors who give positive comments and feedback, show a genuine interest in teachers, and make teachers feel valued also promote trust. Other supervisory practices that encourage trust include frequent ongoing observations, prompt and specific feedback, active listening, and support for teachers.

Rutherford (1985) found that effective principals take time to discover what is going on in classrooms and gather information through formal observations as well as informal methods. Edmonds (1981) and Good and Brophy (1984) noted that the principal needs to create opportunities for teachers to improve their skills by visiting their classes, systematically observing instruction, and providing feedback. Huff and Schoalman (1982) identified coaching skills as a competency found in high performing principals. They found that outstanding principals believe that teachers, given detailed and specific suggestions, will improve their performance.

The school principal has emerged in research and other writings as the primary catalyst for change and school improvement (Berman & McLaughlin, 1978; Goodlad, 1984; Guthrie & Willower, 1973; Lipham, 1982; Sweeney, 1982). Nearly all the literature on effective schools has identified the role of the school principal as an agent of change and emphasized the importance of strong instructional leadership in bringing about high levels of student

achievement (Brookover et al., 1982; Cawelti, 1982; Edmonds, 1979; Lipham, 1982; Sweeney, 1982; Weber, 1971). Weber (1971) observed that principals of effective schools are instrumental in setting the tone of the school, organizing and distributing the school's resources, and assisting staff in decisions on instructional strategies.

The role of the principal has become more complex over the past two decades (Fullan, 1982). The pressure of accountability laws, competency tests, collective bargaining, and mandates for equity and a more responsive curriculum have placed demands on school administrators to improve their skills (Cawelti, 1982). Due to the principal's critical position within the organization and the response of school districts to public pressure for improvement, considerable attention has been devoted to improving the performance of principals (Grippa, 1987).

As attention has focused on the principal as the instructional leader and key to school improvement, studies have shown that principals frequently feel uncertain of their role. Natriello and Dornbusch (1981) noted that many supervisors are unsure how to perform meaningful evaluation. These authors concluded that many administrators lack background in evaluation procedures and are unprepared to add systematic evaluation of teachers to their already crowded schedules.

Robinson (1978) observed that most supervisors receive no ongoing inservice to improve their classroom observational skills, and districts do little to support them in these functions. He also found that supervisors do little preparation before observing a teacher by either reviewing lesson plans, conferencing with the teacher, or reviewing previous observation notes. In most cases, supervisors are unable to define effective teaching and provide the necessary intervention skills (Snyder & Anderson, 1986).

Medley and Coker (1987) researched the accuracy of principals as predictors of teacher effectiveness. They found that principal estimates of teacher effectiveness correlated only .20 with objective data that measured student knowledge. These results are consistent with the findings of earlier studies. Medley and Coker's research did not support the idea that the average principal is a good judge of teacher performance.

Edmonds (1981) and Sweeney and Lindsey (1987) agreed with other researchers that teacher evaluation can make a significant impact on the quality of instruction. Lezotte and Bancroft (1985) stated that increased student learning can be achieved by the professional development of principals in effective schools research and effective teaching. Cawelti (1982) predicted that if principals improve their supervision skills and focus on the characteristics of effective teaching, one could anticipate more successful

schools, improved reliability of those evaluating teachers, less teacher dissatisfaction with supervision, and a reduction in political moves to legislate learning.

Howe (1983) also emphasized the need to focus on the education and professional development of school principals. Brookover and Lezotte (1979) and Cawelti (1982) recommended a program of professional development for school administrators which would emphasize instructional leadership. In order to improve instruction, the principal would need expertise in curriculum development, clinical supervision, staff development, and teacher evaluation (Cawelti, 1982). An Educational Research Survey on teacher evaluation reported by Kowalski (1978) indicated that 61.4 percent of school districts provide inservice for evaluators in the form of workshops, outside consultants, university or college courses, and administrator internship programs.

Bryne et al. (cited in Fullan, 1982) surveyed a sample of principals in the United States who rated how essential certain types of preservice or inservice courses were to their roles. The results were compared with a similar survey carried out twelve years earlier. In 1977, 71 percent of the principals rated supervision of instruction as essential knowledge compared to a 56 percent rating in 1965.

Brookover et al. (1982) concluded that it isn't enough for the principal to convey the expectation of academic

achievement without also supporting inservice education that stresses teaching strategies and behavior useful in achieving those expectations. Berman and McLaughlin (1978) reported that moral support by the principal is essential to the implementation and continuation of a new project and the creation of a school climate that gives the project legitimacy. One indicator of that support is whether the principal attends workshop sessions with teachers. Involvement of the principal in staff development provides necessary information and skills to help teachers implement the project.

Evaluation of teachers has traditionally been the role of the administrator (Wolf, 1973). Kowalski (1978) reported that 92.5 percent of the principals at the elementary level formally evaluate classroom teachers; 86.7 percent of those at the junior high school level formally evaluate classroom teachers; and 81.9 percent of those at the senior high school level formally evaluate classroom teachers.

Brookover et al. (1982) pointed out that although it is easier for the principal to assume the role of instructional leader and evaluator, other members of the school organization might be effective in this position. The evaluator might be the principal, assistant principal, department chairperson, instructional leader, or an influential teacher. However, when the purpose of evaluation is accountability and personnel action may

result, it is often required by law and teacher contract that the supervisor and evaluator be the same.

EVALUATION PROCEDURES

Most teachers do not like to be supervised, react defensively to supervision, and do not see anything to be gained from the process (Acheson & Gall, 1980; Wolf, 1973). Teachers are often critical of evaluation procedures and have expressed the need for more specific performance criteria, more frequent classroom observations, and feedback that is communicated as soon as possible following an observation (Bolton, 1973; Borich & Fenton, 1977; Natriello & Dornbusch, 1981). Natriello and Dornbusch indicated that evaluation criteria are not always shared with teachers, that teachers are sometimes uninformed about the information collected to evaluate their performance, and that insufficient time is taken to communicate evaluation results with them. Wolf (1973) found that teachers want reassurance that the criteria and procedures of evaluation will produce credible results.

Performance Standards

Based on the results of studies on teacher evaluation procedures, Natriello and Dornbusch (1981) found that communicating the criteria or standards by which teacher performance is evaluated is critical if teacher evaluation

is to have a positive impact on teacher performance. Teachers reported that they do not always know what criteria or standards are being used to evaluate their teaching performance. The data showed that in regard to teaching subject matter, 53 percent of the teachers knew the criteria used to evaluate them. In regard to character development, 43 percent of the teachers knew the criteria for teacher evaluation, 63 percent knew the criteria for maintaining classroom control, and 52 percent knew the criteria for record keeping.

Natriello and Dornbusch (1981) also concluded that teachers would be more satisfied if there was agreement among evaluators regarding the criteria and standards used for teacher evaluation. Teachers complained that the criteria vary between schools within the same district.

Performance standards and criteria vary with the purpose of evaluation. Duke and Stiggins (1986) suggested that competency evaluations require different standards than professional development evaluations. Wise et al. (1984) concluded that evaluation used for personnel decisions regarding tenure, dismissal, pay, and promotion require the highest reliability of results. Therefore, the evaluation criteria must be standardized and used with consistency.

In a content analysis of teacher evaluation instruments, Wood and Pohland (1979) found that only 28 percent of the items related to the instructional role of the teacher.

Almost 40 percent of teacher evaluation procedures appeared to place a high value on organizational maintenance rather than helping teachers improve their teaching performance.

McGreal (1982) observed that even when districts claimed that improvement of instruction was the primary purpose of teacher evaluation, their procedures, instruments, and standardized criteria were heavily weighted toward administrative concerns. McGreal suggested that up to 75 percent of the criteria on evaluation instruments are administrative in nature and have nothing to do with the type of data collected in the classroom.

Sources of Teacher Performance Data

Traditionally, classroom observation has been the predominant method for collecting data about teaching performance (McGreal, 1983). Teacher evaluation procedures have generally relied on supervisor ratings and the administrative checklist (Levin, 1979; Lewis, 1982; Robinson, 1978; Wood & Pohland, 1979). Furthermore, Lewis (1982) found that most school districts use the same evaluation checklist for regular and specialized teachers. McLaughlin and Pfeifer (1986) stated that efforts to reduce effective teaching to a series of checklists have not been successful. Levin (1979) concluded that "reliance on a single evaluation technique is unwise" (p. 244) because of the need to counteract biases found in the more commonly used evaluation methods. McGreal

(1983) noted that schools are increasingly using other sources of data for teacher evaluation.

Growth-oriented teacher evaluation procedures must provide rich descriptive information that highlights sources of difficulty as well as courses for change (Wise et al., 1984). Shulman (1987) and his research team are currently working on an approach to teacher assessment that would reflect the complexity and richness of teaching and inspire teachers to aim higher in creating curricula and designing programs. They are developing prototypes of teacher assessment that can serve as models for the National Board of Professional Teaching Standards.

Multiple Sources of Data. Given the complexity of teaching and learning, a number of researchers have recommended the use of multiple sources of data in teacher evaluation (Levin, 1979; Stiggins and Bridgeford, 1985). There is no one evaluation approach that is likely to capture enough information to successfully measure teaching competence, performance, or effectiveness (Darling-Hammond et al., 1983).

Popham (1986) found that current teacher evaluation techniques are lacking and recommended a multi-data model that would involve judgments by a review team of experienced master teachers. Team members would need to be familiar with alternative methods of collecting information and consider evidence from many sources, including classroom

observations, competency tests, and interviews. Popham concluded that a data source may be untrustworthy by itself but of value if used in combination with other sources, especially if consistent patterns emerge.

Multidimensional evaluation procedures need to be developed that include input from a variety of sources because no single source of data can be depended on by itself (Macdonald, 1981). Peterson (1984) reported that the complexity of the teaching act calls for the use of multiple and variable documentation approaches or data sources. Shulman (1987) recommended that multiple sources of information be used to assess the effect of new teaching strategies on students. Alternative sources of data about the quality of teaching performance include classroom observations, student evaluations, student performance, teacher competency tests, artifact collections, teacher portfolios, teacher self-evaluation, teacher interviews, peer evaluations, and assessment centers. This section addresses some of the sources for teacher performance information that are currently being advocated in teacher evaluation research and development.

Clinical Supervision. Clinical supervision is a highly respected and recommended supervision model (McGreal, 1982). Snyder and Anderson (1986) viewed clinical supervision as the most important and potentially useful tool that supervisors have for sampling instruction and

shaping teacher behavior. An increasing number of school districts throughout the country are using clinical supervision as a way to overcome the inadequacies of token and checklist approaches to supervision (Murphy, 1987).

The goal of clinical supervision is the professional growth and development of teachers, with an emphasis on the improvement of instruction (Acheson & Gall, 1980; Drake, 1984; Flanders, 1976). The purpose is to assist teachers in modifying existing patterns of teaching in a way that the teacher has selected (Flanders, 1976; Sergiovanni, 1987). Clinical supervision promotes a school climate in which continuous improvement becomes an essential part of every teacher's job (Ellis, 1984).

Clinical supervision is a data-based, analytical approach to supervision that has the potential to improve teaching. However, the process of frequent observations and conferences is seldom practiced to any significant extent (Snyder & Anderson, 1986). Cawelti and Reavis (1980) found that only 15 percent of teachers reported having any experience with clinical supervision. Sergiovanni (1987) concluded that clinical supervision is demanding in the time required by the supervisor and teacher and may be too much supervision for some teachers. McGreal (1982) noted that the complete clinical supervision model is not always practical to use in schools due to the inservice requirement and time commitment needed to go through the steps.

Clinical supervision was developed by Morris Cogan for use with students in the Master of Arts and Teaching Program at the Harvard School of Education in the 1960s. Cogan's primary concern was to improve the professional status of teachers, and he regarded reflective practice as the heart of clinical supervision (Garman, 1986).

The "clinic of the classroom" was a way of describing the activities where supervisor and teacher work together every day for a prolonged period of time (a practice not generally done in other forms of supervision). Cogan also liked the term "clinical" because it had an element of realism associated with it, as well as referring to someone who is trained to observe and analyze events in an empirical fashion. (Garman, 1986, pp. 4-5)

Garman (1986) noted that clinical supervision encourages a collegial relationship between supervisor and teacher and assumes that teachers should be treated as professionals. Clinical supervision depends on face-to-face interaction between teacher and supervisor, effective communication, and observation of a teacher's classroom behavior (Acheson & Gall, 1980; Krajewski & Anderson, 1980). In clinical supervision the teacher and supervisor work together to set goals for professional growth and determine evaluation procedures, process, and progress (Duke & Stiggins, 1986).

Research in clinical supervision settings confirmed the fact that systematic observation and feedback of classroom events enable teachers to change their teaching

performance (Snyder & Anderson, 1986). Drake (1984) noted that teacher attitudes are more positive toward clinical supervision than toward traditional supervision. In addition, Drake found that clinical supervision produces a more self-directed teacher and a more positive relationship between the teacher and supervisor.

The classic clinical supervision cycle addresses both long-range and short-range goals. Cogan (1973) originally described the clinical supervision cycle with eight stages; Goldhammer described the model with five stages; while Acheson and Gall (1980) emphasized three phases of the clinical supervision cycle. However, Cogan recognized that certain phases of the clinical supervision cycle might need to be altered or omitted, or new procedures instituted, depending on the relationship between the supervisor and the teacher (Garman, 1982). Sergiovanni (1987) noted that most authorities suggest clinical supervision contain the following five structured and systematic stages:

1. Preobservation Conference. The initial purpose of the preobservation conference is to establish rapport and trust between the teacher and the observer. This step also increases the information an observer has prior to an observation and establishes an agreement between the teacher and the observer regarding the purpose of the observation. The teacher's concerns are identified as well as possible

solutions to these concerns, and an appropriate observation instrument or approach is selected.

2. Observation. In order for the data to be accepted by both the teacher and supervisor as reliable and useful, the collection of detailed observational classroom data is based on the agreement made in the preobservation conference. The observer records a sample of behavior in a systematic and objective manner as unobtrusively as possible.

3. Analysis and Strategy. The observer analyzes the data collected from the classroom observation as they relate to the agreement made in the preobservation conference and to pedagogical theory and research. Teaching patterns and critical incidents are identified that have a particularly positive or negative effect on teaching and learning. With the information analyzed and organized, the observer develops a strategy for conferencing with the teacher.

4. Conference. The conference provides feedback to the teacher on the classroom teaching observation and focuses on issues previously agreed upon in the preobservation conference. The teacher and supervisor analyze and interpret the data and determine alternative approaches for the future.

5. Postobservation Conference. The postobservation conference is a joint analysis of the observation cycle experience in order to gain perspective and make long-range

professional growth more probable. This stage encourages reflective thinking about issues in teaching, provides reinforcement, and improves the level of teacher satisfaction.

Murphy (1987) pointed out potential problems for those who use clinical supervision as the primary way of working with teachers to improve instruction. A tendency exists for the process and procedure to assume greater significance than the objectives and content. Murphy emphasized that the process should be grounded on a firm knowledge base about effective teaching and learning. In addition, clinical supervision depends on formal observations which can be too limiting of a strategy for effective supervision. Most clinical supervision models use a narrow database of formal observations and conferences that focus on only a few specific objectives. Based on these concerns, Murphy concluded that principals need to collect as much data as possible by using a variety of strategies.

Snyder and Anderson (1986) emphasized the need for clinical supervision to be linked directly to school goals, performance standards, teacher goals and evaluation, and a clear definition of instruction and learning. Murphy (1987) found that most clinical supervision programs treat everyone the same and that all teachers are subject to the same set of procedures and methods. He indicated that supervisory procedures and methods should vary depending on the

characteristics of individual staff members. Glickman (1987) recommended that principals emphasize either directive, collaborative, or nondirective supervision depending on a teacher's level of commitment to teaching, level of abstraction, or ability to reason abstractly and solve instructional problems.

Popham (1986) noted that classroom observations are regarded as necessary in a teacher evaluation system. However, he felt that direct observation of teaching can be reactive in that it usually distorts a teacher's performance and results in a carefully rehearsed lesson. Furthermore, effective classroom observations depend on lengthy training which many classroom observers lack.

Student Evaluations. Tyler (cited in Mickler, 1985) stated that student evaluation of teachers is one source of information that could help teachers improve if used in conjunction with other sources for evaluation. However, the average teacher is uncomfortable with student evaluation and generally lacks faith in the ability of students to rate a teacher's performance accurately (McGreal, 1983). McGreal (1988) indicated that student evaluations can provide reliable information if they focus on describing life in the classroom rather than making judgments about the teacher.

Popham (1986) noted that one difficulty with students rating their teachers is that a student's perception of a teacher's skill is often biased by the teacher's popularity

or the level of student interest in a subject. Macdonald (1981) also reported that student ratings are subject to some popularity pull but partially reflect the amount of learning that takes place in the classroom.

Popham (1986) reported that in order for student ratings to have validity, student rating forms must be carefully designed and student anonymity protected. He found little evidence regarding the level of maturity required by students for making judgments about a teacher's skill. However, Macdonald (1981) concluded that student rating scales are usually appropriate from fourth grade up, and interview or reaction sheets can be used with younger children.

Natriello, Hoag, Deal, and Dornbusch (1977) and Levin (1979) found that most of the studies on student evaluation of teachers involved college students, although Levin found some studies using students from grade six through high school. Levin indicated that there are several highly reliable student rating forms available and that the ratings of teachers tend to be consistent among students over time. He found that evidence of student bias in teacher ratings is incomplete, but the findings appear to indicate that student characteristics such as age and sex are not related to evaluation results and that easier courses do not receive higher ratings. Factors that influence ratings include class size, teacher reputation, student interest in the

subject, attitude toward school in general, and grades received.

Kauchak, Peterson, and Driscoll (1985) interviewed and surveyed elementary and secondary teachers to investigate their attitudes toward student evaluations and found teachers were almost equally divided into three positions. One third saw student evaluations as being a valuable source of information if professional judgment was used to interpret the results. The middle third was more cautious in their acceptance of student reports. The remaining third doubted the validity or reliability of student feedback on teaching performance.

Kauchak, Peterson, and Driscoll (1985) found that elementary teachers had the greatest doubts about the use of student evaluations while 73 percent of the secondary teachers approved of student reports. The teachers recommended limiting the topics in student evaluations to those in the affective domain rather than instructional competencies. Teachers who were negative toward student evaluations stated that elementary students were unable to understand the complexities of teaching due to their young age. They were also concerned about student emotions influencing teacher evaluations and the inability of students to differentiate between favorite teachers and teaching competency.

Levin (1979) noted that validity is the most difficult problem with student evaluation of teachers. When student ratings of teachers are compared with supervisors or other teachers, results show substantial differences between the groups. In addition, Levin found that feedback from student ratings is not effective in changing teaching behavior.

Achievement Tests. Popham (1986) suggested that using student test performance for measuring teacher effectiveness presents problems because many standardized achievement tests do not take into account differences in instructional methods or students' abilities and attitudes. Other problems include teaching to the achievement test, limitations of achievement tests, and the regression effect (Soar, Medley, & Coker, 1983). Macdonald (1981) noted that standardized tests are not a valid measure of student learning because they do not necessarily reflect what is being taught in the classroom and are correlated highly with IQ tests.

Kauchak, Peterson, and Driscoll (1985) indicated that teachers are against the use of student achievement tests to evaluate teachers. They found that teachers question the validity of achievement tests to assess student learning and to measure teacher ability or performance.

Levin (1979) reviewed the research on linking teacher evaluation with student learning and concluded that the use of student gains to evaluate teaching appears to be a seldom

used and undesirable approach. Levin noted that the disadvantages and dangers of such a system include teaching to the test and the loss of long-range objectives to short-term gains on test scores.

Teacher Competency Tests. The recent concern about the quality of teaching has been accompanied by interest in the National Teacher Examinations (NTE) and teacher competency tests. Soar, Medley, and Coker (1983) found no convincing evidence that NTE scores predict success in teaching when compared with ratings of teaching competency or with student gains on achievement tests. Several states have mandated teacher competency tests for certification. These tests focus on minimum literacy and knowledge of subject matter. Although there is no evidence that scores on such tests are related to student outcomes, Soar, Medley and Coker concluded that colleges should consider such tests when admitting students to teacher education. Popham (1986) suggested that it may be reasonable for state officials to use teacher competency tests before hiring a teacher. However, he found no evidence that such tests reflect the actual classroom applications of a specific teacher's skills.

Artifact Collections. McGreal (1982) emphasized that an artifact collection should be a regular part of an evaluation system. Artifacts would include study guides, question sheets, homework assignments, practice sets,

experiments, descriptions of drill and practice activities, quizzes, and tests. Collecting and reviewing teacher artifacts takes on importance when one considers the teacher effectiveness research which shows that 50-70 percent of the average student's day is spent in seat work and related activities. McGreal recommended that at least once each semester, for a two to three week period or for a unit of work, all artifacts used or produced by the teacher be collected and reviewed with the supervisor.

Teacher Portfolios. Portfolios provide a way for teachers to document their own diverse situations and performances and can be combined with other data sources in an assessment system. The portfolio can be a place to collect artifacts and evidence of actual classroom practices in a teacher's current assignment and can include all instructional materials used to facilitate learning. It can be related to a teacher's efforts to improve his or her teaching, or it can be a display of a teacher's best efforts to teach in a given context (TAP Begins, 1988).

Recently, some teacher evaluation programs have required teachers to submit professional portfolios containing such items as lesson plans, quizzes, and descriptions of classroom projects (Popham, 1986). However, Popham was concerned that portfolios might become contrived extravaganzas designed to win recognition and promotions.

Shulman (1987) saw teacher assessment as an ongoing process which includes written assessments, assessment center exercises, documentation of supervised field experiences, and direct observation of teaching by trained observers. Evidence of these proceedings can be kept in a cumulative portfolio to document teaching ability.

Teacher Self-Evaluation. Self-evaluation involves teachers making judgments about their own teaching and improving or modifying teaching practices based on personal reflection. Natriello et al. (1977) found that many administrators see self-evaluation as an essential part of any teacher evaluation system. However, Popham (1986) concluded that teachers' self-evaluations are useful for purposes of formative evaluation but not as useful for summative evaluation. Barber (1987) noted that its greatest value is for self-understanding and self-improvement. McGreal (1983) reported that self-evaluation data is most effective when shared and discussed with someone else.

Levin (1979) found that teacher attitudes toward self-evaluation range from neutral to slightly favorable and that only a few studies indicated that some teachers are self-directed in their learning. However, Stark and Lowther (1984) found that 89 percent of teachers agreed or strongly agreed that they should assess their own work. Wolf (1973) reported that 58 percent of the teachers questioned believed

that they were not encouraged to evaluate their classroom programs.

Macdonald (1981) asserted that teachers by themselves have not shown any special awareness of difficulties in their own teaching. Studies show that teachers think they are doing fine until they are presented with direct feedback. Then, most teachers are open to change.

Teacher appraisal interviews probe the teacher's approaches to instruction. Teacher interviews are a recent development in evaluation, and Popham (1986) expressed concern about the likelihood of contrived, meaningless declarations. However, Lortie (1975) stated that interview data provide a rich source of information for analyses that cannot be found from self-administered teacher questionnaires.

Peer Evaluations. Kauchak, Peterson, and Driscoll (1985) found that most teachers in their study were positive about teachers evaluating other teachers. If seen as formative evaluation, teachers saw potential benefits from feedback shared through peer evaluations. However, some teachers were concerned over the possible disruption of professional relations within a school and the potential for increased professional competition and isolation. Over one third of the respondents suggested that this problem could be minimized by recruiting peer evaluators from other schools. Another concern centered around the background

evaluators would need in order to relate to the problems and conditions encountered by the teachers being evaluated. A number of respondents felt that this issue could be resolved by having peers from the same subject matter area or from the same grade level do the evaluating.

Lortie's (1975) data clearly rated informal peer exchanges above formal teacher evaluation systems as an important source of assessment. Lortie found that teachers consider each other their most important form of assistance and tend to adapt the classroom practices of others to their own style and context.

Frequency of Observations

An Educational Research Survey reported by Kowalski (1978) indicated that 46.4 percent of the responding districts conduct classroom observations of tenured or continuing teachers only once a year and 80.9 percent conduct at least two classroom observations a year of probationary teachers. In studies conducted or reviewed by Natriello and Dornbusch (1981), teachers expressed a need for more frequent sampling of their teaching performance and more frequent evaluations. The data showed a positive relationship between teacher satisfaction and the frequency of teacher evaluation. The more frequently teachers are evaluated the more likely they are to accept the evaluation process (Natriello, 1983). However, Natriello and Dornbusch

found that evaluations perceived as too frequent led to teacher dissatisfaction. They concluded that, overall, it is not the frequency of negative evaluations but the infrequency of any evaluation that causes teacher dissatisfaction.

Rothberg and Buchanan (1981) surveyed elementary and secondary teachers regarding evaluation procedures and found that in order to make assessment procedures more likely to improve classroom instruction, observations and follow-up conferences need to be more frequent and lengthy but less threatening. Many teachers suggested that observations be preceded by goal-setting conferences so that teachers knew which objectives to emphasize. Nearly half of the teachers reported that the postobservation conference was the part of the evaluation process most likely to lead to instructional improvement. Respondents recommended that follow-up conferences be scheduled as closely as possible to the classroom observation. The most frequently mentioned concerns were teacher stress (16 percent) followed by infrequency and brevity of observations (15 percent). In order to make teacher evaluation more likely to improve instruction, 42 percent of the respondents called for classroom observations that were more frequent, longer, more informal, and less threatening.

EVALUATION FEEDBACK

Research findings indicate that providing feedback is one of the most powerful tools that administrators and managers have in the evaluation process (Sweeney & Lindsey, 1987). The extent to which teachers grow as a result of teacher evaluation depends on the quality and perceived usefulness of the feedback they receive (Duke & Stiggins, 1986). McGreal (1983) stated that the manner in which feedback is presented to the teacher affects the teacher's willingness to participate in instructional improvement. In successful teacher evaluation, formal feedback occurs in the postobservation conference and the final conference at the end of the evaluation period. Little (1981) found that professional growth appears most likely in schools where teacher evaluation includes frequent structured observations and useful feedback. To be valuable, feedback must relate to performance, stress objectivity, and not threaten teachers beyond their ability to cope (Harris, 1986).

Feedback practices play an important role in teacher evaluation outcomes. Based on a study describing four districts which have made substantial progress in initiating and organizing teacher evaluation programs, McLaughlin and Pfeifer (1986) suggested that effective feedback needs to be timely, specific, credible, and perceived as nonpunitive. They found that immediate feedback has maximum learning

potential because motivation to change and anxiety about outcomes are highest right after a classroom observation.

McLaughlin and Pfeifer (1986) concluded that providing specific information is important because data based on classroom observations allow teachers to draw their own conclusions. In the case of a disagreement, evaluators and teachers can refer to the data and interpret it together. Feedback that relates specifically to the classroom observation also indicates that the evaluator has taken the evaluation process seriously.

Duke and Stiggins (1986) reported that successful evaluation requires that feedback procedures be planned carefully, delivered in a sensitive manner, and shared in a private, neutral setting. Feedback should originate from a credible source, describe specific aspects of the observed teaching, provide ideas and suggestions for improvement, be regularly scheduled, and be balanced between informal and formal observations. Stiggins and Duke (1986) found that those attributes of feedback most crucial to the quality and impact of teacher evaluation are the quality of the ideas contained in the feedback as well as the depth and specificity of information provided.

McLaughlin and Pfeifer (1986) stated that the most critical feature of effective feedback is the teachers' perceptions of the intended role of evaluation. Where important outcomes hang in the balance, evaluation will

produce anxiety for those involved. If teachers perceive evaluation as punitive rather than an occasion for reflection and growth, then teachers might hide their shortcomings, become defensive, and minimize risk taking. Teachers must believe that they will be supported in their efforts to change.

THE CONTEXT OF TEACHER EVALUATION

Teacher evaluation occurs in an organizational context with every school and district having its own unique culture. Duke and Stiggins (1986) listed six factors that influence the evaluation context: amount of time spent on evaluation, resources available for professional development, state law, district policy, contractual obligations, and history of labor relations.

Amount of Time Spent on Teacher Evaluation

According to Mackenzie (1970), time is an organization's scarcest and most critical resource. The typical evaluation system is complex, time-consuming for the principal and teacher, and does not produce direct benefits for the teacher or the district (Carey, 1983). Wise et al. (1984) commented that a major obstacle to successful evaluation is the lack of time for observing, conferencing, and supporting teachers who need intensive help. Bridges and Graves (1984) also observed that time is a severe

problem for principals in their responsibility for evaluating teachers.

Stiggins and Bridgeford (1985) noted that the issues of time and money might prevent districts from helping teachers improve. They found that teachers want an evaluation system that provides accurate information on their teaching effectiveness, an opportunity to acquire and master new teaching strategies, and collegial support when transferring newly acquired skills to the classroom. However, these activities demand more time, instructional involvement, and assessment than many principals are able to manage. Duke and Stiggins (1986) reported that it takes time to support teachers as they react to evaluation data, draw their own inferences, respond to the evaluator's analyses, and speculate on growth strategies.

Conflicting time and role demands are common reasons why the principal is unable to assume the role of instructional leader (Lipham, 1982). Research studies have indicated that principals spend most of their workday on managerial tasks that are unrelated to the supervision of instruction. Peterson (1978) found that principals are mainly involved in service, advisory, and auditing activities. They spend the greatest proportion of their time working with students on discipline problems, with teachers who have noninstructional needs, and on clerical activities required by their supervisors. Principals seldom

engage in classroom teaching issues or activities involving change and innovation.

Howell (1981) demonstrated that principals spend less than one third of their time providing instructional leadership and that most of their time is spent on operational duties and other noninstructional functions. He found that elementary principals spend an average of 10 percent of their time on classroom supervision and 10 percent on teacher evaluation.

Stronge (1988) analyzed the results of an on-the-job time allocation study of principals and discovered that elementary school principals spent only 4.4 percent of their time in curriculum planning and 6.6 percent in instructional supervision. High school principals spent 11.3 percent of their time on these same two activities.

Miller and Lieberman (1982) noted that there isn't much time built into the principal's day for meaningful educational dialogue, planning, and evaluation. Cawelti (1982) concluded that the reason for the lack of time spent on instructional leadership is due to the low level of confidence principals have in instructional matters and the fact that they are normally rewarded for running a "tight ship." Principals must strike an appropriate balance between the emphasis they place on instruction as opposed to managerial functions and activities (Lipham, 1982).

Professional Development Resources

Staff development provides a major opportunity to improve instructional effectiveness in a school district and facilitate lasting change (Berman & McLaughlin, 1978). Inservice activities contribute to an ethos in which professional growth is expected and valued (Duke, 1982). Principals can treat teachers as professionals and maintain an effective teaching staff by encouraging them to develop their skills through workshops and an inservice program designed to keep teachers in touch with developments in their fields (Duke, 1982).

Hunter and Russell (1987) noted that staff development which promotes a basis for making skilled teaching decisions is an essential aspect of effective schooling. They claimed that districts supportive of ongoing inservice for instructional effectiveness showed gains in student learning, parent support, and the professional satisfaction of teachers.

Little (1981) conducted a study that explored ways in which the social organization of the school relates to teacher attitude toward new teaching practices and learning on the job. She found that staff development appears to have the greatest influence where there is a norm of collegiality and where teachers value and participate in a broad range of professional interactions with fellow teachers or administrators.

Berman and McLaughlin (1978) reported that teacher visitations to other schools or districts seem to aid in the implementation of new projects. Peers appear to be effective counselors to colleagues new to a project when it comes to advising them about problems they could expect, suggesting remedies, and offering encouragement.

Little (1981) reported that in schools characterized by the staff as collegial, teachers view the principal as actively endorsing and participating in collegial work. She concluded that school improvement is more likely achieved when the following conditions are present:

1. Teachers engage in frequent and continuous talk about teaching and develop a shared language that becomes increasingly concrete and precise.
2. Teachers and administrators frequently observe each other teaching and provide useful feedback.
3. Teachers and administrators plan, design, research, evaluate, and prepare teaching materials together.
4. Teachers and administrators teach each other about topics or tasks related to the practice of teaching.

Little (1981) noted that staff development has greater influence where there are expectations for analysis, evaluation, and experimentation in the school. Her findings suggest that the more clearly principals state their expectations and openness to alternative approaches, the more likely teachers will experiment and be innovative.

Teachers develop an attitude toward teaching practices by the advice they receive and the evaluations they are given.

Principals appear to build norms of collegiality and experimentation by the way they control the resources of a school (Little, 1981). First, principals control the distribution of internal resources and rewards such as schedules, assignments, the budget for materials, and meeting agendas. Second, principals are able to limit or expand a teacher's access to outside resources in regard to special proposals, consultants, and release time. Third, principals evaluate teaching performance and make judgments as to the level of teacher competency.

Rutherford (1985) and Brookover et al. (1982) also noted that in order to maximize teaching effectiveness and student achievement, effective principals ensure that the necessary instructional materials and resources are available for teachers. Duke and Stiggins (1986) maintained that resources for the professional improvement of teachers must be available if evaluations are to contribute to teacher growth. They recommended that districts provide the following resources:

1. Release time for teachers to visit other classrooms, model a lesson in a colleague's classroom, or attend workshops;
2. Technical assistance from consultants and district experts;

3. Data collecting tools to provide teachers regular feedback on teaching performance;
4. Videotaping lessons for the teacher's use;
5. Staff development activities;
6. Peer mentors;
7. Professional library resources.

Pfeifer (1986) concluded that when a school district commits the resources of time, expertise, and money to coordinate evaluation and staff development, it sends a powerful message to teachers about the value of their work and district priorities. Pfeifer recommended that if a district values teacher growth, it should provide teachers with available resources to act on evaluative recommendations. Evaluative feedback supported by staff development resources enables teachers to improve their instructional skills (McLaughlin & Pfeifer, 1986). Given evidence that teaching effectiveness may decline after five years of experience (Berman & McLaughlin, 1978), evaluation supported with appropriate assistance plays a critical role in the teaching profession.

Pfeifer (1986) explored the functions of staff development and its impact on evaluation outcomes. Pfeifer found that school districts have failed to coordinate teacher evaluation and staff development. His findings indicated that the impact of evaluation on teaching performance varies depending on the extent to which staff development practices

are related to evaluative feedback. Pfeifer concluded that staff development helps to construct an environment where evaluation is important to teachers, and it places the evaluator as a manager of opportunities for the professional growth of teachers. Taking time to document effective as well as incompetent teaching and providing resources to maintain teaching effectiveness tells teachers that excellence is valued.

Wise et al. (1984) also found that few districts coordinate the results of formative teacher evaluations with the planning and design of district staff development activities. Districts that tie teacher evaluation to curriculum goals tend to see evaluation and the development of instructional skills as integrated. Pfeifer (1986) noted that both teachers and districts benefit when schools assist teachers in identifying areas for professional growth that are consistent with organizational needs and then provide teachers with available resources, time, and incentives.

Factors Regulating Teacher Evaluation

The purpose of evaluation is to promote good education by selecting and retaining competent teachers and by improving the skills of teachers. Strike and Bull (1981) maintained that an evaluation program must be both fair and effective. Therefore, procedural rules for teacher

evaluation must consider the interests of the evaluator, the school, and society in promoting good education.

State laws and collective bargaining agreements require evaluation for personnel management purposes (Duke & Stiggins, 1986). State evaluation laws or administrative regulations mandate the evaluation of teachers, and collective bargaining agreements and contracts between teachers and school districts specify the evaluation procedures (Stiggins & Duke, 1986). Accountability evaluation systems rely on the law and contractual obligations to require teachers to participate and the threat of personnel action to promote growth when needed (Stiggins, 1986).

State Law and Teacher Evaluation. The operation of public schools is a responsibility reserved to the states and delegated to the local districts (Bolton, 1973). The right to use teacher evaluation as a basis for employment decision making in public schools is primarily governed by state laws (Beckham, 1981; McNeil, 1981). These laws serve as a context for teacher evaluation in districts and schools. State laws regarding teacher evaluation specify minimum standards and are usually written to protect students from incompetent and unethical educational practices as well as to maintain the due process rights of teachers (Duke & Stiggins, 1986).

Education expenditures and the income of educators have risen, but there has been no comparable improvement in

educational performance (March, 1978). Concerned with higher school expenditures but unproductive results, state legislatures began enacting laws requiring teacher evaluation in 1970 (McNeil, 1981). Duke and Stiggins (1986) found that forty-six states have a law or administrative regulation mandating teacher evaluation. This is a significant increase since only six states required teacher evaluation before 1971 (Wuhs & Manatt, 1983). Thirty-six of the states mandating teacher evaluation cite teacher improvement as a purpose of evaluation (Duke & Stiggins, 1986).

The specifics of evaluation requirements vary from state to state (Wuhs & Manatt, 1983; Duke & Stiggins, 1986; Popham, 1986). State laws and regulations vary in designating the personnel to be evaluated, the evaluator, performance standards, procedures, the frequency of evaluation, time lines, and grounds for dismissal (Beckham, 1981; Carey 1983; Duke & Stiggins, 1986). In most states, the control of teacher evaluation procedures is left to the local district (Stiggins & Duke, 1986).

Court Decisions Regarding Teacher Evaluation. The legal context of teacher evaluation is designed to promote fairness in decision making and pertains mainly to summative evaluation. Recent court decisions have forced standardization and rigor in teacher evaluation practices. Traditionally, courts have accepted the authority of school

boards or administrators to determine standards for teacher performance (Beckham, 1981). However, courts have insisted on strict compliance with the procedural requirements outlined in state statutes, board policies, or employment contracts. The cases surveyed by Beckham (1981) emphasize the importance of definitive standards for teacher evaluation and the consistent and uniform application of those standards in decision making. First, the school system must determine the knowledge, skills, and competencies it will require. Second, the school system must develop evaluative processes for identifying and preventing incompetence through early recognition of unsatisfactory teaching performance. Third, the teacher must be adequately informed of the standards and provided with an opportunity to correct identified problems.

Bridges and Graves (1984) noted that in choosing the criteria for evaluating the competence of classroom teachers, superintendents and local boards of education generally have considerable leeway. Courts realize that the evaluation of teachers is highly subjective and that there is no consensus within the teaching profession as to what constitutes adequate or competent teaching performance. In absence of state legislation to the contrary, courts have been inclined to accept the criteria employed by local school districts in evaluating classroom teachers.

The Impact of Collective Bargaining Agreements.

Collective bargaining agreements and contracts have been a major force in specifying teacher evaluation procedures (McNeil, 1981; Stiggins & Bridgeford, 1985). The main impact of these agreements has been to provide due process protection for teachers in case of personnel action by standardizing requirements in summative evaluation (Armiger, 1981; Stiggins & Bridgeford, 1985). In most states, teacher evaluation procedures have been considered a permissive issue for local teachers to negotiate (Beckham, 1981). Contracts vary across states and districts (Duke & Stiggins, 1986). However, the overall effect of these agreements is to promote uniformity and make procedures more specific (Stiggins & Bridgeford, 1985). In some cases the criteria for evaluation are negotiated, and in other cases the procedures for developing criteria are negotiated (Armiger, 1981).

Collective bargaining has done little to promote links between teacher evaluation and professional development. The National Education Association favors evaluation of teachers for instructional improvement but not for accountability or control (McNeil, 1981). However, by focusing on fair practices in personnel decision making, teacher organizations have directed attention toward legal requirements and away from promoting teacher growth and development (Stiggins & Duke, 1986).

EVALUATION ISSUES REGARDING ELEMENTARY AND SECONDARY TEACHERS

Teacher Attributes

The assessment of elementary and secondary teachers presents different challenges to supervisors in that elementary teachers are typically generalists and secondary teachers are more specialized (Shulman, 1987). Berman and McLaughlin (1978) concluded that secondary school teachers are "subject-oriented" in contrast to the "child-centered" orientation attributed to elementary teachers. Elementary teachers usually teach the same group of students for the day while secondary teachers meet with as many as 170 students each day and change the courses they teach two or three times a year. Stark and Lowther (1984) noted that elementary teachers are more likely than secondary teachers to favor the judgments of administrators, and secondary teachers are more favorable than elementary teachers toward peer assessment. Berman and McLaughlin (1978) found that change, for the most part, is more difficult to obtain and maintain at the secondary level. Shulman concluded that these differences between elementary and secondary teachers suggest that there will be differences in their knowledge, skills, dispositions, and educational orientation.

Evaluator Attributes

Few supervisors possess the breadth and depth of knowledge required to evaluate the subject matter competence

of teachers in such diverse fields as language, foreign language, mathematics, science, art, and music (Bridges & Graves, 1984). The issue of credibility at the secondary level is related to the observer's knowledge of content and familiarity with a teacher's classroom and students (Duke & Stiggins, 1986). To increase content credibility at the secondary level, some districts have relied on evaluation feedback from other sources such as peer review (Duke & Stiggins, 1986; Wise et al., 1984).

Respondents in the Wise et al. (1984) study reported problems with the evaluation of secondary school staff and specialists. Most of the respondents felt that the inability of their evaluation system to recognize differences in the performance of elementary, secondary, and specialist teachers is an unresolved issue related to the difficulty of an evaluator who is a generalist assessing the competence of a teacher who is a specialist.

Goodlad (1984) challenged the concept that principals need to be the instructional leaders in schools, acquiring the necessary specialized preparation, teaching pedagogical skills to teachers, and evaluating teachers' performance. Based on the findings of his study, Goodlad concluded that it would be difficult for principals to acquire and maintain teaching expertise beyond that of full-time teachers, especially at the secondary level with its diverse subjects. Goodlad recommended that highly qualified head teachers of

teaching teams be employed who would teach part time, serve as role models to fellow teachers, provide teachers with inservice assistance, diagnose severe learning problems, and work with full-time and part-time teachers, students in teacher preparation programs, and aides.

Evaluation Procedures

Shulman (1987) reported that most current teacher evaluations have grown out of a behavioral and generic view of teaching based on the effective-teaching literature. This literature has defined teaching skill in terms of observable classroom behavior and has been interpreted by many as being generic across ages, levels, and school subjects. Shulman found that some direct form of teacher observation is frequently used for teacher assessment. However, the rating scale that is employed by the evaluator generally fails to take into account differences in subject matter content or the age or level of the students. Shulman argued that teaching usually occurs in specific subject areas or skills and that assessment must examine the applications of pedagogy to the context in which teaching occurs. The subject matter, the learners, and the setting influence the kind of instruction that takes place.

Stodolsky (1984) reported that elementary teachers, who are essentially generalists, create a broad repertoire of organizational and pedagogical arrangements in the classroom. The variation of instructional approaches that

Stodolsky found in teachers at the elementary level is tied to subject matter and curriculum. These findings support the growing recognition that teacher behavior and instructional approaches vary depending on the purposes and subject matter content. The data have implications for current teacher evaluation practices. Since most teacher evaluations are based on a limited number of classroom observations, any given observation will not be representative of the range of teaching behaviors and skills used by an elementary teacher.

Professional Development Resources

The professional development activities of elementary and secondary teachers appear significantly different. Goodlad (1984) found that elementary teachers draw on a wider range of resources for their teaching than secondary teachers do. Furthermore, elementary teachers attend a greater variety of inservice activities for personal and professional improvement than secondary teachers do. Elementary teachers attend sessions that cut across the curriculum, whereas high school teachers usually attend more subject-specific sessions. The inservice activities attended by secondary teachers are more university based than those for elementary teachers. The differences between elementary and secondary teachers suggest the need for inservice strategies that address the specific issues of each group.

A SUMMARY OF THE LITERATURE REVIEW

Public concern over teacher evaluation has assumed increasing importance since the publication of national reports on education in the early 1980s. The public views teacher evaluation as a major problem. As a result, state legislatures have been trying to mandate more effective evaluation of teachers. In some cases, public pressure for accountability in education has resulted in a focus on summative teacher evaluation. The goal of summative evaluation is to facilitate management decision making in order to determine rehiring or firing, tenure, promotion or demotion, assignment, or salary schedule.

Until recently, much of the research on teacher evaluation has focused on evaluation as a source of information for accountability purposes. However, evidence is beginning to accumulate that suggests teacher evaluation can be a tool for teacher improvement. This has led to interest in the development of new teacher evaluation systems. Educators need to address the question of how teaching performance can be formally evaluated in a consistent and objective manner that encourages the professional growth of teachers.

The goal of formative teacher evaluation is to assist teachers in improving their performance by developing effective teaching skills. In formative evaluation the focus is on developing the art of teaching. The evaluator

assumes the role of collaborator and facilitator, assisting the teacher by providing coaching, support, and the necessary resources. The literature consistently shows that principals make a difference, and appropriate leadership is essential for educational change and improvement. However, the motivation to participate and demonstrate skills above the level of minimum competency must come from within each teacher.

A review of the literature on teacher evaluation illustrates the dissatisfaction of teachers, the frustration of administrators, and the confusion related to the purposes and methods of teacher evaluation. Teachers want an evaluation system that provides specific and accurate data on their teaching performance, an opportunity to master new learning techniques, and support when implementing change.

Clinical supervision is a collaborative evaluation model that promotes teacher growth. The teacher and supervisor work together to set goals and determine progress. Although clinical supervision has been implemented by many districts, frequent observations and conferences have not been successfully incorporated in most teacher evaluation systems. In response to concerns about many of the current evaluation systems, some educators are encouraging the use of multiple sources of data in order to provide a more comprehensive picture of a teacher's performance.

This study will help address the need in the literature for further research on how teacher evaluation procedures might promote teacher growth and development. More studies on effective evaluation systems are needed in order to broaden our view regarding the conditions under which teachers are open to growth, development, and change. A review of the literature also shows that although there have been a large number of studies on teacher perceptions of the evaluation process, there has been little research that specifically examines elementary and secondary teachers' perceptions or classroom and nonclassroom teachers' perceptions of the evaluation process.

CHAPTER III

METHODOLOGY

This chapter describes the methods and procedures used to investigate teachers' perceptions of the evaluation process. The following areas are discussed: (a) Method of Research; (b) Statistical Hypotheses; (c) Description of the School District; (d) Description of Subjects and Sample Size; (e) The Instrument; (f) Data Collection; and (g) Data Analysis.

METHOD OF RESEARCH

This study was designed to measure the perceptions of teachers regarding their most recent evaluation experience and to determine if teachers perceived any significant relationship between attributes of teacher evaluation and its quality and impact on professional growth. Differences in teachers' perceptions of the evaluation process were described according to the level of instruction (elementary or secondary) and the role of the teacher (classroom or nonclassroom).

This descriptive study dealt with relationships between nonmanipulated variables. The independent variables were the level of instruction and the role of the teacher.

The dependent variable was the measurement of teacher attitude as described on a questionnaire.

A survey research method was utilized to answer the research questions posed in this study. This method allowed the researcher to gather data from a relatively large number of subjects during one time frame.

Best and Kahn (1986) stated that the purpose of descriptive research is to describe, record, analyze, and interpret current relationships, practices, or trends. A descriptive study involves some type of comparison or contrast and deals with an analysis of the relationships between nonmanipulated variables, the testing of hypotheses, and the development of generalizations, principles, or theories.

STATISTICAL HYPOTHESES

The following null hypotheses are to be tested in this study:

1. There are no significant relationships between teachers' perceptions of the quality of their evaluation and the specific attributes of the evaluation process.
2. There are no significant relationships between teachers' perceptions of the impact of their evaluation experience and the specific attributes of the evaluation process.

3. There are no significant differences between permanent elementary and secondary teachers in the perception of their evaluation experience.

4. There are no significant differences between permanent classroom and nonclassroom teachers in the perception of their evaluation experience.

DESCRIPTION OF THE SCHOOL DISTRICT

The suburban school district in this study serves approximately 22,000 students in 26 elementary schools, 6 intermediate schools, and 3 high schools. Students score above the state and national averages on standardized tests. The district's average teacher/student ratio at the elementary level (grades K-6) is 1:24 and at the secondary level (grades 7-12) is 1:23. The cost of the education program per student in 1987-1988 is \$4,132 as compared to \$4,007 for the state average.

The community that the district serves has been supportive of bond and levy elections over the years. In a 1987 survey by the local chamber of commerce, 27 percent of the community gave the district an excellent rating, 38 percent gave a superior rating, and 29 percent gave an average rating.

To promote educational excellence, the school district developed a program designed to improve the instruction of students as well as attract, retain, and motivate

outstanding teachers. Teachers are encouraged to apply for professional development opportunities through a program established by the school board in 1985. Participants receive grants, developmental leaves, and awards for innovative curriculum.

The teacher population of this district has the following profile:

1. 468 teachers have a Bachelor's Degree.
2. 727 teachers have a Master's Degree.
3. 6 teachers have a Ph.D.
4. The average length of employment with the district is 11 years.
5. The average years in teaching is 14 years.
6. Teachers in this district make more money than those in most other metropolitan districts in the state.
7. The district has a collective bargaining agreement with its teachers.
8. 3,439 teachers applied for a teaching position for the 1987-1988 school year, and 200 teachers were hired.

This district has highly developed teacher evaluation procedures already in place. In 1970 the superintendent requested that the local teacher education association work with the district to develop a new personnel evaluation program. Teachers, administrators, community members, and high school students were members of the original development committee. The program was piloted in eleven schools,

representing the elementary, intermediate, and high school levels, prior to its adoption by the school board in 1971.

Over the next three years, this program was reviewed and modified. New teachers and administrators on the committee also provided assistance in its revision. These revisions were adopted by the school board in 1975.

New state requirements in 1979 changed the evaluation procedures in Oregon. At that time a district committee of administrators selected by the superintendent and teachers selected by the local teacher education association revised the personnel evaluation program to meet those procedural changes. The revised program was adopted by the school board in 1981.

During the 1985-1986 and 1986-1987 school years, a committee of administrators selected by the superintendent and teachers selected by the local teacher education association reviewed the Standards of Competent Performance to ensure that they reflected the current research of the profession and to include standards of competent performance for all categories of certificated staff. The Personnel Evaluation and Professional Development Program, as revised, was adopted by the school board in 1987.

The purposes of the Personnel Evaluation and Professional Development Program are to assure high quality instruction and to appropriately manage the operations of the district. The program serves as an evaluation tool as

well as an incentive toward improved instructional skills. The Evaluation Program is a cooperative endeavor in which the evaluator works with each teacher to promote the teacher's continual growth.

DESCRIPTION OF SUBJECTS

The population from which the sample for this study was selected included all teachers who have obtained permanent status in a suburban school district located near a large metropolitan area in Oregon. Permanent teachers were defined as those teachers who have completed three years of probationary status and have met all of the Standards of Competent Performance for years one, two, and three. The permanent teachers in this study included both elementary and secondary teachers as well as classroom and nonclassroom teachers. Teachers who retired during the 1987-1988 school year, left the school district, or were on sabbatical or leave of absence were not included in this study.

In accordance with the policy of the district where this study took place, teachers were categorized by role according to the standards by which they were evaluated. Classroom teachers are defined as those involved in the direct instruction of students and are evaluated on professional, communication, and instructional responsibilities. Nonclassroom teachers include media specialists, counselors, psychologists, and specialists. Media specialists are

accountable for professional, communication, instructional, and media responsibilities. Counselors and psychologists are evaluated on professional, communication, and counselor and psychologist responsibilities. Specialists include program specialists, evaluation specialists, teachers on special assignment, and staff development specialists. They are evaluated on professional, communication, and specialist responsibilities.

The district's personnel office provided the names of all members of the population being studied: 1,081 permanent teachers. Teachers were listed in alphabetical order by grade level taught (K-6), subject taught, or by title of position. This listing ensured that the following subgroupings would be satisfactorily represented in the sample: elementary teachers (K-6), secondary teachers (7-12), classroom teachers, and nonclassroom teachers.

Members of the sample were selected based on a stratified random sampling to guarantee that all subgroups in the population were proportionately represented. A stratified sampling was used because the research problem required comparison between various subgroups, and this method assured the researcher of an adequate number of respondents for subgroup analysis.

Teachers' names were systematically selected for this study. First, a name on page one of the computer printout of subjects was selected at random. Then from that

beginning point, every third name was selected. Three hundred sixty names were selected using this procedure. In order to select 42 more names and have them evenly distributed throughout the population, two numbers were selected at random using the following process. Given 51 names on a page, with 17 already selected during the first random sample, numbers 13 and 28 were selected at random out of the 34 possible names left on each page. By selecting the 13th and 28th unused names on each page, a total of 402 names was reached for the sample.

The 402 subjects selected for this study were grouped according to the grade level which they taught (elementary or secondary teachers), as well as by their role (classroom or nonclassroom teachers). See Appendix A for a distribution of teachers included in the survey sample.

THE INSTRUMENT

The Teacher Evaluation Profile (TEP), a data collection instrument developed by Stiggins and Duke from Northwest Regional Educational Laboratory, was used to collect data for this study. This instrument was chosen because it addresses many of the research questions in this study and allows the researcher to analyze the growth-producing potential of a particular teacher evaluation environment. The internal consistency reliability of the TEP is .93 (Duke & Stiggins, 1986) and, therefore, capable

of producing dependable or consistent data. The expected time to complete the survey is 15 minutes which makes it more likely to gain the cooperation of the subjects. This closed form questionnaire is easy to respond to on a National Computer Systems (NCS) response form, keeps the respondent on the subject, is relatively objective, and is fairly easy to tabulate and analyze.

The TEP questionnaire evolved from a program of research by Northwest Regional Educational Laboratory that identified the important attributes of a teacher evaluation environment that promoted the professional development of teachers. The items on the questionnaire resulted from case studies of successful teacher evaluations and interviews with teachers who had experienced growth through effective evaluation. Permission to reproduce this instrument for use in this study was granted by Stiggins in November of 1987. Agreement to minor modification of the directions and wording changes within the survey was given in February of 1988. The TEP questionnaire is available through the Northwest Regional Educational Laboratory.

The TEP questionnaire asks teachers to describe their most recent experience with their evaluation system. On the first part of the survey, teachers rate the overall quality of the evaluation using a 10-point scale, with 0 representing very poor quality and 9 very high quality.

Teachers also rate the overall impact of their last evaluation experience on their professional practices using a 10-point scale. A high rating of 9 reflects a strong impact leading to profound changes in their teaching practices, attitudes about teaching, and/or understanding of the teaching process. A low rating of 0 reflects no impact at all and no changes in their practices, attitudes, and/or understanding.

Teachers describe their most recent teacher evaluation experience on a scale of 1 to 5 in terms of the following 44 key attributes:

1. Attributes of the Teacher

- o Professional expectations of the teacher
- o Orientation to risk taking
- o Orientation to changing
- o Willingness to experiment in the classroom
- o Openness to criticism
- o Knowledge of technical aspects of teaching
- o Knowledge of subject matter
- o Years of teaching experience
- o Helpfulness of prior teacher evaluation

2. Attributes of the Evaluator

- o Credibility as a source of feedback
- o Working relationship with the teacher
- o Level of trust
- o Interpersonal manner

- o Temperament of the evaluator
- o Flexibility of the evaluator
- o Knowledge of the technical aspects of teaching
- o Capacity to demonstrate or model improvements
- o Familiarity with the teacher's classroom
- o Experience with classrooms in general
- o Usefulness of suggestions for improvement
- o Persuasiveness of rationale for changes

3. Attributes of the Evaluation Procedures

- o The manner in which standards were communicated
- o The clarity of standards
- o Endorsement of standards as appropriate
- o Uniformity of standards for all teachers

Extent of reliance on the following sources:

- o Observation of classroom performance
- o Examination of classroom or school records
- o Examination of student achievement

Extent of observation in teacher's classroom:

- o The number of formal classroom observations
- o The number of informal classroom observations

4. Attributes of the Feedback

- o Amount of information received
- o Frequency of formal feedback
- o Frequency of informal feedback
- o Depth of information provided
- o Quality of ideas and suggestions

- o Specificity of information provided
- o Nature of information provided
- o Timing of feedback
- o Focus of feedback on district standards

5. Attributes of the Evaluation Context

- o Amount of time spent by all participants involved in the evaluation process
- o Time allocated during the teaching day for professional development
- o Available inservice programs
- o Clarity of district policy statements regarding the purpose for evaluation
- o Intended role of evaluation

DATA COLLECTION

Permission to conduct a study which focused upon personnel evaluation was obtained on February 1, 1988, from the Director of Planning and Program Evaluation of the school district involved. Conditions included establishing a system that assured the district that the researcher would not be able to associate teacher responses with particular administrators and meeting with the director to work out data collection procedures. In addition, the leader of the local education association gave support to the collection of data from teachers. Agreement was made to share the

results of the survey with both the school district and the local education association.

The Portland State University Human Subjects Research Review Committee waived the application and full review of this research project because the data used in this study were institutional records that were the property of the district being studied. The data were obtained from employees as part of their job duties.

A cover letter and copy of the questionnaire were distributed to teachers and returned through the interschool mail. The cover letter explained the purpose of the study, described the procedures for maintaining confidentiality, stressed the value of the information teachers could supply, stated who would have access to the data, and offered an incentive to the respondent for completing the questionnaire. To increase the number of returns, a reminder letter was sent 12 days later and printed on buff colored paper to stand out among other notices that teachers received. The cover letter and questionnaire were xeroxed to produce high-quality copies to indicate the importance of the study and to increase the number of replies. Copies of the cover letter to the subjects and the follow-up letter are found in Appendix B.

Teachers participated in this study on a voluntary basis. In an effort to increase participation, a \$100 savings bond was purchased and awarded to a respondent who

was selected at random. Borg and Gall (1983) claimed that including a small cash reward, a small gift, or premium with the letter increases the response rate. Usually the reward should be a token of appreciation rather than a payment.

The time frame of the study was restricted to the month of March minus a week for spring vacation. March was chosen because it occurred before three major district surveys were to be sent to all teachers. In addition, the district's personnel office wanted the data collected and tabulated by late spring so that it could provide input on the current status of teacher perceptions regarding evaluation.

A cover letter and accompanying questionnaire were sent to 402 teachers selected as the original sample for this study. Twenty-one teachers (5.2 percent of the sample) were ineligible respondents due to retirement, resignation, medical disability, maternity, and leave of absence. That left 381 eligible respondents out of the original sample of 402.

All subjects of this study were asked to respond to the questionnaire between the dates of March 2, 1988, and March 11, 1988. The first mailing of questionnaires resulted in 219 responses from 57.5 percent of the eligible teachers contacted or 54.5 percent of all teachers in the original sample. A follow-up letter sent March 14, three days after the time limit set in the initial cover letter,

resulted in an additional 60 responses. This increased the response rate to 74.5 percent of eligible teachers or 70.6 percent of all teachers in the original sample. There was a 76.6 percent response rate for eligible elementary teachers or 72.1 percent for all elementary teachers in the original sample. There was a 71.0 percent response rate for eligible secondary teachers or 67.9 percent for all secondary teachers in the original sample. No pattern existed among nonrespondents as to whether they were elementary or secondary teachers; therefore, it was assumed that the nonrespondents did not represent biased data. Response rate tables may be found in Chapter IV that present data for elementary and secondary respondents as well as for classroom and nonclassroom respondents. Babbie (1973) suggested that a survey response rate of 50 percent is adequate, 60 percent is good, and 70 percent is very good.

Due to the sensitive nature of the questionnaire, steps were taken to support confidentiality in the hope of producing objective and honest responses. In order to protect the confidentiality of teacher responses on the TEP questionnaire, all responses and information were reported by category of respondents rather than by individual subjects or place. A code number at the top of each response form identified nonrespondents for an additional mailing. Research data were collected so that no one, including the researcher, could link the data to

specific subjects. A principal in the school district held the master list of code numbers and names but did not have access to individual response forms. The researcher had a list of code numbers but was unable to relate respondents' names to code numbers. This procedure also protected those responsible for teacher evaluation from being identified and compared. Soon after the cutoff date for returning response forms, a code number was drawn for the savings bond, and the master list of code numbers and names was sent to the researcher's advisor at Portland State University.

DATA ANALYSIS

The data collected on the Teacher Evaluation Profile are summarized and analyzed in Chapter IV in four forms: a frequency distribution summary, a profile of means and standard deviations, correlation analyses, and an analysis of variance. The frequency distribution summary reports the percent of all respondents selecting a given response option for each item on the questionnaire. The mean and standard deviation were tabulated for individual questionnaire items and reported for all respondents as well as for elementary and secondary teachers and classroom and nonclassroom teachers.

The Pearson product-moment coefficient of correlation was computed to examine the relationship between the teachers' ratings for each of the 44 attribute items on the

questionnaire and their ratings of the overall quality and impact of their most recent evaluation experience. The data identified those attribute items that were most highly correlated and predictive of positive impact and quality and those that were relatively low.

A two-tailed t test was used to determine whether the significance of difference between the means of two given groups differed by more than chance. The means of elementary and secondary teachers were compared as well as the means of classroom and nonclassroom teachers. An alpha of .05 was used to determine statistical significance.

CHAPTER IV

PRESENTATION AND ANALYSIS OF DATA

INTRODUCTION

The purpose of this study is to provide descriptive data and examine teacher perceptions of evaluation and its relationship to professional growth. This chapter reports the results of teachers' responses to the Teacher Evaluation Profile (TEP) questionnaire and furnishes evidence for accepting or rejecting the four research hypotheses in Chapter I. The presentation of data and discussion of findings are organized and reported for each of the following four research questions:

1. Is there a relationship between specific attributes of teacher evaluation and its perceived quality by teachers?
2. Is there a relationship between specific attributes of teacher evaluation and its perceived impact by teachers?
3. Is there a statistically significant difference between permanent elementary and secondary teachers in the perception of their evaluation experience?

4. Is there a statistically significant difference between permanent classroom and nonclassroom teachers in the perception of their evaluation experience?

The following sections are included in this chapter:

(a) Introduction; (b) Description of Respondents; (c) Teacher Perceptions Regarding Evaluation; (d) The Relationship Between Attributes of Teacher Evaluation and Outcomes; (e) Differences Between Elementary and Secondary Teachers Regarding Evaluation; (f) Differences Between Classroom and Nonclassroom Teachers Regarding Evaluation; and (g) Summary of Results.

DESCRIPTION OF RESPONDENTS

The percentages of teacher respondents by level--elementary and secondary--are presented in Table I. Of the 402 teachers surveyed, 70.6 percent (284) completed the questionnaire. The highest response rate was from elementary teachers; of the 204 elementary teachers surveyed, 72.1 percent (147) returned the questionnaire. The lowest response rate was from secondary teachers; of the 184 secondary teachers surveyed, 67.9 percent (125) returned the questionnaire. The 14 teachers who work with both elementary and secondary students had the highest response rate with 85.7 percent (12) of that group responding to the questionnaire. Out of the sample responding, 51.8 percent of the respondents were elementary teachers; 44 percent were

secondary teachers; and 4.2 percent were teachers who teach at both the elementary and secondary levels.

TABLE I
DISTRIBUTION OF RESPONDENTS BY
LEVEL OF INSTRUCTION

Categories of Respondents	Number Surveyed	Number Responding	Percent Responding	Percent of Sample Responding
Elementary Teachers	204	147	72.1	51.8
Secondary Teachers	184	125	67.9	44.0
Multiple Levels	14	12	85.7	4.2
Total	402	284	70.6	100.0

The percentages of teacher respondents by role--classroom and nonclassroom--are presented in Table II. Classroom teachers consisted of 351 of the 402 permanent teachers surveyed and 69.2 percent (243) responded. Nonclassroom teachers totaled 51 of those surveyed and 80.4 percent (41) responded. The highest response rate was from nonclassroom teachers (80.4 percent); the lowest response rate was from classroom teachers (69.2 percent). Out of the sample responding, 85.6 percent of the respondents were classroom teachers; 14.4 percent were nonclassroom teachers.

TABLE II
DISTRIBUTION OF RESPONDENTS BY ROLE

Categories of Respondents	Number Surveyed	Number Responding	Percent Responding	Percent of Sample Responding
Classroom Teachers	351	243	69.2	85.6
Nonclassroom Teachers	51	41	80.4	14.4
Total	402	284	70.6	100.0

TEACHER PERCEPTIONS REGARDING EVALUATION

One of the purposes of this study, as stated in Chapter I, was to examine the perceptions of teachers regarding their evaluation experience. This section will provide descriptive data on the perceptions of teachers regarding their evaluation experience by presenting a distributional report as well as the scale means and standard deviations for each item on the Teacher Evaluation Profile (TEP). Responses to the questionnaire will be examined for the total group of teacher respondents.

Frequency Distribution Summary of District Results

An item-by-item frequency distribution summary is reported in Table III. The distribution of responses for

each item on the TEP reports the percent of respondents selecting each response option.

TABLE III
FREQUENCY DISTRIBUTION SUMMARY
OF DISTRICT RESULTS

Number of respondents: 284

A. Describe these attributes of you as a teacher:

		Percent					
		1	2	3	4	5	
1. Professional expectations of yourself	I demand little	0	0	3	34	62	I demand a great deal
2. Orientation to risk taking	I avoid risks	1	6	29	42	22	I take risks
3. Orientation to change	I'm relatively slow to change	0	4	17	41	38	I'm relatively flexible
4. Orientation to experimentation in classroom	I don't experiment	0	4	21	44	31	I experiment frequently
5. Openness to criticism	I'm relatively closed	0	7	28	46	19	I'm relatively open
6. Knowledge of technical aspects of teaching	I know a little	0	0	11	52	37	I know a great deal
7. Knowledge of subject matter	I know a little	0	0	5	40	55	I know a great deal
		<u>Percent</u>					
8. Years of teaching experience	0 to 1 year	0					
	2 to 5 years	3					
	6 to 10 years	17					
	11 to 15 years	25					
	15 or more years	56					
9. Experience with teacher evaluation prior to most recent experience	Waste of time	7	22	42	25	5	Very helpful

TABLE III
FREQUENCY DISTRIBUTION SUMMARY
OF DISTRICT RESULTS
(continued)

B. Describe your perceptions of the person who evaluated your performance (most recently):

		Percent					
		1	2	3	4	5	
10. Credibility as a source of feedback	Not credible	3	11	24	28	35	Very credible
11. Working relationship with you	Adversary	2	6	17	30	45	Helper
12. Level of trust	Not trustworthy	5	6	17	28	45	Trustworthy
13. Interpersonal manner	Threatening	3	7	14	24	52	Not threatening
14. Temperament	Impatient	1	7	19	28	44	Patient
15. Flexibility	Rigid	4	11	22	34	30	Flexible
16. Knowledge of technical aspects of teaching	Not knowledgeable	1	8	14	36	42	Knowledgeable
17. Capacity to demonstrate or model needed improvements	Low	6	14	30	34	16	High
18. Familiarity with your particular classroom	Unfamiliar	9	17	29	29	16	Very familiar
19. Experience with classrooms in general	Little	5	9	24	30	32	A great deal
20. Usefulness of suggestions for improvements	Useless	6	13	30	36	15	Useful
21. Persuasiveness of rationale for suggestions	Not persuasive	6	11	39	32	12	Very persuasive

TABLE III
FREQUENCY DISTRIBUTION SUMMARY
OF DISTRICT RESULTS
(continued)

C. Describe these attributes of the procedures used during your most recent evaluation:

What procedures were used to address the district's Standards of Competent Performance by which you were evaluated?

		Percent				
		1	2	3	4	5
22. Were standards communicated to you?	Not at all	5	5	31	38	21
23. Were standards clear to you?	Vague	3	5	18	31	43
24. Were standards endorsed by you as appropriate for your classrooms?	Not endorsed	4	6	20	35	36
25. Were the standards . . .	The same for all teachers?	26	13	27	19	15

To what extent were the following sources of performance information tapped as part of the evaluation?

		Percent				
		1	2	3	4	5
26. Observation of your classroom performance	Not considered	4	12	19	26	38
27. Examination of classroom or school records (lesson plans, etc.)	Not considered	38	22	27	10	3
28. Examination of student achievement	Not considered	33	26	26	12	3

TABLE III
FREQUENCY DISTRIBUTION SUMMARY
OF DISTRICT RESULTS
(continued)

Extent of observation in your classroom, based on your most recent experience:

		<u>Percent</u>
29. Number of FORMAL (prescheduled) observations per year	0	12
	1	31
	2	46
	3	8
	4 or more	3
30. Approximate frequency of INFORMAL (unannounced drop-in) observations	None	24
	Less than 1 per month	50
	Once per month	16
	Once per week	8
	Daily	2

D. Please describe these attributes of the feedback you received:

			<u>Percent</u>				
			<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>
31. Amount of information received	None	5	16	37	29	13	Great deal
32. Frequency of formal feedback	Infrequent	20	27	31	14	7	Frequent
33. Frequency of informal feedback	Infrequent	21	22	30	18	9	Frequent
34. Depth of information provided	Shallow	13	15	33	27	12	In-depth
35. Quality of the ideas and suggestions contained in the feedback	Low	10	15	31	32	13	High
36. Specificity of information provided	General	10	13	29	28	21	Specific
37. Nature of information provided	Judgmental	5	8	22	39	27	Descriptive
38. Timing of the feedback	Delayed	6	8	12	34	41	Immediate
39. Feedback focused on district teaching standards	Ignored them	5	9	21	35	30	Reflected them

TABLE III
FREQUENCY DISTRIBUTION SUMMARY
OF DISTRICT RESULTS
(continued)

E. Describe these attributes of the evaluation context:

		Percent				
		1	2	3	4	5
40. Amount of time spent on the evaluation process including your time and that of all other participants	None	1	31	49	15	5
						Great deal

Resources available for professional development:

41. Time allotted during the teaching day for professional development	None	29	35	24	9	5	Great deal
42. Availability of training program and models of good practice	None	4	25	32	22	16	Many

District values and policies in evaluation:

43. Clarity of policy statements regarding purpose for evaluation	Vague	3	12	28	30	28	Clear
44. Intended role of evaluation	Teacher accountability	7	12	30	30	21	Teacher growth

RATING THE QUALITY AND IMPACT OF THE EVALUATION

		Percent											
		Very low	0	1	2	3	4	5	6	7	8	9	Very high
As you think about this experience, how would you rate the Overall Quality of the evaluation?			1	2	7	6	11	15	11	18	18	11	
Rate the Overall Impact of your last evaluation experience on your professional practices.			12	17	12	12	11	16	6	8	5	2	

When analyzing the distribution of teachers' responses for the overall quality of their most recent evaluation on a scale from 0 to 9, 73 percent of the respondents rated it a 5 or higher. A rating of 9 would have represented high quality, and a low rating of 0 would have represented very poor quality. In comparison, 37 percent of the respondents rated the overall impact of their last evaluation experience on professional practices a 5 or higher. A rating of 9 would have reflected a strong impact, and a low rating of 0 would have reflected no impact at all.

Teachers rated the 44 attributes of teacher evaluation on a scale of 1-5. Results are reported in terms of the five subscales of the profile: (a) Teacher Attributes; (b) Evaluator Attributes; (c) Evaluation Procedures; (d) Evaluation Feedback; and (e) Evaluation Context.

Teacher Attributes. Respondents scored teacher attributes consistently higher than other subscales on the profile. Three items seemed noteworthy with responses ranging from 3 to 5 on a 5-point descriptive scale. For professional expectations of self, 96 percent of the teachers rated that item high with a score of 4 or 5; for knowledge of subject matter, 95 percent scored it high; and for knowledge of the technical aspects of teaching, 89 percent scored it high.

In regard to length of teaching career, 98 percent of the teachers had at least six years of teaching experience.

Eighty-one percent of the teachers had at least eleven years of teaching experience.

Evaluator Attributes. For attributes describing the evaluator, only three out of twelve items were rated high (a score of 4 or 5) by at least 75 percent of the teachers. Those items included the evaluator's technical knowledge of teaching (78 percent), the evaluator's interpersonal manner (76 percent), and the evaluator's working relationship with the teacher (75 percent).

Evaluation Procedures. None of the attributes describing evaluation procedures were rated high (a score of 4 or 5) by at least 75 percent of respondents. Responses to the item, uniqueness of the standards, appeared evenly distributed among the response choices. Responses indicated that classroom or school records and student achievement were not often considered as sources of information for evaluation. Eighty-nine percent of the respondents received no more than two formal observations per year. Ninety percent of the respondents were observed informally no more than once per month.

Evaluation Feedback. As a group, the attributes describing feedback procedures appeared the most evenly distributed of all the subscales. Timing of feedback was rated high by 75 percent of the respondents. Frequency of formal feedback was rated low (a score of 1 or 2) by 47

percent of the respondents. Frequency of informal feedback was rated low by 43 percent of the respondents.

Evaluation Context. In regard to attributes describing the evaluation context, clarity of policy statements was rated highest with 58 percent of the respondents giving it a score of 4 or 5. Only 14 percent of the respondents rated time allotted during the teaching day for professional development as high (a score of 4 or 5).

Item Means and Standard Deviations
for Total Teacher Sample

Table IV reports the scale means and standard deviations for each item on the TEP questionnaire. Teachers rated overall quality and impact on a scale of 0-9. The mean for the overall quality of their most recent evaluation experience was 5.91, and the mean for the overall impact of their last evaluation experience on professional practices was 3.44. The 44 attributes of teacher evaluation were rated on a scale of 1-5. The means ranged from 2.15 to 4.59.

Teacher Attributes. Teachers tended to rate the attributes describing themselves as high. On a scale of 1-5, the highest rated behaviors were professional expectations of self ($\bar{M} = 4.59$, $SD = .57$) and knowledge of subject matter ($\bar{M} = 4.50$, $SD = .59$). Other attributes with means higher than 4.0 included knowledge of the technical aspects of teaching ($\bar{M} = 4.25$, $SD = .66$), orientation to

TABLE IV
ITEM MEANS AND STANDARD DEVIATIONS FOR
TOTAL TEACHER SAMPLE

	Mean	SD	N
Overall Quality (Scale 0-9)	5.91	2.21	239
Overall Impact (Scale 0-9)	3.44	2.45	239
<u>Item Number</u> (Scale 1-5)			
<u>TEACHER</u>			
1. Expectations of Self	4.59	.57	284
2. Orientation to Risk Taking	3.79	.88	284
3. Orientation to Change	4.11	.86	284
4. Experiment in Classroom	4.03	.82	283
5. Openness to Criticism	3.76	.84	284
6. Technical Knowledge	4.25	.66	283
7. Knowledge of Subject	4.50	.59	282
8. Years Teaching Experience	4.33	.86	284
9. Helpfulness of Prior Eval.	2.98	.97	284
<u>EVALUATOR</u>			
10. Credibility	3.81	1.11	284
11. Relationship to Teacher	4.11	1.01	284
12. Level of Trust	4.03	1.13	284
13. Interpersonal Manner	4.14	1.10	284
14. Patience	4.06	1.03	284
15. Flexibility	3.75	1.11	284
16. Technical Knowledge	4.09	.97	281
17. Capacity to Model	3.40	1.10	280
18. Familiarity with Classroom	3.25	1.19	283
19. Experience in Classrooms	3.73	1.16	284
20. Usefulness of Suggestions	3.42	1.07	283
21. Persuasive Rationale	3.33	1.02	282
<u>PROCEDURES</u>			
22. Standards Communicated	3.65	1.02	284
23. Clarity of Standards	4.06	1.05	283
24. Standards Endorsed	3.92	1.07	281
25. Uniformity of Standards	2.84	1.39	278
26. Observation of Classroom	3.81	1.19	282
27. Examination of Records	2.17	1.13	283
28. Examination of Achievement	2.27	1.14	281
29. # of Formal Observations	2.57	.91	283
30. # of Informal Observations	2.15	.95	283

TABLE IV
ITEM MEANS AND STANDARD DEVIATIONS FOR
TOTAL TEACHER SAMPLE
(continued)

	Mean	SD	N
<u>FEEDBACK</u>			
31. Amount of Information	3.30	1.05	283
32. Frequency of Formal	2.61	1.17	283
33. Frequency of Informal	2.72	1.24	283
34. Depth of Information	3.11	1.19	282
35. Quality of Ideas	3.23	1.14	281
36. Specificity of Information	3.37	1.22	281
37. Nature of Information	3.75	1.08	279
38. Appropriateness of Timing	3.95	1.18	281
39. Focused on Standards	3.77	1.12	280
<u>CONTEXT</u>			
40. Time Spent on Evaluation	2.93	.84	282
41. Time for Prof. Development	2.26	1.10	284
42. Available Training Programs	3.22	1.11	283
43. Clarity of Eval. Purposes	3.69	1.09	283
44. Intended Role of Evaluation	3.46	1.15	278

change ($M = 4.11$, $SD = .86$), and orientation to experimentation in the classroom ($M = 4.03$, $SD = .82$). The lowest rated teacher attributes were helpfulness of prior evaluation experiences ($M = 2.98$, $SD = .97$) and openness to criticism ($M = 3.76$, $SD = .84$). The standard deviations of the items within this subscale were consistently smaller than for the other items on the questionnaire.

Evaluator Attributes. The attributes describing the evaluator that received the highest ratings were interpersonal manner ($M = 4.14$, $SD = 1.10$), working

relationship with the teacher ($\underline{M} = 4.11$, $\underline{SD} = 1.01$), knowledge of the technical aspects of teaching ($\underline{M} = 4.09$, $\underline{SD} = .97$), patience ($M = 4.06$, $SD = 1.03$), and level of trust ($\underline{M} = 4.03$, $\underline{SD} = 1.13$). Attributes rated the lowest were persuasiveness of rationale for suggestions ($\underline{M} = 3.33$, $\underline{SD} = 1.02$) and familiarity with the teacher's particular classroom ($\underline{M} = 3.25$, $\underline{SD} = 1.19$).

Evaluation Procedures. In regard to evaluation procedures, the highest rated attributes were clarity of standards ($\underline{M} = 4.06$, $\underline{SD} = 1.05$) and endorsement of standards as appropriate for the teacher's classroom ($\underline{M} = 3.92$, $\underline{SD} = 1.07$). Attributes rated the lowest were the number of informal observations ($\underline{M} = 2.15$, $\underline{SD} = .95$) and examination of classroom or school records ($\underline{M} = 2.17$, $\underline{SD} = 1.13$).

Evaluation Feedback. Among the attributes related to evaluation feedback, immediacy of timing ($\underline{M} = 3.95$, $\underline{SD} = 1.18$) and feedback that focused on district teaching standards ($\underline{M} = 3.77$, $\underline{SD} = 1.12$) were rated the highest. Attributes rated the lowest were frequency of formal feedback ($\underline{M} = 2.61$, $\underline{SD} = 1.17$) and frequency of informal feedback ($\underline{M} = 2.72$, $\underline{SD} = 1.24$).

Evaluation Context. The two highest rated items describing evaluation context were clarity of policy statements regarding the purpose for evaluation ($\underline{M} = 3.69$, $\underline{SD} = 1.09$) and the intended role of evaluation ($\underline{M} = 3.46$, $\underline{SD} = 1.15$). Attributes rated the lowest were the amount of

time allotted during the teaching day for professional development ($\bar{M} = 2.26$, $SD = 1.10$) and the amount of time spent on the evaluation process ($\bar{M} = 2.93$, $SD = .84$).

Subscale Data. When the items on the Teacher Evaluation Profile questionnaire were grouped into five subscales and scored, the results produced the data described in Table V. Summary statistics for the means and standard deviations were calculated. The subscale describing teacher attributes had the highest mean score ($\bar{M} = 4.04$, $SD = .37$). The next highest rating was given to attributes describing the evaluator ($\bar{M} = 3.75$, $SD = .82$). The subscale describing evaluation procedures had the lowest mean score ($\bar{M} = 3.06$, $SD = .57$).

TABLE V
DESCRIPTIVE DATA ON THE FIVE SUBSCALES

Subscales	Mean	Standard Deviation
Teacher Attributes (Questions 1-9)	4.04	.37
Evaluator Attributes (Questions 10-21)	3.75	.82
Evaluation Procedures (Questions 22-30)	3.06	.57
Evaluation Feedback (Questions 31-39)	3.31	.84
Evaluation Context (Questions 40-44)	3.11	.70

THE RELATIONSHIP BETWEEN ATTRIBUTES OF TEACHER EVALUATION AND OUTCOMES

The Pearson product-moment coefficient of correlation was used to identify attributes of teacher evaluation that have the strongest and weakest association with perceived quality and impact. A high absolute value indicates a strong relationship, whereas a near zero value indicates a very weak relationship between the two variables being measured. Best and Kahn (1986) recommended the following criterion for evaluating the magnitude of a correlation:

COEFFICIENT (r)	RELATIONSHIP
.00 to .20	Negligible
.20 to .40	Low
.40 to .60	Moderate
.60 to .80	Substantial
.80 to 1.00	High to very high

The first research question, "Is there a relationship between specific attributes of teacher evaluation and its perceived quality by teachers?" is answered in Table VI. This table presents data on the relationship between individual questionnaire items and the overall quality of evaluation as perceived by the total group of respondents.

The attribute with the highest correlation to overall quality of evaluation was the quality of ideas and suggestions contained in the feedback ($r = .73$). Other items with a substantial correlation include usefulness of the

evaluator's suggestions for improvement ($r = .69$), depth of information provided in the feedback ($r = .66$), persuasiveness of the evaluator's rationale for suggestions ($r = .65$), specificity of information provided in the feedback ($r = .65$), amount of information received in the feedback ($r = .62$), and credibility of the evaluator as a source of feedback ($r = .61$).

The attributes with the lowest correlation to overall quality of teacher evaluation were a teacher's openness to criticism ($r = .04$), a teacher's orientation to change ($r = .06$), uniformity of standards ($r = .09$), and a teacher's expectations of self ($r = .10$). The attributes with the strongest negative correlation to overall quality of teacher evaluation included the teacher's knowledge of subject matter ($r = -.12$) and the teacher's willingness to experiment in the classroom ($r = -.07$).

The second research question, "Is there a relationship between specific attributes of teacher evaluation and its perceived impact by teachers?" is addressed in Table VI for the total group of respondents. Table VI presents data on the relationship between attributes of teacher evaluation and their overall impact on teacher growth.

None of the questionnaire items had a substantial or high correlation to overall impact of evaluation on teacher growth. The attribute with the highest correlation to teacher growth was the usefulness of the evaluator's

suggestions for improvement ($\underline{r} = .51$). Other attributes with a moderate correlation included the following items: persuasiveness of the evaluator's rationale for suggestions ($\underline{r} = .48$); the quality of ideas and suggestions contained in the feedback ($\underline{r} = .48$); the amount of information received in the feedback ($\underline{r} = .47$); the depth of information provided in the feedback ($\underline{r} = .45$); the helpfulness of prior evaluation ($\underline{r} = .41$); and the specificity of information provided in the feedback ($r = .40$).

TABLE VI
CORRELATIONS BETWEEN INDIVIDUAL QUESTIONNAIRE
ITEMS AND OUTCOME RATINGS

	Overall Quality	Overall Impact
<u>TEACHER</u>		
1. Expectations of Self	.0987	.0378
2. Orientation to Risk Taking	-.0109	-.0162
3. Orientation to Change	.0639	.0975
4. Experiment in Classroom	-.0697	.0047
5. Openness to Criticism	.0421	.0477
6. Technical Knowledge	-.0130	-.0485
7. Knowledge of Subject	-.1201	-.0559
8. Years Teaching Experience	-.0596	-.0397
9. Helpfulness of Prior Eval.	.4234	.4092
<u>EVALUATOR</u>		
10. Credibility	.6114	.3300
11. Relationship to Teacher	.5218	.2648
12. Level of Trust	.5510	.2665
13. Interpersonal Manner	.3594	.0800
14. Patience	.3875	.2345
15. Flexibility	.3880	.1954
16. Technical Knowledge	.4794	.2951
17. Capacity to Model	.4948	.2604

TABLE VI
CORRELATIONS BETWEEN INDIVIDUAL QUESTIONNAIRE
ITEMS AND OUTCOME RATINGS
(continued)

	Overall Quality	Overall Impact
18. Familiarity with Classroom	.5172	.2527
19. Experience in Classrooms	.4852	.3159
20. Usefulness of Suggestions	.6875	.5144
21. Persuasive Rationale	.6542	.4840
<u>PROCEDURES</u>		
22. Standards Communicated	.3777	.2466
23. Clarity of Standards	.3066	.2237
24. Standards Endorsed	.3885	.2290
25. Uniformity of Standards	.0935	.1289
26. Observation of Classroom	.4616	.2846
27. Examination of Records	.3467	.2613
28. Examine Achievement	.2582	.1880
29. # of Formal Observations	.3122	.2089
30. # of Informal Observation	.1952	.1035
<u>FEEDBACK</u>		
31. Amount of Information	.6211	.4703
32. Frequency of Formal	.5092	.3093
33. Frequency of Informal	.4137	.3068
34. Depth of Information	.6560	.4542
35. Quality of Ideas	.7278	.4766
36. Specificity of Information	.6512	.3953
37. Nature of Information	.4695	.2268
38. Appropriateness of Timing	.4375	.1792
39. Focused on Standards	.4477	.1871
<u>CONTEXT</u>		
40. Time Spent on Evaluation	.4818	.2902
41. Time for Prof. Development	.2103	.2199
42. Available Training Programs	.2076	.1448
43. Clarity of Eval. Purposes	.2133	.1628
44. Intended Role of Evaluation	.2880	.2040

The attributes with the lowest correlation to impact on teacher growth were the teacher's willingness to experiment in the classroom ($r = <.01$) and the teacher's expectations of self ($r = .04$). The attributes with the strongest negative correlation to impact on growth included the teacher's knowledge of subject matter ($r = -.06$) and knowledge of the technical aspects of teaching ($r = -.05$).

The items on the Teacher Evaluation Profile were grouped into five subscales and used to predict overall quality and impact. The results are presented in Table VII. Attributes of feedback ($r = .77$) and the evaluator ($r = .67$) were the best predictor of overall quality, while attributes of feedback ($r = .48$) and evaluation procedures ($r = .43$) most accurately predicted impact. The subscale describing teacher attributes appears unrelated to the perceived quality and impact of evaluation.

DIFFERENCES BETWEEN ELEMENTARY AND SECONDARY TEACHERS REGARDING EVALUATION

This section addresses the research question, "Is there a statistically significant difference between permanent elementary and secondary teachers in the perception of their evaluation experience?" The Teacher Evaluation Profile questionnaire was administered to determine whether teacher perceptions of the evaluation process varied between elementary and secondary teachers.

Tables VIII-XII present the scale means and standard deviations for each item as well as the t value, degrees of freedom, and the two-tail probability. Data are reported separately for elementary teachers (grades K through 6) and secondary teachers (grades 7 through 12). A two-tailed t test was used to determine whether the significance of difference between the means of elementary and secondary teachers differed by more than chance at $p < .05$ level of significance.

TABLE VII
CORRELATIONS BETWEEN THE FIVE SUBSCALES
AND OUTCOME RATINGS

Subscales	Overall Quality	Overall Impact
Teacher Attributes (Questions 1-9)	.1015	.1249
Evaluator Attributes (Questions 10-21)	.6727	.3869
Evaluation Procedures (Questions 22-30)	.5926	.4317
Evaluation Feedback (Questions 31-39)	.7673	.4797
Evaluation Context Questions (40-44)	.4170	.3045

Table VIII reports the means of elementary teachers' responses as higher in overall quality and impact than the

means of secondary teachers. Elementary teachers scored the overall quality of their most recent evaluation experience higher ($\bar{M} = 6.28$, $SD = 2.15$) as compared to secondary teachers ($\bar{M} = 5.52$, $SD = 2.14$). Elementary teachers scored the overall impact of their last evaluation experience on professional practices higher ($\bar{M} = 3.62$, $SD = 2.49$) as compared to secondary teachers ($\bar{M} = 3.22$, $SD = 2.38$). There was a statistically significant difference as elementary teachers rated the overall quality of their most recent evaluation higher ($t = 2.68$, $p = <.01$) than secondary teachers did.

Teacher Attributes

For both elementary and secondary teachers, the highest rated teacher attribute was expectations of self: elementary ($\bar{M} = 4.57$, $SD = .60$) and secondary ($\bar{M} = 4.57$, $SD = .56$). The second-highest rated teacher attribute for both groups was knowledge of subject matter: elementary teachers ($\bar{M} = 4.40$, $SD = .59$) and secondary teachers ($\bar{M} = 4.62$, $SD = .58$).

The data in Table VIII show that there were statistically significant differences between elementary and secondary teachers for the following teacher attributes:

1. Elementary teachers rated themselves higher in orientation to change ($t = 3.23$, $p = <.01$) than secondary teachers.

2. Secondary teachers rated themselves higher in knowledge of subject matter ($t = -3.02$, $p = <.01$) than elementary teachers.

3. Secondary teachers had more years of teaching experience ($t = -2.12$, $p = .04$) than elementary teachers.

Evaluator Attributes

For elementary teachers, the highest rated evaluator attributes were working relationship with the teacher ($M = 4.23$, $SD = .94$) and interpersonal manner ($M = 4.10$, $SD = 1.61$). The two highest rated attributes describing the evaluator by secondary teachers were interpersonal manner ($M = 4.18$, $SD = 1.05$) and patience ($M = 4.16$, $SD = .94$).

The data in Table IX show that there were statistically significant differences between elementary and secondary teachers for the following evaluator attributes:

1. Elementary teachers rated the credibility of the evaluator as a source of feedback as higher ($t = 3.24$, $p = <.01$) than secondary teachers.

2. Elementary teachers rated their working relationship with the evaluator as higher ($t = 2.32$, $p = .02$) than secondary teachers.

3. Secondary teachers rated their evaluators as having more experience with classrooms in general ($t = -2.15$, $p = .03$) than elementary teachers.

TABLE VIII
TEACHER PERCEPTIONS OF SELF
REGARDING EVALUATION

Attributes	Elementary Teachers (<u>n</u> = 147)		Secondary Teachers (<u>n</u> = 125)		<u>t</u>	<u>df</u>	<u>p</u>
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>			
Overall Quality	6.28	2.15	5.52	2.14	2.68	230	.008*
Overall Impact	3.62	2.49	3.22	2.38	1.26	229	.208
<u>TEACHER</u>							
Expectations of Self	4.57	.60	4.57	.56	.05	270	.961
Orientation to Risk Taking	3.75	.93	3.78	.82	-.26	270	.796
Orientation to Change	4.25	.82	3.92	.88	3.23	270	.001*
Experimentation in Classroom	4.06	.85	3.94	.76	1.27	269	.206
Openness to Criticism	3.74	.85	3.72	.84	.21	270	.835
Technical Knowledge	4.27	.64	4.20	.69	.88	269	.382
Knowledge of Subject	4.40	.59	4.62	.58	-3.02	268	.003*
Years Teaching Experience	4.22	.89	4.44	.83	-2.12	270	.035*
Helpfulness of Prior Evaluation	3.07	.98	2.89	.96	1.53	270	.128

* statistically significant

4. Elementary teachers rated the usefulness of the evaluator's suggestions for improvement as higher ($t = 2.15$, $p = .03$) than secondary teachers.

Evaluation Procedures

For elementary teachers, the highest rated attributes describing evaluation procedures were clarity of standards ($M = 4.04$, $SD = 1.06$) and observation of classroom ($M = 4.00$, $SD = 1.20$). Secondary teachers also rated clarity of standards ($M = 4.05$, $SD = 1.06$) the highest and endorsement of standards as appropriate for the teacher's classroom ($M = 3.86$, $SD = 1.13$) was rated the next highest.

The data in Table X show that there were statistically significant differences between elementary and secondary teachers for the following questionnaire items on evaluation procedures:

1. Elementary teachers rated the extent that classroom observations were used as a source of information for evaluation as higher ($t = 2.79$, $p = <.01$) than secondary teachers.
2. Elementary teachers had more formal observations per year ($t = 2.21$, $p = .03$) than secondary teachers.
3. Elementary teachers had more informal observations per year ($t = 2.44$, $p = .02$) than secondary teachers.

TABLE IX
TEACHER PERCEPTIONS OF THE EVALUATOR

	Elementary Teachers (<u>n</u> = 147)		Secondary Teachers (<u>n</u> = 125)				
Attributes	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>t</u>	<u>df</u>	<u>p</u>
<u>EVALUATOR</u>							
Credibility	4.01	1.05	3.58	1.14	3.24	270	.001*
Relationship to Teacher	4.23	.94	3.95	1.04	2.32	270	.021*
Level of Trust	4.02	1.57	4.04	1.10	-0.14	270	.887
Interpersonal Manner	4.10	1.61	4.18	1.05	-0.66	270	.512
Patience	4.01	1.08	4.16	.94	-1.24	270	.217
Flexibility	3.68	1.14	3.82	1.07	-1.07	270	.288
Technical Knowledge	4.09	.97	4.10	.98	-0.07	268	.948
Capacity to Model	3.38	1.13	3.45	1.04	-0.48	267	.634
Familiarity w/Classroom	3.23	1.15	3.29	1.22	-0.41	269	.683
Experience in Classrooms	3.61	1.16	3.91	1.13	-2.15	270	.032*
Usefulness of Suggestions	3.55	1.04	3.27	1.08	2.15	269	.033*
Persuasive Rationale	3.39	1.00	3.25	1.03	1.15	268	.250

* statistically significant

TABLE X
TEACHER PERCEPTIONS OF EVALUATION PROCEDURES

	Elementary Teachers (<u>n</u> = 147)		Secondary Teachers (<u>n</u> = 125)				
Attributes	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>t</u>	<u>df</u>	<u>p</u>
<u>PROCEDURES</u>							
Standards Communicated	3.61	1.09	3.67	.98	-0.47	270	.637
Clarity of Standards	4.04	1.06	4.05	1.06	-0.05	269	.957
Standards Endorsed	3.98	1.04	3.86	1.13	0.88	267	.379
Uniformity of Standards	2.88	1.42	2.73	1.37	0.89	265	.374
Observation of Classroom	4.00	1.20	3.60	1.16	2.79	268	.006*
Examination of Records	2.24	1.21	2.04	1.01	1.47	269	.144
Examination of Achievement	2.38	1.19	2.16	1.06	1.56	267	.119
# of Formal Observations	2.69	.87	2.44	.94	2.21	269	.028*
# of Informal Observations	2.27	.97	1.99	.86	2.44	269	.015*
* statistically significant							

Evaluation Feedback

For both elementary and secondary teachers, the highest rated attributes describing feedback were immediacy

of timing: elementary ($\bar{M} = 4.09$, $\underline{SD} = 1.17$) and secondary ($\bar{M} = 3.86$, $\underline{SD} = 1.16$) and feedback that focused on district teaching standards: elementary ($\bar{M} = 3.89$, $\underline{SD} = 1.08$) and secondary ($\bar{M} = 3.67$, $\underline{SD} = 1.17$).

The data in Table XI show that there were statistically significant differences between elementary and secondary teachers for the following questionnaire items on evaluation feedback:

1. Elementary teachers rated the amount of feedback information as higher ($t = 2.74$, $p = <.01$) than secondary teachers.
2. Elementary teachers received more formal feedback ($\underline{t} = 2.35$, $p = .02$) than secondary teachers.
3. Elementary teachers rated the depth of information provided in the feedback as higher ($\underline{t} = 2.40$, $p = .02$) than secondary teachers.
4. Elementary teachers rated the quality of ideas and suggestions contained in the feedback as higher ($\underline{t} = 2.46$, $p = .01$) than secondary teachers.

Evaluation Context

For both elementary and secondary teachers, the highest rated attributes describing evaluation context were the clarity of purposes for evaluation: elementary ($\bar{M} = 3.64$, $\underline{SD} = 1.13$), and secondary ($\bar{M} = 3.74$, $\underline{SD} = 1.03$) and the intended role of evaluation: elementary ($\bar{M} = 3.46$, $\underline{SD} = 1.16$) and secondary ($\bar{M} = 3.45$, $\underline{SD} = 1.14$).

TABLE XI
TEACHER PERCEPTIONS OF EVALUATION FEEDBACK

	Elementary Teachers (<u>n</u> = 147)		Secondary Teachers (<u>n</u> = 125)				
Attributes	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>t</u>	<u>df</u>	<u>p</u>
<u>FEEDBACK</u>							
Amount of Information	3.47	1.03	3.12	1.06	2.74	269	.006*
Frequency of Formal	2.77	1.22	2.44	1.09	2.35	269	.020*
Frequency of Informal	2.84	1.28	2.56	1.21	1.84	269	.067
Depth of Information	3.28	1.20	2.94	1.15	2.40	268	.017*
Quality of Ideas	3.38	1.16	3.04	1.10	2.46	267	.014*
Specificity of Information	3.47	1.25	3.23	1.20	1.65	268	.100
Nature of Information	3.87	1.07	3.61	1.11	1.95	266	.053
Appropriate Timing	4.09	1.17	3.86	1.16	1.59	268	.113
Focused on Standards	3.89	1.08	3.67	1.17	1.57	267	.118
* statistically significant							

The data in Table XII show that a statistically significant difference existed between elementary and secondary teachers for one attribute describing the

evaluation context. Elementary teachers rated the amount of time spent on the evaluation process as greater ($t = 4.32$, $p = <.01$) than secondary teachers.

TABLE XII
TEACHER PERCEPTIONS REGARDING THE
CONTEXT OF EVALUATION

	Elementary Teachers (<u>n</u> = 147)		Secondary Teachers (<u>n</u> = 125)				
Attributes	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>t</u>	<u>df</u>	<u>p</u>
<u>CONTEXT</u>							
Time Spent on Evaluation	3.12	.90	2.69	.73	4.32	269	.000*
Time for Professional Development	2.33	1.10	2.13	1.05	1.52	270	.130
Available Training	3.31	1.10	3.06	1.10	1.85	269	.065
Evaluation Purposes	3.64	1.13	3.74	1.03	-0.70	269	.486
Intended Role of Evaluation	3.46	1.16	3.45	1.14	0.09	265	.929
* statistically significant							

DIFFERENCES BETWEEN CLASSROOM AND NONCLASSROOM TEACHERS REGARDING EVALUATION

This section addresses the research question, "Is there a statistically significant difference between

permanent classroom and nonclassroom teachers in the perception of their evaluation experience? Tables XIII-XVII present the means, standard deviations, t value, degrees of freedom, and two-tail probability for each item. Data are reported separately for classroom and nonclassroom teachers. A two-tailed t test was used to determine whether the significance of difference between the means of classroom and nonclassroom teachers differed by more than chance at the $p < .05$ level of significance.

Table XIII reports the means of classroom teachers in regard to the overall quality of their most recent evaluation experience as higher ($M = 5.99$, $SD = 2.17$) than the means of nonclassroom teachers ($M = 5.43$, $SD = 2.43$). Nonclassroom teachers rated the overall impact of their last evaluation experience on professional practices higher ($M = 3.56$, $SD = 2.18$) as compared to classroom teachers ($M = 3.41$, $SD = 2.50$).

Teacher Attributes

For both classroom and nonclassroom teachers, the highest rated teacher attribute was expectations of self: classroom teachers ($M = 4.56$, $SD = .57$) and nonclassroom teachers ($M = 4.71$, $SD = .56$). The second-highest rated teacher attribute was knowledge of subject matter for classroom teachers ($M = 4.49$, $SD = .60$) and years of teaching experience for nonclassroom teachers ($M = 4.63$, $SD = .62$).

TABLE XIII
CLASSROOM AND NONCLASSROOM TEACHER PERCEPTIONS
OF SELF REGARDING EVALUATION

Attributes	Classroom Teachers (<u>n</u> = 243)		Nonclassroom Teachers (<u>n</u> = 41)		<u>t</u>	<u>df</u>	<u>p</u>
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>			
Overall Quality	5.99	2.17	5.43	2.43	-1.39	237	.166
Overall Impact	3.41	2.50	3.56	2.18	0.32	237	.750
<u>TEACHER</u>							
Expectations of Self	4.56	.57	4.71	.56	1.49	282	.138
Orientation to Risk Taking	3.76	.89	3.98	.82	1.47	282	.142
Orientation to Change	4.13	.86	4.00	.87	-0.88	282	.381
Experimentation in Classroom	4.03	.81	4.05	.90	0.15	281	.880
Openness to Criticism	3.74	.83	3.88	.93	0.96	282	.336
Technical Knowledge	4.24	.66	4.28	.64	0.29	281	.775
Knowledge of Subject	4.49	.60	4.58	.55	0.82	280	.411
Years Teaching Experience	4.28	.89	4.63	.62	3.18	71	.002*
Helpfulness of Prior Eval.	2.99	.98	2.93	.88	-0.40	282	.692

* statistically significant

The data in Table XIII show that there was a statistically significant difference between classroom and nonclassroom teachers for one teacher attribute. Nonclassroom teachers had more years of teaching experience ($t = 3.18$, $p = <.01$) than classroom teachers.

Evaluator Attributes

For classroom teachers, the highest rated evaluator attributes included working relationship with the teacher ($M = 4.15$, $SD = .96$) and interpersonal manner ($M = 4.13$, $SD = 1.11$). The two highest rated evaluator attributes scored by nonclassroom teachers were interpersonal manner ($M = 4.22$, $SD = 1.06$) and technical knowledge ($M = 4.03$, $SD = 1.06$). The data in Table XIV show that there was no significant difference between classroom and nonclassroom teachers for any of the evaluator attributes.

Evaluation Procedures

For both classroom and nonclassroom teachers, the highest rated evaluation procedure attributes were clarity of standards: classroom teachers ($M = 4.04$, $SD = 1.06$) and nonclassroom teachers ($M = 4.15$, $SD = .96$) and endorsement of standards as appropriate for the teacher's classroom: classroom teachers ($M = 3.95$, $SD = 1.03$) and nonclassroom teachers ($M = 3.78$, $SD = 1.27$).

The data in Table XV show that there were statistically significant differences between classroom and

* statistically significant

TABLE XV

[illegible]

nonclassroom teachers for the following two questionnaire items on evaluation procedures:

1. Nonclassroom teachers rated uniformity of standards higher ($t = 3.21$, $p = <.01$) than classroom teachers.

2. Classroom teachers had more formal observations per year ($t = -3.24$, $p = <.01$) than nonclassroom teachers.

Evaluation Feedback

For both classroom and nonclassroom teachers, the highest rated attribute describing evaluation feedback was immediacy of timing: classroom teachers ($M = 4.01$, $SD = 1.16$) and nonclassroom teachers ($M = 3.54$, $SD = 1.25$). The second-highest rated attribute for classroom teachers was feedback that focused on district teaching standards ($M = 3.84$, $SD = 1.11$). Nature of information ($M = 3.46$, $SD = 1.21$) was the second-highest rated attribute for non-classroom teachers.

The data in Table XVI show that there were statistically significant differences between classroom and nonclassroom teachers for the following questionnaire items on evaluation feedback:

1. Nonclassroom teachers received more informal feedback ($t = 2.82$, $p = <.01$) than classroom teachers.

2. Classroom teachers rated the immediacy of feedback as higher ($t = -2.34$, $p = .02$) than nonclassroom teachers.

3. Classroom teachers rated feedback that focused on district teaching standards as higher ($t = -2.66$, $p = <.01$) than nonclassroom teachers.

TABLE XVI
CLASSROOM AND NONCLASSROOM TEACHER PERCEPTIONS
OF EVALUATION FEEDBACK

	Classroom Teachers (<u>n</u> = 243)		Nonclassroom Teachers (<u>n</u> = 41)				
Attributes	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>t</u>	<u>df</u>	<u>p</u>
<u>FEEDBACK</u>							
Amount of Information	3.32	1.04	3.23	1.10	-0.51	281	.607
Frequency of Formal	2.65	1.19	2.38	1.03	-1.38	281	.168
Frequency of Informal	2.64	1.23	3.23	1.14	2.82	281	.005*
Depth of Information	3.13	1.19	2.98	1.19	-0.75	280	.451
Quality of Ideas	3.24	1.14	3.18	1.15	-0.31	279	.753
Specificity of Information	3.38	1.23	3.26	1.16	-0.61	279	.546
Nature of Information	3.79	1.06	3.46	1.21	-1.77	277	.078
Appropriateness of Timing	4.01	1.16	3.54	1.25	-2.34	279	.020*
Focused on Standards	3.84	1.11	3.33	1.08	-2.66	278	.008*

* statistically significant

Evaluation Context

Classroom teachers rated clarity of evaluation purposes the highest ($\bar{M} = 3.71$, $SD = 1.05$), and nonclassroom teachers rated the intended role of evaluation the highest ($\bar{M} = 3.65$, $SD = 1.21$). The second-highest rated attributes describing evaluation context were intended role of evaluation for classroom teachers ($\bar{M} = 3.43$, $SD = 1.14$) and clarity of evaluation purposes for nonclassroom teachers ($\bar{M} = 3.54$, $SD = 1.27$).

The data in Table XVII show that the only statistically significant difference between classroom and nonclassroom teachers was the amount of time allotted during the teaching day for professional development. Nonclassroom teachers rated that attribute as higher ($t = 2.08$, $p = .04$) than classroom teachers.

SUMMARY OF RESULTS

Results are reported to furnish evidence for accepting or rejecting the four hypothesis statements. The data supported three of the four research hypotheses.

Hypothesis 1. There are significant relationships between teachers' perceptions of the quality of their evaluation experience and the specific attributes of the evaluation process.

The hypothesis that there are significant relationships between teachers' perceptions of the quality of their evaluation experience and the specific attributes of the evaluation process was not rejected. The null hypothesis

TABLE XVII
CLASSROOM AND NONCLASSROOM TEACHER PERCEPTIONS
REGARDING THE CONTEXT OF EVALUATION

	Classroom Teachers (<u>n</u> = 243)		Nonclassroom Teachers (<u>n</u> = 41)				
Attributes	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>t</u>	<u>df</u>	<u>p</u>
<u>CONTEXT</u>							
Time Spent on Evaluation	2.95	.87	2.74	.60	-1.91	67	.060
Time for Professional Development	2.20	1.08	2.59	1.16	2.08	282	.039*
Available Training	3.23	1.09	3.10	1.24	-0.71	281	.480
Clarity of Evaluation Purposes	3.71	1.05	3.54	1.27	-0.95	281	.344
Intended Role of Evaluation	3.43	1.14	3.65	1.21	1.07	276	.286
* statistically significant							

was rejected. There was a substantial correlation ($r = .60$ or higher) with the overall quality of evaluation and the following seven attributes describing teacher evaluation:

1. Quality of ideas and suggestions contained in the feedback;
2. Usefulness of the evaluator's suggestions for improvement;

3. Depth of information provided in the feedback;
4. Persuasiveness of the evaluator's rationale for suggestions;
5. Specificity of information provided in the feedback;
6. Amount of information received in the feedback;
7. Credibility of the evaluator as a source of feedback.

Hypothesis 2. There are significant relationships between teachers' perceptions of the impact of their evaluation experience and the specific attributes of the evaluation process.

The hypothesis that there are significant relationships between teachers' perceptions of the impact of their evaluation experience and the specific attributes of the evaluation process was rejected. There are no significant relationships between teachers' perceptions of the impact of their evaluation experience and the specific attributes of the evaluation process. The null hypothesis was retained.

Hypothesis 3. There are significant differences between permanent elementary and secondary teachers in the perception of their evaluation experience.

The hypothesis that there are significant differences between permanent elementary and secondary teachers in the perception of their evaluation experience was not rejected at the .05 level of significance. The null hypothesis was rejected. There was a significant difference between the perceptions of elementary and secondary teachers for the following 16 teacher evaluation attributes:

1. Elementary teachers perceive the overall quality of their most recent evaluation experience as higher than do secondary teachers.

2. Elementary teachers perceive themselves as more flexible and oriented to change than do secondary teachers.

3. Secondary teachers feel stronger in their knowledge of subject matter than do elementary teachers.

4. Secondary teachers have more years of teaching experience than do elementary teachers.

5. Elementary teachers perceive evaluators as more credible than do secondary teachers.

6. Elementary teachers perceive their evaluator as more of a helper than do secondary teachers.

7. Secondary teachers feel that their evaluator has had more experience with classrooms in general than do elementary teachers.

8. Elementary teachers believe that their evaluators' suggestions are more useful than do secondary teachers.

9. Elementary teachers believe that classroom observations are used as a source of performance information to a greater extent than do secondary teachers.

10. Elementary teachers have more formal observations per year than do secondary teachers.

11. Elementary teachers have more informal observations per year than do secondary teachers.

12. Elementary teachers believe they receive more feedback than do secondary teachers.

13. Elementary teachers receive formal feedback more frequently than do secondary teachers.

14. Elementary teachers believe they receive more in-depth feedback than do secondary teachers.

15. Elementary teachers believe the quality of the ideas and suggestions contained in the feedback is higher than do secondary teachers.

16. Elementary teachers perceive that the amount of time spent on the evaluation process is greater than do secondary teachers.

Hypothesis 4. There are significant differences between permanent classroom and nonclassroom teachers in the perception of their evaluation experience.

The hypothesis that there are significant differences between permanent classroom and nonclassroom teachers in the perception of their evaluation experience was not rejected at the .05 level of significance. The null hypothesis was rejected. There were significant differences between the perceptions of classroom and nonclassroom teachers for the following seven teacher evaluation attributes:

1. Nonclassroom teachers have more years of teaching experience than do classroom teachers.

2. Nonclassroom teachers believe that the standards by which they are evaluated are unique to their situation.

Classroom teachers feel less strongly about the uniqueness of the standards as applied to their situation.

3. Classroom teachers have more formal observations per year than do nonclassroom teachers.

4. Nonclassroom teachers believe they receive informal feedback more frequently than do classroom teachers.

5. Classroom teachers believe that the timing of their feedback is more immediate than do nonclassroom teachers.

6. Classroom teachers believe that their feedback focuses more on district teaching standards than do nonclassroom teachers.

7. Nonclassroom teachers believe that the amount of time they receive during the teaching day for professional development is greater than do classroom teachers.

The above summary includes the significant relationships between overall quality and impact of the evaluation experience and the 44 evaluation attributes. In addition, the summary addresses all the significant differences between elementary and secondary teachers as well as classroom and nonclassroom teachers in their perceptions of teacher evaluation. The following chapter will include a discussion of the conclusions and implications from this study.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

INTRODUCTION

This study examined teacher perceptions of evaluation in regard to professional growth and improvement of instruction. The following sections will be covered in this chapter: (a) Summary; (b) Conclusions and Implications; (c) Limitations of the Study; and (d) Recommendations.

SUMMARY

The purpose of this study was to measure the perceptions of teachers regarding their evaluation experience and to determine whether teachers perceived any significant relationships between the attributes of teacher evaluation and its quality and impact on teacher growth. The study also looked at whether significant differences existed between elementary and secondary teachers' perceptions of evaluation as well as between classroom and nonclassroom teachers' perceptions of evaluation. Four major research questions were addressed:

1. Is there a relationship between specific attributes of teacher evaluation and its perceived quality by teachers?

2. Is there a relationship between specific attributes of teacher evaluation and its perceived impact by teachers?

3. Is there a statistically significant difference between permanent elementary and secondary teachers in the perception of their evaluation experience?

4. Is there a statistically significant difference between permanent classroom and nonclassroom teachers in the perception of their evaluation experience?

To answer these questions, the Teacher Evaluation Profile (TEP), an assessment instrument developed by Stiggins and Duke from Northwest Regional Educational Laboratory, was used to measure respondents' perceptions of their evaluation experience. The TEP questionnaire is a dependable and consistent instrument with an internal consistency reliability of .93 (Stiggins & Duke, 1986). First, the questionnaire asked teachers to rate the overall quality and impact of their most recent evaluation experience by using a 10-point scale. Teachers also responded to 44 specific attributes that described their most recent evaluation experience by using a 5-point scale. Each item had its own descriptive rating scale.

The sample for this study was composed of 402 teachers who were randomly selected from 1,081 permanent teachers in one suburban school district. The sample included

elementary and secondary teachers as well as classroom and nonclassroom teachers.

Data were reported in terms of frequency distributions, means, and standard deviations. Data analysis consisted of correlational analyses and analysis of variance. An alpha of .05 was used to determine statistical significance.

Results

Results are reported to furnish evidence for accepting or rejecting the four hypotheses statements. The data from this study supported three of the four hypotheses.

Hypothesis 1. There are significant relationships between teachers' perceptions of the quality of their evaluation experience and the specific attributes of the evaluation process.

The hypothesis that there are significant relationships between teachers' perceptions of the quality of their evaluation experience and the specific attributes of the evaluation process was not rejected. The null hypothesis was rejected. There is a substantial correlation ($r = .60$ to $.80$) between teachers' perceptions of the quality of their evaluation and 7 out of 44 attributes describing the evaluation process.

Hypothesis 2. There are significant relationships between teachers' perceptions of the impact of their evaluation experience and the specific attributes of the evaluation process.

The hypothesis that there are significant relationships between teachers' perceptions of the impact of their evaluation experience and the specific attributes of the evaluation process was rejected. There are no significant relationships between teachers' perceptions of the impact of their evaluation experience and the specific attributes of the evaluation process. The null hypothesis was retained.

Hypothesis 3. There are significant differences between permanent elementary and secondary teachers in the perception of their evaluation experience.

The hypothesis that there are significant differences between permanent elementary and secondary teachers in the perception of their evaluation experience was not rejected at the .05 level of significance. The null hypothesis was rejected. There are significant differences between the perceptions of elementary and secondary teachers for 15 out of 44 evaluation attributes regarding the overall quality of their evaluation experience.

Hypothesis 4. There are significant differences between permanent classroom and nonclassroom teachers in the perception of their evaluation experience.

The hypothesis that there are significant differences between permanent classroom and nonclassroom teachers in the perception of their evaluation experience was not rejected at the .05 level of significance. The null hypothesis was rejected. There are significant differences between the

perceptions of classroom and nonclassroom teachers for 7 out of 44 teacher evaluation attributes.

Teacher Attributes. Teachers rated attributes describing themselves consistently higher than other attributes on the TEP. The data suggest that teachers believe they demand a great deal of themselves as professionals. As a group, teachers tend to view themselves as strong in knowledge of subject matter, knowledge of the technical aspects of teaching, and orientation to change.

The findings indicate that attributes describing the teacher appear unrelated to perceptions regarding the overall quality and impact of evaluation. With the exception of how teachers viewed their prior experience with evaluation, there appears to be a negligible relationship between specific attributes of the teacher and the outcomes of evaluation.

Elementary teachers see themselves as significantly more oriented to change than do secondary teachers. Secondary teachers have significantly more years of teaching experience and perceive themselves as significantly stronger in knowledge of subject matter than do elementary teachers. Nonclassroom teachers have significantly more years of teaching experience than do classroom teachers.

Evaluator Attributes. Teachers tend to view their evaluator as nonthreatening, trustworthy, and patient.

Evaluators are seen in the role of a helper and as knowledgeable about the technical aspects of teaching.

Teachers perceive the following evaluator attributes as substantially related to the overall quality of their evaluation: the credibility of the evaluator as a source of feedback, the usefulness of the evaluator's suggestions for improvement, and the persuasiveness of the evaluator's rationale for suggestions.

Elementary teachers see their evaluator as significantly more credible and more of a helper than do secondary teachers. In addition, elementary teachers find the evaluator's suggestions for improvement as significantly more useful than secondary teachers find them. However, secondary teachers perceive their evaluators as having more experience with classrooms, in general, than do elementary teachers.

Evaluation Procedures. Teachers believe that the teaching standards used for evaluation are clear. They also feel that classroom or school records and student achievement are not considered as part of their evaluation. Most teachers are observed formally no more than two times per year and are observed informally less than once per month. There appears to be only a moderate to negligible relationship between attributes describing evaluation procedures and teachers' perceptions of the overall quality and impact of evaluation.

When comparing elementary teachers with secondary teachers, elementary teachers perceive that classroom observations are used significantly more often as a source of information for evaluation than do secondary teachers. In addition, elementary teachers have significantly more formal and informal observations per year than do secondary teachers.

Nonclassroom teachers perceive the standards by which they are evaluated as more specifically designed for them than do classroom teachers. Classroom teachers have significantly more formal observations per year than do nonclassroom teachers.

Evaluation Feedback. Teachers believe that the feedback they receive is usually immediate, descriptive in nature, and focused on district teaching standards. Teachers perceive the following evaluation feedback attributes as substantially related to the overall quality of their evaluation: amount and depth of information received in the feedback, the quality of ideas and suggestions contained in the feedback, and the specificity of information provided in the feedback.

Elementary teachers perceive that the amount of information they receive, the frequency of formal feedback, the depth of information provided in the feedback, and the quality of ideas and suggestions contained in the feedback are significantly greater than do secondary teachers.

Nonclassroom teachers receive significantly more informal feedback than do classroom teachers. Classroom teachers perceive the timing of feedback as more immediate and the focus of feedback as more of a reflection of district teaching standards than do nonclassroom teachers.

Evaluation Context. Teachers generally believe that policy statements regarding the purpose for evaluation are clear. However, only a small portion of teachers feel that they receive much time during the teaching day for professional development. There appears to be only a moderate to negligible relationship between attributes describing the evaluation context and the overall quality and impact of evaluation.

Elementary teachers perceive the amount of time spent on the evaluation process as significantly greater than secondary teachers perceive it. Nonclassroom teachers feel that they have significantly more time during the teaching day for professional development than do classroom teachers.

Subscale Data. The TEP describes 44 distinct attributes of teacher evaluation which are grouped into the following five subscales: (a) Teacher Attributes; (b) Evaluator Attributes; (c) Evaluation Procedures; (d) Evaluation Feedback; and (e) Evaluation Context. The subscales were analyzed for means, standard deviations, and correlations with the overall quality and impact of teacher evaluation.

It appears that when the respondents rated the attributes within each subscale, they tended to collapse the individual items and consider each subscale holistically. The results indicate that teachers feel the most positive about attributes describing themselves and the evaluator. When the five subscales were correlated with overall quality and impact, the subscales describing evaluation feedback and the evaluator had a substantial correlation with teachers' feelings about the overall quality of evaluation.

CONCLUSIONS AND IMPLICATIONS

A major outcome of this study shows that teachers appear to judge the quality of their evaluation experience based on the attributes of the person who evaluates them and the feedback they receive. The results indicate that teacher attributes, evaluation procedures, and the context of the evaluation system do not have a strong influence on how teachers perceive the quality of their evaluation.

The data show that teachers determine the quality of evaluation based on the strength of the ideas and suggestions contained in the feedback. They look for the depth of information received in the feedback as well as the specificity and amount of information they receive. The frequency of formal feedback appears moderately related to how teachers perceive the quality of their evaluation.

Since the data indicate that effective feedback is related to the quality of evaluation, one implication might be that the evaluation system should focus on the specific attributes that describe feedback. Considerations include feedback that is provided in a timely manner after a classroom observation and data that are specific and tied closely to the observation. If multiple sources of data are used, they might provide a more complete picture of teaching performance. Feedback should be perceived as an occasion for reflection and growth.

The findings indicate that evaluators impact the quality of evaluation when they share novel and insightful information with teachers about their teaching. Teachers appear to value someone who gives useful suggestions for improvement, has a persuasive rationale for suggestions, and is a credible source of feedback.

Since teachers appear to perceive the person who evaluates them as having a substantial impact on the quality of their evaluation, then perhaps there needs to be a focus on ways to increase the effectiveness of the evaluator. It is the opinion of this researcher that evaluators should receive ongoing inservice to develop their communication and interpersonal skills as well as their ability to give meaningful feedback. If the level of the evaluator's credibility as a source of ideas for effective teaching is raised, then perhaps the perceived quality of evaluation

will also increase. Furthermore, since the data indicate the importance of the evaluator in teacher evaluation, perhaps the use of peer panels or multiple judges might be appropriate in order to balance out the biases of individuals.

LIMITATIONS OF THE STUDY

A self-administered questionnaire was used to collect data for this study. An advantage to using the Teacher Evaluation Profile was that closed-ended questions required respondents to answer within a given format. This uniformity of responses made the questionnaire easier to process. However, this type of survey instrument relies on the respondent to state accurately what he or she feels. Furthermore, a descriptive scale poses a problem in that teachers might assign different values to the words used to describe the 44 attributes of the evaluation process.

Collecting only one line of evidence limits the amount and kind of data that can be obtained. Open-ended questions at the end of the survey would have provided teachers an opportunity to share their concerns about evaluation and make recommendations for change.

In order to maintain confidentiality, this study did not attempt to find the reasons for the responses or verify the accuracy of the respondents' perceptions. An advantage to interviewing teachers would have been the ability to ask

follow-up questions. A disadvantage to the interview method might have been the personal presence of the interviewer which could have reduced the respondent's spontaneity. When anonymity is lost, the willingness to be completely frank and honest might be reduced when dealing with a subject such as teacher evaluation.

Interviews with the superintendent and director of personnel would have provided information regarding the district's governing values toward teacher evaluation. Interviews with the president of the local education association and the teachers' contract representative would have provided insight on how those involved in the negotiation of evaluation procedures perceive the evaluation process.

Finally, this study was limited to permanent elementary and secondary teachers from one suburban school district. The results do not necessarily reflect the perceptions of all teachers.

RECOMMENDATIONS

Based on a review of the literature and the conclusions and implications from this study, it is recommended that evaluation methods supportive of professional growth for competent, permanent teachers include the following characteristics:

1. Differentiated evaluation where formative evaluation remains distinct from summative evaluation;
2. Opportunities for teacher involvement in developing the evaluation system and functioning within it;
3. Evaluation procedures that focus on well-defined effective teacher characteristics;
4. Inservice for evaluators on effective communication skills and interpersonal relations;
5. Inservice for teachers and evaluators on the evaluation process including the collection and analysis of descriptive data on teaching;
6. Cooperative identification of goals by teachers and supervisors which contribute to the professional growth and development of teachers;
7. Coordination of staff development resources and activities with the identified goals of teachers;
8. Opportunities for collegial supervision, teacher mentors, peer coaching, peer review, and self-evaluation;
9. Multiple sources of data to provide a more complete picture of teaching performance;
10. Feedback activities about teaching performance that are timely, specific, and useful.

This researcher suggests that districts consider alternative forms of evaluation to match the employment status, level of experience, and degree of success teachers are experiencing. One recommendation made by Stiggins and

Duke (1986) is to establish a separate four-year evaluation system for competent, permanent teachers. Teachers would be formally evaluated on specific criteria during the first year of the cycle. Those teachers viewed as successful during the first year would focus on professional development and instructional improvement during years two through four. In order to encourage risk taking, no official records would be kept and the supervisor would act in a supportive, collegial role. This cycle would provide teachers time to study, practice, internalize, and refine their new skills as well as present opportunities for working with peers.

The research supports the need to improve current teacher evaluation practices (Levin, 1979; Petersen, 1989). Traditionally, classroom observation has been the predominant method for collecting data about teaching (McGreal, 1983). Teacher evaluation procedures have generally relied on the administrative checklist (Lewis, 1982; Wood & Pohland, 1979). This researcher recommends that in addition to some form of clinical supervision, other data sources need to be used for teacher evaluation. A multiple data source model might include student evaluations, student achievement tests, artifact collections, teacher competency tests, teacher portfolios, teacher self-evaluations, and peer evaluations.

The effects of teacher evaluation on minority teachers are complex and often reflect biases. The teacher portfolio, if designed appropriately, can reveal strengths and give credit to the work of teachers that ordinarily might not appear in other types of assessment. Portfolios should be a vehicle for teachers to reveal their teaching practices and display diversified samples of work (Minority Participation, 1988).

The school district in this study has already implemented many of the criteria recommended in the literature for establishing an effective teacher evaluation system. The data indicate that teachers in this study appear satisfied with the overall quality of current evaluation practices.

Ninety-eight percent of the respondents in this study have at least six years of teaching experience. Berman and McLaughlin (1978) found evidence that many teachers become less effective as their length of teaching experience increases and that the average teacher is most productive from approximately the third to sixth year of teaching. Based on the research of Berman and McLaughlin (1978) and the results of this study, it is recommended that this district continue to focus on staff development activities that promote the professional growth and instructional improvement of teachers. However, it is this researcher's opinion that a more formal process needs to be established

that coordinates staff development resources and activities with the identified goals of teachers involved in formative evaluation.

Future Study

Teacher evaluation needs further research in order to clarify how it can increase the overall quality and impact of evaluation and promote professional growth. Teachers in this study tended to rate the overall quality of their evaluations higher than the overall impact of their evaluations. These results match the findings of Stiggins and Duke (1986) in their analyses of school districts they surveyed using the Teacher Evaluation Profile (TEP). Based on this outcome, related research questions deserving investigation include: Why do teachers tend to rate the quality of their evaluations higher than the impact of their evaluations? What are some effective ways of reporting the impact of evaluation on teachers?

The TEP measures 44 key attributes that contribute to the quality and impact of teacher evaluation. In this study, the attributes describing the teacher have, for the most part, a negligible correlation to overall quality and impact of teacher evaluation. These results also match the findings of Stiggins and Duke (1986). Further studies of effective teacher evaluation systems need to address the following questions: What attributes of the teacher have a

strong relationship to perceived quality and impact of evaluation? Are there other attributes of an effective teacher evaluation system that contribute to its overall quality and impact?

Preliminary research in other studies suggests that teachers might exhibit different teaching behaviors and skills based on the purpose of instruction, the subject matter, and grade level of the students (Shulman, 1987; Stodolsky, 1984). Berman and McLaughlin (1978) found differences in the characteristics of elementary and secondary schools in terms of their orientation to change and to subject matter. The results of this study indicate that teachers have different perceptions about evaluation based on their level (elementary or secondary) as well as their role (classroom or nonclassroom).

More research is needed to understand why significant differences exist between elementary and secondary teachers as well as classroom and nonclassroom teachers in their perceptions of the evaluation process. A future study might examine the factors that contribute to variations in teaching and determine ways teaching is seen as context-bound. We need to answer the following questions: How should teacher evaluation systems address the importance of context? How might the interests of elementary and secondary teachers as well as those of classroom and

nonclassroom teachers be addressed in growth-oriented teacher evaluation?

McGreal (1988) reported increasing evidence that the role of the superintendent and central office staff is crucial to the successful implementation of instructional improvement efforts. Although this study did not address this attribute of the evaluation context, further research might explore the impact of the superintendent and central office staff in determining the governing values of a district in regard to teacher evaluation and how that policy is supported at the school level.

This study should be replicated in urban, rural, and other suburban school settings. Researchers must determine the key attributes of effective evaluation in other systems and settings and make recommendations as to how growth-oriented teacher evaluation can be best implemented.

Increasingly, the public and many educators feel that the key to improving schools is through developing the skills of teachers. It is the belief of this researcher that growth-oriented teacher evaluation has the potential to improve the quality of education by promoting professional development and improving instructional skills.

REFERENCES

- Acheson, K. A., & Gall, M. D. (1980). Techniques in the clinical supervision of teachers. New York: Longman.
- Anderson, R. H. (1982). Creating a future for supervision. In T. J. Sergiovanni (Ed.), Supervision of teaching (pp. 181-190). Alexandria, VA: Association for Supervision and Curriculum Development.
- Armiger, M. L. (1981). The political realities of teacher evaluation. In J. Millman (Ed.), Handbook of teacher evaluation (pp. 292-302). Beverly Hills, CA: Sage Publications.
- Babbie, E. R. (1973). Survey research methods. Belmont, CA: Wadsworth Publishing.
- Bacharach, S. B., Conley, S. C., & Shedd, J. B. (1987). A career developmental framework for evaluating teachers as decision-makers. Journal of Personnel Evaluation in Education, 1, 181-194.
- Barber, L. (1987). Teacher evaluation. Unpublished report submitted to the Teacher Quality Program, Indiana Department of Education.
- Beckham, J. C. (1981). Legal aspects of teacher evaluation. Topeka, KS: National Organization on Legal Problems of Education.
- Bell, T. H. (1986). A nation still at risk. Principal, 66(1), 5-6.
- Berman, P., & McLaughlin, M. W. (1978). Federal programs supporting educational change: Vol. VIII. Implementing and sustaining innovations. Santa Monica, CA: The Rand Corporation.
- Best, J. W., & Kahn, J. V. (1986). Research in education. Englewood Cliffs, NJ: Prentice-Hall.
- Blumberg, A., & Greenfield, W. (1986). The effective principal. Boston, MA: Allyn and Bacon.
- Bolton, D. L. (1973). Selection and evaluation of teachers. Berkeley, CA: McCutchan.

- Borg, W. R., & Gall, M. D. (1983). Educational research. New York: Longman.
- Borich, G. D., & Fenton, K. S. (1977). The appraisal of teaching: Concepts and process. Menlo Park, CA: Addison-Wesley.
- Bridges, E. M., & Graves, B. (1984). Managing the incompetent teacher. Stanford, CA: Stanford University, Institute for Research on Educational Finance and Governance.
- Brookover, W. B., Beamer, L., Efthim, H., Hathaway, D., Lezotte, L., Miller, S., Passalacqua, J., & Tornatzky, L. (1982). Creating effective schools: An in-service program for enhancing school learning climate and achievement. Holmes Beach, FL: Learning Publications.
- Brookover, W. B., & Lezotte, L. W. (1979). Changes in school characteristics coincident with changes in student achievement (Occasional Paper No. 17). East Lansing: Michigan State University, Institute for Research on Teaching, College of Education.
- Carey, W. C. (1983). Redesigning your teacher evaluation system. Salem, OR: Options Press.
- Cawelti, G. (1982). Training for effective school administrators. Educational Leadership, 39, 324-329.
- Cawelti, G., & Reavis, C. (1980). How well are we providing instructional improvement services? Educational Leadership, 38, 236-240.
- Chunn, G. F. (1985). Perceptions of teachers and principals concerning behaviors and attitudes that contribute to an effective supervisory cycle. Dissertation Abstracts International, 46, 2494-A. (University Microfilms No. DA8524040)
- Cogan, M. L. (1973). Clinical supervision. Boston: Houghton Mifflin.
- Darling-Hammond, L., Wise, A. E., & Pease, S. R. (1983). Teacher evaluation in the organizational context: A review of the literature. Review of Educational Research, 53, 285-328.

- Drake, J. M. (1984, February). Improving teacher performance through evaluation and supervision. Paper presented at the Annual Meeting of the National Association of Secondary School Principals, Las Vegas, NV.
- Duke, D. L. (1982). Leadership functions and instructional effectiveness. Portland, OR: Lewis and Clark College.
- Duke, D. L. (1985). What is the nature of educational excellence and should we try to measure it? Phi Delta Kappan, 66, 671-674.
- Duke, D. L., & Stiggins, R. J. (1986). Teacher evaluation: Five keys to growth. Washington, DC: National Education Association Professional Library.
- Edmonds, R. (1979). Effective schools for the urban poor. Educational Leadership, 37, 15-24.
- Edmonds, R. R. (1981). Making public schools effective. Social Policy, 12(2), 56-63.
- Educational Research Service. (1978). Evaluating teacher performance. Arlington, VA: Educational Research Service.
- Ellis, T. I. (1984). Teacher competency: What administrators can do. Eugene, OR: ERIC Clearinghouse on Educational Management. (ERIC Document Reproduction Service No. ED 259 452)
- Flanders, N. A. (1976). Interaction analysis and clinical supervision. Journal of Research and Development in Education, 9(2), 47-57.
- Fullan, M. (1982). The meaning of educational change. New York: Teachers College Press.
- Gallup, A. M., & Clark, D. L. (1987). The 19th annual Gallup poll of the public's attitudes toward the public schools. Phi Delta Kappan, 69, 17-30.
- Gallup, G. (1979). The eleventh annual Gallup poll of the public's attitudes toward the public schools. Phi Delta Kappan, 61, 33-46.
- Garman, N. B. (1982). The clinical approach to supervision. In T. J. Sergiovanni (Ed.), Supervision of teaching (pp. 35-52). Alexandria, VA: Association for Supervision and Curriculum Development.

- Garman, N. B. (1986). Reflection, the heart of clinical supervision: A modern rationale for professional practice. Journal of Curriculum and Supervision, 2, 1-24.
- Glickman, C. D. (1987). Instructional improvement and the K-8 principal. Alexandria, VA: National Association of Elementary School Principals. (EDRS Document Reproduction Service No. ED 280 146)
- Good, T. L., & Brophy, J. E. (1984). Looking in classrooms. San Francisco, CA: Harper & Row.
- Goodlad, J. I. (1984). A place called school. San Francisco, CA: McGraw-Hill.
- Grippa, P. A. (1987). Developing an evaluation system for principals: A project and a process. Dissertation Abstracts International, 49, 1014-A. (University Microfilms No. DA 8808341)
- Gudridge, B. M. (1980). Teacher competency: Problems and solutions. Sacramento, CA: American Association of School Administrators.
- Guthrie, H. D., & Willower, D. J. (1973). The ceremonial congratulation: An analysis of principals' observation reports of classroom teaching. The High School Journal, 56, 284-290.
- Hackman, J. R., & Lawler, E. E. (1971). Employee reactions to job characteristics (monograph). Journal of Applied Psychology, 55, 259-286.
- Hackman, J. R., & Oldham, G. R. (1980). Work redesign. Menlo Park, CA: Addison-Wesley.
- Ham, M. C., & Oja, S. N. (1987, February). A collaborative approach to leadership in supervision. Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education, Arlington, VA.
- Harris, B. M. (1986). Developmental teacher evaluation. Boston, MA: Allyn and Bacon.
- Herzberg, F. (1968). One more time: How do you motivate employees? Harvard Business Review, 46, 53-62.
- Howe, H., II. (1983). Education moves to center stage: An overview of recent studies. Phi Delta Kappan, 65, 167-172.

- Howell, B. (1981). Profile of the principalship. Educational Leadership, 38, 333-336.
- Huff, S., Lake, D., & Schoalman, M. D. (1982). Principal differences: Excellence in school leadership and management. Boston, MA: McBer and Company.
- Hunter, M. (1988). Effecting a reconciliation between supervision and evaluation--A reply to Popham. Journal of Personnel Evaluation in Education, 1, 275-279.
- Hunter, M., & Russell, D. (1987, Summer). Critical attributes of a staff development program to increase instructional effectiveness. Teaching-Coaching-Supervising Newsletter, p. 1.
- Kauchak, D., Peterson, K., & Driscoll, A. (1985). An interview study of teachers' attitudes toward teacher evaluation practices. Journal of Research and Development in Education, 19(1), 32-37.
- Kowalski, J. P. S. (1978). Evaluating teacher performance. Arlington, VA: Educational Research Service.
- Krajewski, R. J., & Anderson, R. H. (1980). Goldhammer's clinical supervision a decade later. Educational Leadership, 37, 420-423.
- Levin, B. (1979). Teacher evaluation--A review of research. Educational Leadership, 37, 240-245.
- Lewis, A. C. (1982). Evaluating educational personnel. Arlington, VA: American Association of School Administrators.
- Lezotte, L. W., & Bancroft, B. A. (1985). Growing use of the effective schools model for school improvement. Educational Leadership, 42(6), 23-27.
- Lipham, J. M. (1982). Leadership of the principal for educational improvement. Portland, OR: Northwest Regional Educational Laboratory.
- Little, J. W. (1981, April). The power of organizational setting: School norms and staff development. Paper presented at the Annual Meeting of the American Educational Research Association, Los Angeles, CA.
- Lortie, D. C. (1975). Schoolteacher: A sociological study. Chicago, IL: University of Chicago Press.

- Lyman, L. (1987, March). Principals and teachers: Collaboration to improve instructional supervision (building trust, fostering collaboration, encouraging collegiality). Paper presented at the Annual Meeting of the Association for Supervision and Curriculum Development, Orlando, FL.
- Macdonald, J. B. (1981). Planning for the evaluation of teaching. In W. R. Duckett (Ed.), Planning for the evaluation of teaching (pp. 2-12). Bloomington, IN: Phi Delta Kappa.
- Mackenzie, R. A. (1970). Managing time at the top. New York: The Presidents Association.
- March, J. G. (1978). American public school administration: A short analysis. School Review, 86, 217-250.
- McGreal, T. L. (1982). Effective teacher evaluation systems. Educational Leadership, 39, 303-305.
- McGreal, T. L. (1983). Successful teacher evaluation. Alexandria, VA: Association for Supervision and Curriculum Development.
- McGreal, T. L. (1988). Evaluation for enhancing instruction: Linking teacher evaluation and staff development. In S. J. Stanley & W. J. Popham (Eds.), Teacher evaluation: Six prescriptions for success (pp. 1-29). USA: Association for Supervision and Curriculum Development.
- McLaughlin, M. W., & Pfeifer, R. S. (1986). Teacher evaluation: Learning for improvement and accountability and case studies. Stanford, CA: Stanford University, Stanford Education Policy Institute. (EDRS Document Reproduction Service No. ED 271 839)
- McNeil, J. D. (1981). Politics of teacher evaluation. In Jason Millman (Ed.), Handbook of teacher evaluation (pp. 272-291). Beverly Hills, CA: Sage Publications.
- Medley, D. M., & Coker, H. (1987). The accuracy of principals' judgments of teacher performance. The Journal of Educational Research, 80, 242-247.
- Mickler, M. L. (1985). Interviews with Ralph W. Tyler. The Educational Forum, 50(1), 23-46.

- Miller, L., & Lieberman, A. (1982). School leadership between the cracks. Educational Leadership, 39, 362-367.
- Millman, J. (Ed.). (1981). Handbook of teacher evaluation. Beverly Hills, CA: Sage Publications.
- Minority participation in teaching. (1988, Fall-Winter). The Teacher Assessment Project Newsletter, pp. 7-9.
- Murphy, J. (1987). Teacher evaluation: A comprehensive framework for supervisors. Journal of Personnel Evaluation in Education, 1, 157-180.
- National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. Washington, DC: U.S. Government Printing Office.
- Natriello, G. (1983). Evaluation frequency, teacher influence, and the internalization of evaluation processes: A review of six studies using the theory of evaluation and authority. Eugene: University of Oregon, Center for Educational Policy and Management. (EDRS Document Reproduction Service No. 242 050)
- Natriello, G., & Dornbusch, S. M. (1981). Pitfalls in the evaluation of teachers by principals. Administrator's Notebook, 29(6), 1-4.
- Natriello, G., Hoag, M., Deal, T. E., & Dornbusch, S. M. (1977). A summary of the recent literature on the evaluation of principals, teachers, and students (Occasional Paper No. 18). Stanford, CA: Stanford Center for Research and Development in Teaching.
- Northwest Regional Educational Laboratory. (1988). The teacher evaluation profile (Contract No. 400-86-0006). Portland, OR: Author.
- Peterson, K. (1984). Technical problems in teacher evaluation. Journal of Research and Development in Education, 7(4), 62-70.
- Peterson, K. D. (1978). The principal's tasks. Administrator's Notebook, 26(8), 1-4.
- Pfeifer, R. S. (1986, April). Integrating teacher evaluation and staff development: An organizational approach. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

- Popham, W. J. (1986). Teacher evaluations: Mission impossible. Principal, 65(4), 56-58.
- Popham, W. J. (1988). The dysfunctional marriage of formative and summative teacher evaluation. Teacher Evaluation Journal of Personnel Evaluation in Education, 1, 269-273.
- Robinson, J. J. (1978). The observation report: A help or a nuisance? National Association of Secondary School Principals Bulletin, 62(416), 22-26.
- Rothberg, R. A., & Buchanan, L. L. (1981). Teacher perceptions of teacher assessment. Phi Delta Kappan, 62, 527.
- Rutherford, W. L. (1985). School principals as effective leaders. Phi Delta Kappan, 67, 31-34.
- Scriven, M. (1981). Summative teacher evaluation. In J. Millman (Ed.), Handbook of teacher evaluation (pp. 244-271). Beverly Hills, CA: Sage Publications.
- Scriven, M. (1983). Evaluation ideologies. In G. F. Madaus, M. Scriven, & D. L. Stufflebeam (Eds.), Evaluation models: Viewpoints on educational and human services evaluation (pp. 229-260). Boston, MA: Kluwer-Nijhoff.
- Sergiovanni, T. J. (1987). The principalship. Boston, MA: Allyn and Bacon.
- Shulman, L. S. (1987). Assessment for teaching: An initiative for the profession. Phi Delta Kappan, 69, 38-44.
- Snyder, K. J., & Anderson, R. H. (1986). Managing productive schools. Orlando, FL: Academic Press.
- Soar, R. S., Medley, D. M., & Coker, H. (1983). Teacher evaluation: A critique of currently used methods. Phi Delta Kappan, 65, 239-246.
- Stark, J. S., & Lowther, M. A. (1984). Predictors of teachers' preferences concerning their evaluation. Educational Administration Quarterly, 20, 76-106.
- Stiggins, R. J. (1986). Teacher evaluation: Accountability and growth systems--different purposes. National Association of Secondary School Principals Bulletin, 70(490), 51-58.

- Stiggins, R. J., & Bridgeford, N. J. (1985). Performance assessment for teacher development. Educational Evaluation and Policy Analysis, 7, 85-97.
- Stiggins, R. J., & Duke, D. L. (1986). Research on teacher evaluation: The case for a commitment to teacher growth. Portland, OR: Northwest Regional Educational Laboratory.
- Stodolsky, S. S. (1984). Teacher evaluation: The limits of looking. Educational Researcher, 13(9), 11-18.
- Strike, K., & Bull, B. (1981). Fairness and the legal context of teacher evaluation. In J. Millman (Ed.), Handbook of teacher evaluation. Beverly Hills, CA: Sage Publications.
- Stronge, J. H. (1988). The elementary principalship: A position in transition? Principal, 67(5), 32-33.
- Sullivan, C. G. (1980). Clinical supervision: A state of the art review. Alexandria, VA: Association for Supervision and Curriculum Development.
- Sweeney, J. (1982). Research synthesis on effective school leadership. Educational Leadership, 39, 346-352.
- Sweeney, R. C., & Lindsey, R. B. (1987, April). Problem solving: A way to operationalize the supervision process. Paper presented at the Annual Meeting of the Association of California School Administrators, San Francisco, CA.
- TAP begins portfolio assessment study. (1988, Fall-Winter). The Teacher Assessment Project Newsletter, pp. 1-2.
- Tye, K. A., & Costa, A. L. (1986). Better teaching through instructional supervision: Policy and practice. Sacramento, CA: California School Boards Association. (EDRS Document Reproduction Service No. ED 277 108)
- Weber, G. (1971). Inner-city children can be taught to read: Four successful schools (Occasional Paper No. 18). Washington, DC: Council for Basic Education.
- Weber, J. R. (1987). Teacher evaluation as a strategy for improving instruction. Eugene, OR: ERIC Clearinghouse on Educational Management.

- Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1984). Teacher evaluation: A study of effective practices. Santa Monica, CA: The Rand Corporation.
- Wolf, R. L. (1973). How teachers feel toward evaluation. In E. R. House (Ed.), School evaluation: The politics and process (pp. 156-168). Berkeley, CA: McCutchan.
- Wood, C. J., & Pohland, P. A. (1979, April). Teacher evaluation: Organizational maintenance versus stimulation of improved teaching performance. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.
- Wuhs, S. K., & Manatt, R. P. (1983). The pace of mandated teacher evaluation picks up. American School Board Journal, 170(5), 28.

APPENDIX A

DISTRIBUTION OF TEACHERS INCLUDED IN THE SAMPLE

DISTRIBUTION OF TEACHERS INCLUDED
IN THE SAMPLE

Following is a distribution of the 402 permanent teachers included in the survey sample:

Elementary Classroom Teachers (K-6) N = 190

135 Classroom Teachers

2 Band Teachers

2 Orchestra Teachers

8 Vocal Teachers

13 Physical Education Teachers

3 Talented and Gifted Teachers

3 Chapter I Teachers

9 Learning Disability Teachers

1 Intervention Teacher

2 Special Education Itinerants

3 Elementary Learning Center Teachers

3 English As a Second Language Teachers

2 Teachers for the Severely Emotionally Disturbed

4 Speech Clinicians

Secondary Classroom Teachers (Grades 7-12) N = 156

144 Subject Matter Teachers

1 English As a Second Language Teacher

1 Chapter I Teacher

1 Teacher for the Teachable Mentally Retarded

9 Resource Room Teachers

Elementary/Secondary Classroom Teachers N = 5

- 1 Speech Clinician for the Teachable Mentally Retarded
- 2 Motor Development Teachers
- 2 Home Instruction Teachers

Elementary Nonclassroom Teachers N = 14

- 8 Elementary Media Specialists
- 1 Elementary Program Specialist
- 2 Elementary Child Development Specialists
- 1 Elementary Student Relations Specialist
- 1 Elementary Teacher on Special Assignment
- 1 Elementary Chapter I Specialist

Secondary Nonclassroom Teachers N = 28

- 14 Secondary Counselors
- 2 Secondary School Alcohol Facilitators
- 2 Secondary Work Experience Coordinators
- 1 Secondary Career Coordinator
- 7 Secondary Media Specialists
- 1 Secondary Student Supervisor
- 1 Secondary Staff Development Specialist

Elementary/Secondary Nonclassroom Teachers N = 9

- 1 Elementary Specialist
- 1 Language Arts Specialist
- 4 Psychologists
- 1 Resource Librarian
- 2 Teachers on Special Assignment

APPENDIX B

CORRESPONDENCE TO TEACHERS

March 2, 1988

Dear Colleague,

My name is Bev Hobson. I am a teacher in the Beaverton School District and a doctoral student at Portland State University. The topic of my dissertation deals with teachers' perceptions of the evaluation process and its impact on teacher growth.

Your name has been randomly selected for voluntary participation in this study. Since the sample size is small, it is especially important that you respond. I would appreciate your support by taking fifteen minutes to read the questionnaire and return the completed response form to me at Aloha Park School through the interschool mail by March 11.

The master list of code numbers and names is being held by Doug Smith, principal at Aloha Park, so that neither one of us will be able to put the returned responses with a specific name. ALL RESPONSES WILL BE KEPT STRICTLY CONFIDENTIAL, AND INFORMATION WILL BE REPORTED BY CATEGORY OF RESPONDENT RATHER THAN BY NAME OR PLACE. The purpose of the code number at the top of each response form is so that I can send an additional mailing to nonrespondents. The list of names and numbers will be destroyed upon return of the response forms and completion of the study.

While this study is not sponsored by the school district, both Beaverton School District and Beaverton Education Association are supportive of this study and results will be shared with those organizations.

Thank you very much for taking time from your busy day to answer these questions. In order to increase participation so that the results will be meaningful, I have purchased a government savings bond for the amount of \$100 which will be issued to a participant in this study. Upon completion of the study, this summer, one of the participants will be randomly selected by a third party and awarded the savings bond.

Please return the response form to Bev Hobson at Aloha Park School by March 11.

Sincerely,

Bev Hobson

March 14, 1988

Dear Colleague,

Just a reminder that I have not yet received the survey I sent you on March 2. Your input on teachers' perceptions of the evaluation process and its impact on teacher growth is important. ALL RESPONSES WILL BE KEPT STRICTLY CONFIDENTIAL, AND INFORMATION WILL BE REPORTED BY CATEGORY OF RESPONDENT RATHER THAN BY NAME OR PLACE.

If you have already mailed your survey, thank you. I realize what a busy time of the year this is. I would appreciate your support by taking fifteen minutes to read the questionnaire and return the completed response form to me at Aloha Park School through the interschool mail by March 18.

A \$100 government savings bond will be issued to one of the participants of this study. A name will be randomly selected by a third party upon completion of this study. Please return the response form to Bev Hobson at Aloha Park School by March 18.

Sincerely,

Bev Hobson