

Portland State University

PDXScholar

Dissertations and Theses

Dissertations and Theses

1989

The Characteristics of the Teacher Evaluation Process as Perceived by Elementary Teachers and Principals

Peter H. Taylor
Portland State University

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

Let us know how access to this document benefits you.

Recommended Citation

Taylor, Peter H., "The Characteristics of the Teacher Evaluation Process as Perceived by Elementary Teachers and Principals" (1989). *Dissertations and Theses*. Paper 1365.
<https://doi.org/10.15760/etd.1364>

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.

THE CHARACTERISTICS OF THE TEACHER EVALUATION PROCESS
AS PERCEIVED BY ELEMENTARY TEACHERS AND PRINCIPALS

by
PETER H. TAYLOR

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

DOCTOR OF EDUCATION
in
PUBLIC SCHOOL ADMINISTRATION AND SUPERVISION

Portland State University

1989

TO THE OFFICE OF GRADUATE STUDIES:

The members of the Committee approve the dissertation
of Peter H. Taylor presented June 5, 1989

[Redacted Signature]

John D. Lind, Chairperson

[Redacted Signature]

Loyde W. Hales

[Redacted Signature]

John P. Heflin

[Redacted Signature]

Meredith D. Gall

[Redacted Signature]

Sandra C. Anderson

APPROVED:

[Redacted Signature]

Robert B. Everhart, Dean [Redacted] School of Education

[Redacted Signature]

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

ACKNOWLEDGEMENTS

This dissertation is the culmination of a professional goal that was made possible by the support, assistance, and guidance of many people.

I wish to thank the Tigard School District and the Parkrose School District for their permission to conduct this research. Denny Westover, Art Rutkin, and Anne Wax have taken time from their busy schedules to give support for this project.

Rick Stiggins, Phil Nickel, Mary Rose, and Maggie Rogers from the Northwest Regional Educational Laboratory have given advice and technical support that made this project possible.

I would like to express appreciation to the members of my committee, Dr. Sandra Anderson and Dr. John Heflin for their support and guidance. I would like to especially acknowledge the assistance of Dr. John Lind, my advisor throughout my doctoral program, for his continual support. I also wish to acknowledge the many hours of guidance and assistance provided by Dr. Loyde Hales in organizing this project. I am grateful to Dr. Meredith Gall for his editorial skills and for his guidance in the logical presentation of this study.

Finally, I wish to thank my wife, Bonnie, for her patience and encouragement throughout this project and I dedicate this dissertation to her.

TABLE OF CONTENTS

	PAGE
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	vii
CHAPTER	
I INTRODUCTION	1
Background of the Study	1
Rationale for the Study	6
Purpose of the Study.....	8
Significance of the Study.....	8
Limitations of the Study	10
Definition of Terms	11
II REVIEW OF THE LITERATURE	12
Introduction	12
Foundations of Teacher Evaluation	12
Clinical Supervision	15
Perceptions of Teachers and Principals	22
Instructional Theory Into Practice (ITIP)	27
Leadership of the Principal	35
Summary of the Literature Review	39

CHAPTER		PAGE
III	METHODOLOGY	41
	Introduction	41
	Sample	41
	The Research Instruments	45
	Data Collection	51
	Data Analysis	52
	Research Questions	53
IV	RESULTS	55
	Introduction	55
	Teacher Respondent Information	56
	Total Teacher Responses	56
	Principal Respondent Information	70
	Total Principal Response	72
	Analysis of the Research Questions	85
	Summary of Chapter IV	115
V	SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS	119
	Introduction	119
	Summary of the Study	120
	Conclusions	122
	Recommendations	132
	Summary of Chapter V	134

REFERENCES.....	135
APPENDICES	143
A: TEACHER EVALUATION PROFILE.....	143
B: PRINCIPAL EVALUATION PROFILE	151
C: COMPARISON WITH NATIONAL TEP STUDY.....	161

LIST OF TABLES

TABLE		PAGE
I	Frequency/Percentage Distributions for Teaching Experience, Employment Status, and ITIP Training for Teachers	43
II	Frequency/Percentage Distributions for Administrators Experience, Evaluation Training, and ITIP Experience for Principals	44
III	Teacher Experience with ITIP	57
IV	Response of Teachers to the Items Describing the Attributes of the Teacher	59
V	Response of Teachers to the Items Describing the Attributes of the Evaluator	61
VI	Response of Teachers to the Items Describing the Attributes of the Procedures	63
VII	Response of Teachers to the Items Describing the Attributes of the Feedback	65
VIII	Response of the Teachers to the Items Describing the Attributes of the Evaluation Context	66
IX	Response of the Teachers to the Items Describing Their Experience with ITIP	67
X	Response of the Teachers to the Criterion Questions on Quality, Impact, and Value of Teacher Evaluation	68

TABLE	PAGE
XI Means, Standard Deviations and Reliability Coefficients for the Total Teacher Sample on the Sections of the TEP	69
XII Principals Experience with Evaluation and ITIP	71
XIII Response of the Principals to the Items on Section A of the TEP	73
XIV Response of the Principals to the Items in Section B of the TEP	75
XV Response of the Principals to the Items in Section C of the TEP	77
XVI Response of the Principals to the Items in Section D of the TEP	79
XVII Response of the Principals to the Items in Section E of the TEP	81
XVIII Response of the Principals to the Items in Section F of the TEP	82
XIX Response of the Principals to the Criterion Questions on Quality, Impact, and Value of Teacher Evaluation	83
XX Means, Standard Deviations, and Reliability Coefficients for the Principals Sample on the Sections of the TEP	84
XXI Correlation of Teacher Responses on the TEP Items and Criterion Ratings for Quality, Impact and Value for Section A	87

TABLE	PAGE
XXII Correlation of Teacher Responses on the TEP Items and Criterion Ratings for Quality, Impact, and Value for Section B	89
XXIII Correlation of Teacher Responses on the TEP Items and Criterion Ratings for Quality, Impact, and Value for Section C	91
XXIV Correlation of Teacher Responses on the TEP Items and Criterion Ratings for Quality, Impact, and Value for Section D	93
XXV Correlation of Teacher Responses on the TEP Items and Criterion Ratings for Quality, Impact, and Value for Section E	94
XXVI Correlation of Teacher Responses on the TEP Items and Criterion Ratings for Quality, Impact, and Value for Section F	95
XXVII Analysis of Variance Table for Teacher Rating of Criterion Questions and Number of ITIP Courses	97
XXVII Analysis of Variance Table for Scale Scores and ITIP Courses	98
XXIX Analysis of Variance Table for Scale A and Number of ITIP Courses Taken by Teachers	100

TABLE		PAGE
XXX	Significant Item Means and Stand Deviations on Scale A for Teachers Grouped by ITIP Courses	101
XXXI	Analysis of Variance Table for Scale F and Number of ITIP Courses Taken by Teachers	102
XXXII	Significant Item Means and Standard Deviations on Scale F for Teachers Grouped by ITIP Courses Taken	103
XXXIII	Analysis of Variance Table for Criterion Questions for Probationary and Tenured Teachers	104
XXXIV	Analysis of Variance Table for Scale Scores for Probationary and Tenured Teachers	105
XXXV	Analysis of Variance Table for Scale A for Probationary and Tenured Teachers	107
XXXVI	Significant Item Means and Standard Deviations on Scale A for Probationary and Tenured Teachers	108
XXVII	Analysis of Variance Table for Scale E for Probationary and Tenured Teachers	109
XXXVIII	Item Means and Standard Deviations on Scale E for Probationary and Tenured Teachers	110
XXIX	Analysis of Variance Table for Teachers and Principals Ratings of the Criterion Questions	111

TABLE	PAGE
XXXX Analysis of Variance Table for Scale Scores for Principal and Teachers	112
XXXXI Analysis of Variance Table for Scale E for Principals and Teachers	113
XXXXII Analysis of Variance Table for Scale F for Principals and Teachers	114

AN ABSTRACT OF THE DISSERTATION OF Peter H. Taylor for
the Doctor of Education in Public School Administration and
Supervision Presented June 5, 1989.

TITLE: The Characteristics of the Teacher Evaluation Process
 as Perceived by Elementary Teachers and Principals

APPROVED BY MEMBERS OF THE DISSERTATION COMMITTEE:

[Redacted Signature]

John D. Lind, Chairperson

[Redacted Signature]

Loyde W. Hales

[Redacted Signature]

John F. Heflin

[Redacted Signature]

Sandra C. Anderson

[Redacted Signature]

Meredith D. Gall

The purpose of this study was to investigate the perceptions of elementary teachers and principals of the evaluation system used in two suburban school districts near Portland, Oregon. The two districts involved in the study used the ITIP model for instruction and teacher evaluation over a period of several years. The Teacher Evaluation Profile (TEP) developed by

the Northwest Regional Education Laboratory was the instrument used in the study of teachers. A modified version of the TEP was used with elementary principals. The population of the study consisted of 233 elementary teachers from grades K-6 and 14 elementary principals. The study examined their perceptions of the teacher characteristics, evaluator characteristics, evaluation procedures, the feedback, the evaluation context, and experience with Madeline Hunter's Instructional Theory Into Practice.

In addition to examining the general perceptions of teachers and the principals, the study also investigated four research questions:

- (1) Is the value that teachers place on teacher evaluation related to the characteristics of the evaluation?
- (2) Does the amount of training that teachers receive in ITIP affect the value they place on teacher evaluation?
- (3) Do probationary and tenured teachers have different perceptions about the nature and value of teacher evaluation?
- (4) Do principals and teachers have different perceptions about the nature and value of teacher evaluation?

The results of the questionnaires were examined using ANOVA and correlational techniques. In addition, alpha coefficients were computed to estimate the internal reliability of the instruments.

The findings suggest that teachers and principals had positive perceptions of the quality, impact, and value of the evaluation process. Teachers perceived a positive working relationship with their principals in the evaluation process, and they rated the principals high on their evaluation skills. Teachers reported that feedback from principals was specific and that

feedback was generally helpful. Teachers and principals agreed that there was value in sharing a common terminology (from ITIP) for discussions in the evaluation conference.

CHAPTER I

INTRODUCTION

Chapter I describes the background of this study and also its significance and purpose. The following sections are included in this chapter: (1) Background of the Study; (2) Rationale for the Study; (3) Purpose of the Study; (4) Significance of the Study; and (5) Limitations of the Study.

BACKGROUND OF THE STUDY

Four developments in the past decade have directed new attention to the importance of teacher evaluation. First, various national commissions, such as the National Commission on Excellence in Education, issued reports calling for school reform. Second, research on effective schools succeeded in identifying educational practices that increase student achievement. Third, the model of instruction developed by Madeline Hunter was adopted by many school systems across the nation. The fourth development was the increasing centralization of power at the state level brought about by the development of policies designed to improve education.

Several of the reports of national commissions emphasized the importance of teacher evaluation. For example, one of the primary recommendations of the National Commission on Excellence in Education report, "A Nation at Risk, The Imperative for Educational Reform," (1983)

was to evaluate teacher performance more rigorously. This recommendation was taken seriously at the state level (Kirst, 1983).

A number of states established teacher evaluation as their cornerstone for educational reform. California and Texas enacted legislation requiring that more time be spent by school administrators in observing and evaluating teachers. Other states subsequently enacted similar legislation. The majority of teacher evaluation programs developed in response to this legislation followed a format of clinical supervision involving a preconference between the teacher and principal, the direct observation of teaching, and a postconference.

Many other proposals for educational reform proved to be dependent upon an effective and efficient system of teacher evaluation. Examples of such proposals are calls for the development of master teacher programs, merit pay systems, and career ladders. All of these proposals are dependent upon teacher evaluation as a means of identifying outstanding teachers.

A second development that directed attention to the importance of teacher evaluation was research on the characteristics of effective schools. This research grew out of concern about a persistent problem of many American schools. Student achievement in the basic skills and other subjects is particularly low in those public schools that serve low-income and minority students. For decades, educators, researchers, and the public explained this low performance by attributing the cause either to characteristics of the student or to the family background of the student. These explanations ignored the influence of the teachers or principals who worked with these students. However, research in the early 1980s demonstrated that some schools were effective in teaching students from

low income and minority backgrounds. This discovery directed attention to investigating characteristics of these effective schools and supported the argument that schools can make a difference in the level of student achievement.

Edmonds (1981), in his pioneering work on effective schools, described five fundamental characteristics:

1. The leadership of the principal characterized by attention to the quality of instruction
2. A pervasive and broadly understood instructional focus
3. An orderly, safe climate conducive to teaching and learning
4. Teacher behaviors that convey the expectation that all students are to obtain at least minimum mastery
5. The use of measures of pupil achievement as the basis for program evaluation (p. 4)

These findings emphasized the importance of both the principal and the teacher. The principal of an effective school was described as an instructional leader who regularly visited classrooms and who spent much of his or her time dealing with the evaluation of teachers. In describing effective schools, Brookover and Lezotte (1982) emphasized the need for teacher growth through supervision and evaluation. The research on effective schools directed attention to the importance of teacher evaluation.

The third factor that has concentrated attention on teacher evaluation has been the broad acceptance of Madeline Hunter's instructional model. Few programs in the history of American education have gained such wide

acceptance as quickly as the Instructional Theory Into Practice (ITIP) program developed by Hunter. McGreal (1987), for example, observed that many southern states have mandated that every teacher be trained in the ITIP model. Hunter, who is a professor of education at the University of California in Los Angeles, is one of the most popular educational trainers in the country. She has given hundreds of training sessions in school districts across the country and has produced training films and books on ITIP. In addition, there is a national network of teacher trainers who support the model.

The ITIP model includes several instructional methods that teachers use in planning and conducting lessons. These methods are based upon commonly accepted psychological theories, such as motivation, reinforcement, and transfer. Some of the methods described in the model are based upon research in learning theory dating back to Thorndike. However, the results of many current research studies also support ITIP methods. For example, research on teacher effects and student achievement (Good and Brophy, 1984; Rosenshine, 1985) support the methods advocated by Hunter. Studies on classroom management (Evertson, 1979) also support many of the ITIP methods.

The adoption of the ITIP instructional model by thousands of school districts across the country has influenced teacher evaluation practices. One of the effects of its adoption has been to narrow the focus on teaching to those methods described in the model. In turn, teacher evaluation has been modified to ensure that the ITIP model for instruction is implemented. Some school districts have modified their clinical supervision model to match the Hunter model. In fact, Hunter (1984) has designed a teacher evaluation

process to support her instructional model. Her teacher evaluation process is a modified form of clinical supervision involving classroom observation and conferencing. Hunter has designed a program to train principals in this teacher evaluation program.

The fourth factor has been the increasing development of policy at the state level to correct local educational problems. This trend to create policies at the state level that are designed to standardize the operation of educational institutions has been described by Wise (1979) as hyperrationalization. He identifies three trends in policy development: federal and state governments are now making policy in areas formerly reserved to local school boards and college boards of trustees; general government is making policy in areas formerly reserved to educational government; and, as other levels of government make educational policy, schools are becoming more bureaucratic. This trend, which Wise identified in the late 1970's continued into the 1980's and, as a result, most of the programs identified with the first wave of educational reform have come from state mandates.

In summary, the following four factors have converged to make teacher evaluation a matter of much attention and importance: (1) national commissions on school reform, (2) research on effective schools, (3) the broad adoption of Hunter's instructional model and (4) the increasing influence of state education policies.

RATIONALE FOR THE STUDY

The preceding discussion makes clear that teacher evaluation is now, and will continue to be, a highly visible process. Therefore, it's important to learn more about teacher evaluation. One area in need of research is an examination of the attitudes of educators toward teacher evaluation. It seems likely that their attitudes would affect the success of teacher evaluation and, in turn, the success of school reforms that depend on teacher evaluation.

Descriptions in the literature of teacher and principal attitudes toward evaluation are discouraging. In reviewing the literature on teacher evaluation, McLaughlin and Pflieger (1986) described the evaluation process as pro-forma, meaningless, and ineffective. Popham (1988) described the process as a high-cost, low-yield endeavor. Dissatisfaction with the process is indicated by the titles of some of the early books on teacher evaluation: A Reluctant Profession by Mosher and Purple (1972), and A Private Cold War by Blumberg (1974). It is important, therefore, to learn more about the attitudes of teachers and principals toward evaluation and to discover ways to improve the process.

McLaughlin (1984) found that teachers placed little value on the evaluation process because it provided feedback which was too general to be useful. Blumberg and Amidon (1965) reported that teacher's attitudes toward evaluation were affected by their perceptions of the principal's manner and skills in communication. These findings suggest that it would be productive to examine the relationship between the value that teachers place on evaluation and specific characteristics of the evaluation process.

There is also a need to examine differences among teacher groups in their attitudes toward evaluation. Do beginning teachers who are still in their probationary period (the first three years of teaching) view the evaluation process differently than the more experienced teachers who have received tenure? Information is lacking in the literature on this aspect of teacher attitudes toward evaluation.

It is clear that the attitude of the principal toward evaluation is critical to the success of the program. McLaughlin (1984) stresses the importance of the principal's commitment to teacher evaluation as central to a successful program. This question must be examined because the extent to which principals and teachers differ in their perceptions about the nature and value of teacher evaluation will determine the success of an evaluation program.

McGreal (1983) suggested that principals and teachers who are trained in the same instructional model have a common perspective on teaching. He also suggested that this common perspective on instructional practice would cause teachers and principals to have similar attitudes toward teacher evaluation. Darling-Hammond (1983), in a Rand Corporation study of school districts with exemplary teacher evaluation programs, also supported this position. One of the assumptions of those who advocate the use of the ITIP model is that teachers who have received training in the model and are evaluated on the basis of the model place a greater value on teacher evaluation. Hunter (1988), who has stressed the importance of training programs for both teachers and principals, supports this position. However, there is little evidence available to support the assumption.

The purpose of the present study is to test these assumptions by investigating the perceptions of teachers and administrators regarding teacher evaluation in two school districts that have made extensive use of the Hunter model of instruction.

PURPOSE OF THE STUDY

The purpose of the study is to examine the general perceptions of teachers and principals toward teacher evaluation. The study was conducted in the elementary schools of two suburban school districts near Portland, Oregon, that have adopted the ITIP model and have trained both teachers and administrators in the model. In addition to examining the general perceptions of teachers and principals, four research questions were developed to examine specific aspects of the evaluation process.

1. Is the value that teachers place on teacher evaluation related to the characteristics of the evaluation?
2. Do probationary and tenured teachers have different perceptions about the nature and value of teacher evaluation?
3. Do principals and teachers have different perceptions about the nature and value of teacher evaluation?
4. Does the amount of training that teachers receive in ITIP affect the value they place on teacher evaluation?

SIGNIFICANCE OF THE STUDY

A study of teacher and administrator perceptions of evaluation in districts that have trained both principals and teachers in the ITIP model can provide information valuable in improving the evaluation process. This

information can be used to design in-service programs for teachers and principals which focus on the specific needs of both groups.

Suppose, for example, the findings relating to the first research question reveal that teachers place high value on evaluations which have certain specific characteristics. These characteristics can be emphasized in training principals to conduct evaluations. If, however, we find that the value that teachers place on evaluation is independent of the characteristics of the evaluation, then it would be necessary to look for other factors influencing teacher attitudes toward evaluation.

The second research question explores whether probationary and tenured teachers have different perceptions about the nature and value of teacher evaluation. This question is prompted by educators such as Stiggins (1988) who argues that programs of formative and summative evaluation should not be combined in one evaluation format and that more experienced teachers should be assessed by peers. If, for example, it is discovered that teachers with more experience find the evaluation performed by the building principal less valuable than do probationary teachers, it may be wise to examine the current system. On the other hand, if tenured and probationary teachers do not differ in their perceptions of the value, it may be wise not to change the current system.

The third research question asks whether principals and teachers have different perceptions about the nature and value of teacher evaluation. Barber (1987) found that teachers view evaluation as aversive control over their professional careers. This finding would predict that teachers view teacher evaluation differently than principals. However, it may be that the existence of a common standard for instructional practice, as in the case of

ITIP, causes teachers and principals to be more congruent in their views of the nature and value of teacher evaluation.

The fourth research asks whether the amount of training that teachers have in ITIP affects the value they place on teacher evaluation. This question is significant because many school districts have expended large amounts money over the years to train both teachers and administrators in ITIP.

The answers to the research questions posed in this study can serve as a base of information which will be valuable not only to the districts involved in this study, but to other school districts as well.

LIMITATIONS OF THE STUDY

1. Generalizations from this study are limited because the study was conducted in elementary schools.
2. The investigation was conducted during one academic year in only two suburban school districts.
3. Data used in the study were collected through the use of a questionnaire with subjects reporting their perceptions of the evaluation process. Self reporting limits the validity of the data.
4. The reliability and validity of the Principal Evaluation Profile (PEP) was tested only in this study.

DEFINITION OF TERMS

Evaluation

Evaluation is a program designed for the appraisal of a teacher's performance. Standards for administering the program are detailed in board of education policy statements.

ITIP

ITIP stands for Instructional Theory Into Practice. This program is a series of teaching decisions based on psychological theory. The program was developed by Madeline Hunter.

Clinical Supervision

Clinical Supervision is a system for collecting appraisal data based on the classroom performance of teachers. The system involves the collection and analysis of written classroom observations and usually involves a preconference, observation, and postconference.

CHAPTER II

REVIEW OF THE LITERATURE

INTRODUCTION

This survey of the literature provides a framework for examining the process of teacher evaluation and for reviewing the reports and research studies that are relevant to the research questions investigated in this study. The following sections are included in this chapter: (1) Foundations of Teacher Evaluation; (2) Clinical Supervision; (3) Perceptions of Teachers and Principals; (4) ITIP; (5) Leadership of the Principal; and (6) Summary of the Literature Review.

FOUNDATIONS OF TEACHER EVALUATION

Contemporary strategies of teacher evaluation have evolved from management practices that date back to the turn of the century. Between 1890 and 1930, school management was greatly influenced by practices in business and industry. Tyack and Hansot (1982) reported that during this period businessmen had a major influence on the operation of schools. They were active in the political movement to abolish ward school boards and to modernize urban school systems.

The pattern of organization that was predominant in business during this time stressed structure and formal relationships. Business managers during this time had been trained in the theories of scientific management developed by Frederick Taylor. Taylor (1912) identified task analysis and task management as the critical elements of an effective organization. Business leaders had also been influenced by the work of Weber (1947) whose model of bureaucratic organization was based on the assumption that the way in which a staff is structured provides the rational means to achieve organizational objectives. A third management theory which influenced business and, in turn, school management came from the work of Fayol (1916). Fayol developed a system of management control which operated to ensure that all elements of an organization including things, people, and actions are in conformity with a master plan. By the 1930s, the concepts of efficiency, task structure, bureaucratic organization, and management control were well established in America's large school systems. These structures were copied by smaller school districts and became part of the curriculum for training school administrators.

Tyack (1975) has documented the influence of scientific management upon the professional training of school administrators. During this period, the major emphasis in training school administrators was upon scientific management for school effectiveness. School administrators trained at institutions such as Stanford and Columbia were described as "school executives." Callihan (1962), in his work on the history of school management, referred to this trend as the "cult of efficiency."

Teacher evaluation practices during this time reflected this concern for effectiveness and efficiency. Many of these evaluation systems consisted of a description of teacher traits. Teacher performance was graded on the basis of these traits, and the teacher was assigned a numeric grade. Davis (1964) described a system used during that period which included such criteria as physical efficiency, social efficiency, dynamic efficiency, and achieved efficiency. These teacher trait systems were used extensively during this era and adaptations of these evaluation systems are still used in some school districts today.

Theories of effective management changed dramatically in the period between 1930 and 1960 and, once again, practices in business and industry influenced school management. In the 1930s, Mayo (1945) demonstrated that workers are motivated by social needs and that meeting these needs can increase worker productivity. For the next three decades worker motivation became an important management priority. Lewin (1958) studied the effects of organizing workers into groups to share ideas about production. McGregor (1960) compared two management assumptions about worker motivation. Maslow (1962) identified levels of need that influence worker behavior. Armed with these new tools of modern management, school principals sought new ways to involve teachers in the evaluation process. Texts on school administration reflected these practices. Two of the most popular school supervision texts of that period, Supervision, by Baar, Burton and Brueckner (1955) and Supervision For Better Schools, by Wiles (1955), emphasized the importance of democratic management by the school principal. Teacher evaluation during this period also changed dramatically.

During this period teacher evaluation became a more democratic process. Checklists of teacher behavior were still used, but principals began to meet with teachers to discuss what aspects of their teaching should be reviewed, and observation of teaching in classrooms became more common

Educators also directed their attention to the process of classroom observation and began to develop systematic ways of observing teacher behaviors. Flanders (1967) developed a complex system of charts and numbers to describe teaching behaviors in the classroom:

It was . . . during July of 1957 that I first studied a ten-by-ten interaction analysis matrix. It was tabulated from some code numbers collected a few days earlier in an elementary classroom.. The notion that the numbers in certain rows and cells could explain a teacher's influence pattern became apparent only gradually. These insights did not burst into full bloom suddenly; they sort of crept into the matrix one at a time. (p. vii)

The influence of democratic management, the increase in the direct observation of teaching behavior, and the development of strategies for collecting data on teaching behaviors were three factors which changed teacher evaluation practices during this period and set the stage for the development of a more effective model for teacher evaluation, clinical supervision.

CLINICAL SUPERVISION

Clinical supervision is described by educational researchers as one of the most effective teacher evaluation models available; (McLaughlin and Pfeifer, 1986; McGreal, 1986; Darling -Hammond, 1984). Clinical supervision

grew out of the work of Morris Cogan and his associates at Harvard University in the 1950s. Cogan, who was Director of Secondary School Apprentice Teaching, applied the term clinical supervision to the form of supervision which is designed to provide feedback on in-class behaviors of teachers. Data are collected, often by verbatim notes taken by a supervisor, and a conference is held with the teacher to interpret these data.

Cogan's work in clinical supervision began during the 1955-56 academic year with students enrolled in a teacher education program. Cogan (1976) reported that his reasons for developing the new system were the inadequacies in the preservice education of teachers, the underdeveloped state of supervisory practice, the need to provide a support system for teachers, and the need to establish an analytical approach to the problems of teaching.

Cogan (1973) described the system as a means of collecting empirical data on classroom teaching:

The word clinical was selected precisely to draw attention to the emphasis placed on classroom observation, analysis of in-class events, and the focus on teachers' and students' in-class behavior. In brief, clinical was designed both to denote and connote the salient operational and empirical aspects of supervision in the classroom.(p.9)

Cogan's work, which was originally designed to supervise student teachers, became a major tool for teacher evaluation. Since the 1960s, clinical supervision has become the standard model for teacher evaluation.

Major texts on the subject were written by Goldhammer(1969), Purpel (1972), Cogan (1976), Anderson and Krajewski (1980), and Acheson and Gall (1987). The authors outline a series of stages in clinical supervision that

follow a sequence. The following sections describe the sequence.

Stage 1 Preobservation Conference

The primary purpose of this stage is to establish rapport, to help the teacher understand the supervisory process, to agree upon the purpose of the observation, to help the supervisor understand the objective of the lesson, to establish parts of the lesson the teacher wants to have observed for specific feedback, and to set the time for the observation.

Stage 2 Observation and Data Collection

During this stage the supervisor takes copious notes and attempts to record data which will become the basis for the conference to follow. There are many strategies for data collection including verbatim notes, instruments for data collection, and audio or video recording.

Stage 3 Data Analysis and Strategy Planning

In this stage, the supervisor converts the data collected into information which is meaningful and manageable. The information should describe teaching patterns and critical incidents which can be used in the conference. After converting the raw data into information, the supervisor designs a strategy for conferencing with the teacher. Some of the factors to be considered in designing a strategy include the competency or experience level of the teacher, the nature of the original agreement with the teacher, the quality of the interpersonal relationship, and the emerging pattern of instruction described by the information collected.

Stage 4 Postobservation Conference

The primary purpose of this stage is to provide information to the teacher regarding the original agreement. This information should be descriptive rather than evaluative, and the teacher and principal should share the responsibility for interpreting the information. Skill in conducting conferences is the most critical attribute of the process, and much has been written describing this process. The post-observation conference ends with the development of an agenda for future action.

Stage 5 Postconference Analysis

This stage serves as an opportunity for the supervisor and the teacher to review and assess the process just experienced. Was the conference useful? Was the information accurate? What can the supervisor do to improve his or her skills? What should be pursued next as new cycles are undertaken?

The five-stage model described incorporates all the phases that are recommended by Cogan (1973) and the three major stages of the planning conference, observation, and feedback conference described by Acheson and Gall (1987). Clinical supervision, as a teacher evaluation model, supports all of the criteria which most authorities describe as critical to a successful evaluation program. Darling-Hammond (1983), in an extensive review of the literature, concluded that four minimal conditions are necessary for the successful operation of a teacher evaluation system:

1. All actors in the system have a shared understanding of the criteria and processes for teacher evaluation.

2. All actors understand how these criteria and processes relate to the dominant symbols of the organization; that is, there is a shared sense that they capture the most important aspects of teaching, that the evaluation system is consonant with educational goals and conceptions of teaching work.
3. Teachers perceive that the evaluation procedure enables and motivates them to improve their performance; and principals perceive that the procedure enables them to provide instructional leadership.
4. All actors in the system perceive that the evaluation procedure allows them to strike a balance between control and autonomy.
(p. 287)

Stiggins and Duke (1986), in their case studies of effective evaluation, identified five attributes which are critical to a successful evaluation. They listed the attributes of the teacher, the evaluator, the procedures used to gather data on teacher performance, the feedback given to the teacher, and the evaluation context. They identified characteristics of effective evaluation under each of these categories which included:

The Teacher

- Instructional Competence
- Personal Expectations
- Openness to Suggestion
- Orientation to Change
- Subject Knowledge
- Experience

The Evaluator

- Credibility
- Persuasiveness
- Patience
- Trust
- Track Record
- Modeling

Procedures

- Performance and Criteria Standards
- Data Collection Procedures

Feedback**Specificity of Feedback****Frequency of Feedback****Descriptive or Judgmental Feedback****Timing of Feedback****Feedback Related to Performance Standards****Context****Time Spent on Evaluation****Resources Available for Evaluation (pp. 79-92)**

Stiggins and Duke (1988) developed a teacher questionnaire based on these characteristics of successful evaluation. The Teacher Evaluation Profile (TEP) asks teachers to describe their most recent evaluation experience in the context of these characteristics. They call for the development of more case studies to describe the characteristics of successful teacher evaluation programs:

More case studies of successful evaluations are needed. Studies of districts conducting effective evaluation systems and studies of teachers whose professional development has been enhanced through effective evaluation.

McLaughlin and Pfeifer (1986) identified important elements of effective evaluation in four case studies of exemplary teacher evaluation programs which had been identified in an earlier study by Wise (1984). They identified four characteristics which are similar to the attributes identified by Duke and Stiggins.

1. Mutual trust between teachers and administrators

2 Open channels of communication

3. A general commitment to the individual and institutional learning

4. A high level of awareness of the evaluation activities (p. 3)

Mc Greal (1986), in surveying approximately 300 teacher evaluation systems with which he had worked as a consultant, identified nine characteristics which represent a set of best practices:

Attitudes. Sound systems are built around the positive attitudes and procedures needed to promote instructional improvement rather than the negative attitudes and procedures associated with evaluation for accountability and teacher dismissal.

Complementary procedures, processes, and instrumentation. Evaluation procedures are flexible, allowing supervisors and teachers to tailor data collection to the individual needs of teachers.

Separation of teacher evaluation from teaching evaluation. Evaluation systems are most effective when they deal with specific dimensions of classroom instruction and teaching behaviors that can be improved; they are least effective when they focus on performance criteria that are more administrative or personal in nature.

Goal setting. Effective, growth-producing evaluation systems replace the standardized criteria of accountability evaluation with the individualized professional development goals identified by teachers and supervisors.

Narrowed focus on teaching. Sound evaluation systems are centered on an agreed-upon and clearly-articulated definition of teaching that gives the teacher and the supervisor a common frame of reference.

Use of modified clinical supervision format. Helpful evaluation relies on preobservation planning, observation of specified behaviors, and a feedback conference.

Use of alternative data sources. Over and above classroom observation by the supervisor, evaluation can be based upon self-evaluation, peer evaluation, parent evaluation, student evaluation, student performance, and examination of classroom artifacts.

Different requirements for tenured and nontenured teachers. The purposes for evaluation differ in fundamental ways, so should evaluation procedures.

A complete training program. All participating teachers and supervisors must possess the skills and understanding needed to carry out an evaluation that will serve the growth purpose, which requires training. (pp. 9-10)

Teacher evaluation has evolved from a system of checklists which described teacher traits to a system that involves the collection of information from classroom observations and a joint analysis of these data by the teacher and the principal. The teacher's perception of the quality and value of the evaluation will be determined by such characteristics as the competence of the evaluator, the procedures used in evaluation, the feedback from the conference, and the context of the evaluation.

PERCEPTIONS OF TEACHERS AND PRINCIPALS

There are relatively few studies that compare the attitudes of probationary and tenured teachers toward evaluation. However, studies of preservice teachers seem to indicate that this group has a more positive attitude toward evaluation than more experienced teachers.

Shinn (1976) asked preservice teachers to rate the ideal frequency with which they would like school principals to use the phases of clinical

supervision and compared this with actual practice. He found that all of the teachers rated all of the phases as useful and wanted them to be used more often.

Research by Zonca (1972) involving intern teachers found that conditions of openness contributed to the interns' overall development and had a positive effect on attitude, ability to analyze classroom teaching behavior, and ability to move toward self-supervision.

A study by McCarty (1986) examined the perceptions of teachers who were evaluated in thirty-six school districts. Over seventy-five percent of these teachers reported that evaluation observations were too limited in number and that the feedback from the principal was too general to be of value in improving instruction.

These studies tend to support the notion that probationary teachers will view the evaluation process more positively than more experienced teachers. Experienced teachers may find the evaluation process less useful than teachers new to the profession because they have a greater mastery of the teaching methods and the subject matter. If this is true, it would support the recommendations of Glickman (1981) who advocates different types of evaluation for teachers who are at different stages in their professional development. It would also support the format developed by Hunter (1980) which uses different types of supervisory conferences for teachers with different needs.

Several researchers have found that principals have a more positive attitude toward teacher evaluation than do teachers. However, this does not mean that teachers necessarily have a negative attitude toward the process.

A statewide study by Goldsberry (1984) examined the attitudes of teachers and principals toward evaluation in Pennsylvania and reached the following conclusions:

1. There is no "cold war" in educational evaluation.
2. Teachers, for the most part, experienced very little direct supervisory assistance for improving their teaching.
3. Women are less likely to hold some jobs in public schools than are men.
4. The younger the learners who are served by the teacher, the more likely the teacher views supervision positively.
5. Principals perceive supervisory services as greatly more positive than do teachers.
6. For most teachers, supervision is a friendly but inconsequential ritual. (pp. 2-5)

Ruck (1984) investigated the implementation of a legislated teacher evaluation plan in Oregon. The plan, which followed the basic pattern of clinical supervision, was perceived by teachers and principals to be useful; but principals rated the plan more positively than did teachers.

A study by Graybeal (1984) in Oregon examined contemporary classroom supervision and found that the majority of teachers and principals perceived the supervisory relationship to be either helpful, collegial, or professional. He also found that the major purpose for classroom evaluation was perceived to be the improvement of instruction. A majority of the teachers and principals in this study indicated that the classroom supervision had contributed to the teachers' self-analysis of instruction.

McIntyre and Morris (1982) reported that while principals view the purpose of evaluation as a means of improving instruction, teachers see evaluation as a means for determining contract renewal. Blumberg (1980) reported that principals believe that evaluation has a high value but that teachers found the process to be of little value.

Barber (1987) describes aversive control as the main issue:

The most significant reason for the failure of teacher evaluation programs involves the perceived use of aversive control. Although designers of evaluation plans insist they do not use aversive control in the evaluation systems, most teachers and administrators think that any form of control (reward and/or punishment, summative systems) over their professional lives is aversive. Teachers' perceptions of aversive control found in evaluation plans are perhaps one of the most basic problems discussed. Teachers and principals clearly perceive any attempt to control their behavior through behavior modification or other means was extremely aversive and reacted as anyone does to aversive control-- they create a ministrike by becoming passive participants in the organization. (p. 15)

Stiggins (1988) argues that evaluation systems fail because traditional evaluation systems are designed to serve two conflicting functions. The first is to provide information for use in personnel management decisions such as hiring, firing, promotion, and tenure. This function has been defined as summative teacher evaluation (Millman, 1981) and serves the purpose of accountability. The second function, often called formative evaluation because it seeks to form or modify a teacher's instructional behaviors, serves the purpose of professional development and growth. This function is designed to help the already competent teacher continue to attain new levels of professional excellence. Stiggins advocates separating the two systems of teacher evaluation.

Popham (1988) agrees with Stiggins and suggests that evaluation tasks be carried out separately by different individuals:

In most school districts, formative and summative teacher evaluation are closely linked. Often these two functions are carried out by the same individual, typically a school's principal. Throughout the school year, the principal is usually obliged to provide teachers with suggestions for improvement as well as to make end-of-year judgements about the very same teachers.

In spite of its prevalence, the blending of formative and summative teacher evaluation represents a grave conceptual error. Both formative and summative evaluation are important functions, but these two teacher evaluation tasks must be carried out separately by different individuals. (p. 58)

An obvious question that comes from this discussion is : Should equal emphasis be placed on both formative and summative evaluation? A study by Bridges (1984) suggests that summative evaluation for purposes of dismissing tenured teachers for incompetence is extremely rare. His search of national publications and court records found only eighty-six cases from the period 1939 to 1982 where summative evaluations were used for the purpose of dismissing a teacher.

Given the dual role of the principal in evaluation, it is predictable that teachers and principals will have different perceptions about the nature and value of teacher evaluation. One of the critical problems associated with any situation in which one individual must judge another lies in the fact that the evaluation is not just a judgement about the teacher's professional competence, but is often perceived as a judgement about the worth of the individual.

INSTRUCTIONAL THEORY INTO PRACTICE (ITIP)

McGreal (1983) suggests that districts whose evaluation systems have been viewed as most effective have often adopted a narrowed focus on teaching. He recommends that the teaching focus should have at least four criteria and suggests that ITIP meets this standard:

1. A strong empirical base
2. A close approximation to standard practice
3. A commonsense orientation
4. Perspectives and skills that are potentially generalizable across subject areas and grade levels

In terms of current teaching research, the focus on teaching that seems to best meet the above criteria, as well as the one that has been the most successfully implemented in school districts, is based on a combination of current teacher effectiveness research and a portion of Madeline Hunter's work. (p. 304)

McGreal suggest that an instructional system should provide a particular perspective on teaching, have a set of definitions and language to describe instruction and should be presented to all teachers and principals at the same time and in the same manner.

Darling-Hammond (1984) examined the actual operation of teacher evaluation programs in four school districts that were recognized for having outstanding teacher evaluation. In describing the Lake Washington School District program, Darling-Hammond concludes:

The ITIP precepts that guide staff development for principals and teachers bring cohesiveness to an activity that is usually fragmented and erratic and helps teachers and their supervisors to identify and clarify problems. The ITIP framework also gives them tools and a common vantage point for developing pragmatic solutions. The investment in staff development thus increases the utility of teacher evaluation. (p. 31)

Hunter (1984) describes her model for instruction as a series of questions to be examined in planning lessons. She also warns against using the model as a rigid, lock-step plan for teaching. Hunter advises teachers to use the model as a guide for planning lessons. Hunter describes seven key elements in instruction and a rationale for each:

1. Anticipatory set. Has the teacher developed in the students a mental set that causes them to focus on what will be learned? An anticipatory set may also give some practice in helping students achieve the learning and yield diagnostic data for the teacher.
2. Objective and purpose. Not only do students learn more effectively when they know what they're supposed to be learning and why that learning is important to them, but teachers teach more effectively when they have that same information. Consequently, in words that are meaningful to the students, the teacher states what will be learned and how it will be useful.

3. Input. Students must acquire new information about the knowledge, process, or skill they are to achieve. Regardless of whether that information comes from discovery, discussion, reading, listening, observing, or being told, the teacher must have task-analyzed the final objective to identify knowledge and skills that need to be acquired. Only then can the input phase of the lesson be designed so that a successful outcome becomes predictable.

4. Modeling. Seeing what is meant is an important adjunct to learning. Usually, it is facilitating for the learners to directly perceive the process or product they are expected to acquire or produce. So that creativity will not be stifled or generalizability impeded, several examples should be a routine part of most (not all) lessons. Demonstrations, live or filmed, of process and products are facilitating rather than restricting to student initiative and creativity.

5. Checking for understanding. Before students are expected to do something, it is wise to ascertain that they understand what it is they're supposed to do and that they have the minimum skills required to do so. Sometimes this checking occurs verbally before actual student action. Sometimes it occurs simultaneously with the next element.

6. Guided practice. Students practice their new knowledge or skill under direct teacher supervision. New learning is like wet cement; it is easily damaged. An error at the beginning of learning can easily "set" so that it is harder to eradicate than if it had been apprehended immediately.

7. Independent practice. Independent practice is assigned only after the teacher is reasonably sure that students will not make serious errors. After an initial lesson, students frequently are not ready to practice independently, and the teacher has committed a pedagogical error if unsupervised practice is expected. (pp. 175-176)

Many of the instructional practices advocated by Hunter have been supported by research studies such as those of Good and Grouws (1979), Emmer and Evertson (1979), Rosenshine (1981), and Brophy (1975). Pratton

and Hales (1986) reported that active participation, which is advocated by Hunter, increased student learning.

However, some of Hunter's critics suggest that while individual elements of her model are supported by research findings, there is no evidence that the model, as a whole, is effective. Slavin (1988) claims that the model does not work:

But in actual fact many of the programs derived directly from the process-product research have failed. One clear example is the Madeline Hunter programs. Everything Madeline Hunter talks about is very well established, either in laboratory research or correlational research or both; but when assembled into an experimental program, will it work in actual practice? We now have enough evidence to say that it doesn't. (p. 27)

Slavin cites a study by Stallings and Krasvage (1986) as an example of a study that found no difference between control classes and those using Hunter's methods. The Napa Project (Stallings and Krasvage, 1986), was a four-year study which examined the effects of ITIP for instruction in reading and math. Robbins and Wolfe (1987) reported that during the fourth year of the study student scores declined after three years of gaining. The authors speculated that teaching practices had not become sufficiently mastered by the project teachers to maintain the growth that students had gained in the first three years of the program. Hunter (1988) agreed with this observation:

The Napa Project, given its flaws, presented hard data that test scores escalated when teachers were applying what they had learned. What the project really validated was that newly learned professional skills do not maintain themselves without encouragement and refinement from coaching. In addition, the teaching "skills" were used only for

reading and math; therefore, the teachers did not perceive them as generic elements in every teaching decision regardless of content or teaching mode. I criticized the project for the omission of the transfer theory that is central to the model. (p. 29)

Much of the criticism of Hunter's work comes from the university community. Gibboney (1987) criticizes ITIP on the grounds that the content of the model is primarily about techniques which are unrelated to a coherent educational theory. He claims that teachers are viewed as technical decision makers, not as professional decision makers.

Retallick (1986) faults the model on the grounds that it is behaviorist learning theory which addresses the question: Does observable behavior validate that learning has occurred? He contends that the supervision model is a very conservative process since it does not seek to generate new forms of information which could enable teachers to question the model or strive toward alternative proposals. Costa (1984) maintains that the Hunter format is expressed in the language of technology and is a narrow model which addresses only direct teaching.

Slavin (1987) recognizes the broad appeal of Hunter's work to teachers and administrators but warns against the institutionalization of the model:

Madeline Hunter's ideas are...common-sense translations of well-founded instructional theory put into practical form. At their best, her ideas give teachers new ways to improve their effectiveness...But the Madeline Hunter Movement has a serious downside. Scores of state and local districts have virtually institutionalized Hunter's model in every lesson. In some places the situation has become extreme. One physical education teacher in Texas received an unsatisfactory evaluation because her explanation of volleyball rules didn't include all steps of the Madeline Hunter lesson. (p. 56)

Over sixty percent of the school districts in Pennsylvania are involved in the Hunter model. In a study funded by the University of Pittsburgh, Garman and Hazi (1988) interviewed more than 200 Pennsylvania teachers who were involved in Hunter programs and found a mixed response to the model. Most of the criticism, however, centered on the implementation of the model by school administrators rather than on the model itself. Teachers in the study reported that the benefits included:

1. More attention to teaching
2. More positive reinforcement of teachers by administrators
3. More communication with their colleagues about teaching
4. A common language to describe instruction (p. 670)

Roughly two-thirds of the teachers in the study expressed concerns about their experiences with the model. Most of these concerns related to the way in which the model was implemented in the school district and included the following:

1. Being forced to adopt a single model of teaching
2. Feelings of rejection because the administration diminished what the teachers had done in the past and pushed for only one model
3. The negative reaction to mandates and the standardization of teaching practice. (p. 671)

Hunter (1980) has created a teacher evaluation program to support her model of instruction. The program is a modified version of clinical supervision involving direct observation of teaching and a conferencing

system. Hunter designed her supervisory conferences to meet the needs of teachers who are at different stages of professional development:

Type A conferences are for the purpose of identifying, labeling and explaining the teacher's effective instructional behaviors.

Type B conferences serve to stimulate the development of a repertoire of effective teaching techniques.

Type C conferences encourage the teacher to identify those aspects of the lesson with which the teacher was not satisfied in order to develop strategies for reducing or eliminating future unsatisfactory outcomes.

Type D conferences occur when the supervisor identifies and labels less effective aspects of the lesson that were not evident to the teacher.

Type E conferences are for the purpose of promoting continued growth of excellent teachers. (p. 410)

Hunter also has designed an evaluation conference that is summative in nature and takes into account the other types of conferences which have been held earlier with the teacher.

In summary, there is a lack of substantive research on the effects of ITIP. Stallings is a capable and respected researcher, but staff changes during the course of the Napa Project raise questions about the results of the study. Robbins (1986), in a presentation to the American Educational Research Association, identified problems in the study and suggested that following steps be taken in replicating the study:

1. Take the time to build a readiness for change.
2. Involve participants in all stages of the planning

3. Secure strong administrative support
4. Describe the training to the participants in advance
5. Monitor project activities and adjust as required
6. Keep good anecdotal records (p. 49)

Mandeville and Rivers (1989), in a study of the Hunter program in South Carolina, concluded that coaching by a principal or trainer was the single most important factor in the success of the program.

The ITIP model is composed of teaching methods that are recognized as effective by teachers and administrators. Wolfe (1987) has noted that the main abuse of the Hunter model has been the application of the seven steps in lesson planning as a recipe rather than as elements to be considered in planning instruction. Some of the teaching methods within the model have been validated by correlational research as indicated earlier, but the wide-spread acceptance of the model is not due to the support from research findings.

Hunter has many critics, but her model continues to grow in popularity in public schools. Several states, including South Carolina, Texas, and Pennsylvania, have conducted state-wide teacher training programs in the model. Much of the criticism of ITIP by teachers relates to the way in which the model has been implemented by administrators.

LEADERSHIP OF THE PRINCIPAL

The leadership of the principal is crucial to the operation of the school. Much of what we know today about leadership behavior is based on recent research. Research on effective management conducted in the 1960s emphasized the important role of leadership. Interest in this topic continues to grow, and new forms of leadership training have been developed. Much of this interest in the role of leadership grew out of the studies on the motivation of workers. Studies by Argyris (1964) and Likert (1967) demonstrated correlations between effective leader behavior and effective worker behavior. Burns (1978) described the effective leader as a transformational leader. A transformational leader is one who looks for potential motives in followers, seeks to satisfy higher needs, and, in the process, engages the full person of the follower. Transformational leadership has become increasingly important as schools face the major changes brought about by the clash of social, political, and economic forces.

Peter Drucker (1978) was one of the first management theorists to predict that organizations would move from an industrial-based environment to a knowledge-based environment. Further, he predicted that it would be necessary to abandon the management practices of an industrial age. He also predicted that workers in a knowledge-base environment would not be motivated by traditional management practices.

In a recent article, Drucker (1988) described the challenge faced by managers in information-based organizations:

Because the players in an information-based organization are specialists, they cannot be told how to do their work. There are probably few orchestra conductors who could coax even one note out of a French horn, let alone show the horn player how to do it. But the conductor can focus the horn player's skill and knowledge on the musicians' joint performance; and this focus is what the leaders of an information-based business must be able to achieve. (p. 47)

Today, principals are faced with this same issue as they attempt to evaluate and motivate teachers in a rapidly changing, information-based educational system. The principal's training in teacher evaluation is critical to the success of the school.

The principal's skill in conducting evaluation conferences is the critical factor in a successful evaluation program. The teacher's perception of the principal's motives, skills, and knowledge will determine the impact of the evaluation. Blumberg and Amidon (1965) studied teachers' perceptions of supervisor behavior in conferencing. They examined the direct and indirect behavior on the part of the supervisor during the conference. Direct behavior was defined as giving directions or comments, giving information or opinion, and giving critical comments. Indirect behavior was seen as praising or encouraging, accepting feelings and ideas, and asking questions. The assumption was that teachers would differ in their views of the value of the conference depending on their perceptions of the supervisor's behavior. The results of the study demonstrated that the conference was viewed as most productive when the supervisor emphasized indirect behavior and did not emphasize direct behavior. Teachers reported that they learned most

about themselves as individuals and as teachers when there was an emphasis on indirect supervisor behavior in conjunction with some direct supervisor behavior. However, they also reported that communication was limited when the supervisor emphasized high direct behavior with some low direct behavior.

Generally positive evaluations by teachers of the quality of their supervisory interpersonal relations appear to develop when a teacher perceives his supervisor's behavior as consisting of a heavy emphasis on both telling, suggesting, and criticizing, and on reflecting, asking for information, opinions, etc.; or when a teacher perceives his supervisor as putting little emphasis on telling and much on reflecting and asking.

Generally less positive or even negative evaluation by teachers of the quality of their supervisory interpersonal relationships appears to develop when a teacher perceives his supervisor as predominantly telling and not doing much reflecting or asking. (p. 169)

A study by Blumberg and Cusick (1980) examined the evaluation skills of the principal. In this investigation they found that many supervisors lacked the necessary skills for effective conferencing. Sahling (1981) reported that supervisors who were conferencing with teachers asked nine out of ten questions and tended to dominate the conference by structuring three out of four topics for discussion. Manatt, Palmer, and Hidlebaugh (1976) found that teacher's job performance was directly related to their perception of the competency of their supervisor. Blumberg (1974) described a pattern which was based upon teacher perceptions of supervisory behaviors:

It is evident that the teacher's perception of the competency of the principal in conducting evaluations will determine the value of the

conference. One of the key issues is the credibility of the principal in terms of knowledge of those things that have direct relevance to a teacher's evaluation such as knowledge of subject matter and a knowledge of the teacher's classroom. Stiggins and Duke (1988) found that the ability to inspire trust is another important element and they have identified several factors that relate to trust:

Supervisors' intentions (what they and the teacher regard as the ultimate purpose of evaluation).

Maintaining confidentiality in communication

How a supervisor handles evidence of performance from sources other than the classroom (e.g., hearsay and complaints)

The consistency with which the supervisor applies evaluation rules and regulations

The extent to which the teacher and supervisor see themselves as partners in the school improvement effort

The honesty and sincerity of interpersonal communications

The extent to which the teacher has an opportunity to interpret evaluation data first before sharing it with others

The extent to which teachers participate in the selection of performance goals (p. 85)

The training of the principal in teacher evaluation is of central importance to the success of the evaluation program. Competence in evaluation skills and the ability to inspire trust are the two most important characteristics of effective principals.

SUMMARY OF THE LITERATURE REVIEW

Teacher evaluation has developed over the years in a manner that parallels the development of management practices within business and industry. Teacher evaluation systems have moved from programs that emphasized compliance to formal tasks, management control, and conformity to bureaucratic mandates to systems that involve joint goal setting, observation of teaching, and conferencing. The characteristics of the evaluation determine the value that teachers place on the process. Teachers favor systems that use the techniques of clinical supervision, and they value specific feedback on their teaching. Reports in the literature indicate that less experienced teachers view the evaluation process more positively than do teachers with several years of experience. Darling-Hammond (1984) and McGreal (1986) reported that clinical supervision is an evaluation system that builds trust between principals and teachers, increases communication, and supports a high level of commitment to growth.

Principals tend to place a higher value upon evaluation as a means of improving instruction than do teachers. Some educators report that both teachers and principals view any evaluation as a threat and an assessment of their personal worth. Barber (1988) suggests that the most significant drawback to evaluation is the issue of aversive control. Stiggins (1988) and Popham (1988) identify the mixing of formative and summative evaluation as the most serious problem.

The work of Madeline Hunter has received mixed reviews from the educational research community, but it continues to grow in popularity in

public schools across the country. Many of the problems associated with the ITIP model appear to be created by the manner in which the model has been implemented in some school districts.

Finally, the literature reviewed supported the need for a clear understanding of the perceptions of both teachers and administrators of the evaluation process. Research related to the characteristics of effective evaluation models, the attitudes of teachers and principals toward teacher evaluation, the ITIP model of instruction, and the evaluation skills required by principals have been reported by previous investigators. However, specific questions and investigations related to the perceptions of teachers and administrators about their experience with ITIP and teacher evaluation are lacking in the research literature. The intent of this study is to examine the perceptions of teachers and administrators concerning the quality, impact, and value of teacher evaluation in two school districts which used the ITIP model of instruction over several years. Chapter III presents the methodology for this study.

CHAPTER III

METHODOLOGY

INTRODUCTION

This chapter describes the population studied in the investigation and the procedures used in the collection and analysis of data from the study. The following sections are included in this chapter: (1) Sample; (2) Research Instruments; (3) Data Collection; (4) Data Analysis; and (5) Research Questions.

SAMPLE

The population of this study consisted of the elementary teachers and principals of two suburban school districts near Portland, Oregon. The 233 elementary teachers and the 14 elementary principals participating in this study completed a questionnaire that described their most recent experiences in teacher evaluation. Teachers in the study were classroom teachers who taught at least half time. Teachers with less than a half-time assignment were not included in the study because their evaluation experiences were too limited for the purposes of the study. Specialists such as counselors and special education teachers were excluded from the study because they were evaluated by both central office administrators and building principals. Principals in the study worked full time in a single school building and were responsible for the evaluation of all classroom

teachers in that building.

The two school districts in this study were selected on the basis of their experience with teacher evaluation and the use of the ITIP model of instruction. A phone survey was conducted with school districts in the region in an attempt to find districts with extensive experience with both topics. The two districts selected for this study have the following characteristics:

1. Each of the two districts have over twelve years of experience in using ITIP as their basic model of instruction and evaluation.
2. All elementary principals in these districts were trained in ITIP and the Hunter model of teacher evaluation.
3. Both districts made a strong financial commitment to staff development programs for teachers and administrators in teacher evaluation and ITIP.
4. The districts have maintained a sustained effort over several years to train new teachers and principals in ITIP and evaluation.
5. The districts have established written standards for teacher evaluation which specify ITIP as the basic model for instruction.

Table I, which describes the teachers' experience levels, shows that about 60% of the teachers in this study had over six years of experience in teaching using the ITIP model. In addition both districts had conducted extensive in-service programs for principals in the Madeline Hunter model of teacher evaluation.

Table II, which describes the principals' experience levels, shows that 85% of the principals completed three or more classes in ITIP and that 80% of them had worked with ITIP for six years or more.

TABLE I

FREQUENCY/PERCENTAGE DISTRIBUTIONS FOR TEACHING EXPERIENCE,
EMPLOYMENT STATUS, AND ITIP TRAINING FOR TEACHERS

Variable	Category	Frequency	Percentage*
Years of Experience	0 to 1	10	4.3
	2 to 5	39	16.7
	6 to 10	38	16.3
	11 to 15	57	24.5
	16 or more	89	38.2
Employment Status	Probationary	48	28.7
	Tenured	185	71.3
Years of experience with ITIP	0 to 1	19	8.2
	2 to 5	75	32.2
	6 to 10	102	43.8
	11 to 15	22	9.4
	More than 15	15	6.4
Classes taken in ITIP	No Classes	9	3.9
	One Class	57	24.5
	Two classes	84	36.1
	Three Classes	48	20.6
	More than three	35	15.0

*The percentage is the frequency of teachers in the category divided by the number of teachers in the sample (N = 233).

TABLE II

**FREQUENCY/PERCENTAGE DISTRIBUTION FOR ADMINISTRATIVE EXPERIENCE,
EVALUTION TRAINING, AND ITIP EXPERIENCE FOR PRINCIPALS**

Variable	Category	Frequency	Percentage*
Years of Experience in	0 to 1	1	7.1
	2 to 5	2	14.3
	6 to 10	9	64.3
	11 to 15	1	7.1
	16 or more	1	7.1
Clock Hours of Training in Teacher Evaluation	0 to 10	0	0.0
	11 to 40	1	7.1
	41 to 80	2	14.3
	81 to 160	6	42.9
	More than 160	5	35.7
Years of Experience with ITIP	0 to 1	1	7.1
	2 to 5	2	14.3
	6 to 10	9	64.3
	11 to 15	2	14.3
	More than 15	0	0.0
Classes Taken in ITIP	No classes	0	0.0
	One class	1	7.1
	Two classes	1	7.1
	Three classes	2	14.3
	More than three	10	71.4

* The percentage is the frequency of teachers in the category divided by the number of teachers in the sample (N = 233).

THE RESEARCH INSTRUMENTS

This section describes the key components of the two questionnaires used in this study. The Teacher Evaluation Profile (TEP) was developed by Stiggins and Duke (1988) at the Northwest Regional Laboratory. The TEP was modified for the purposes of this study. A copy of the modified TEP can be found in Appendix A.

The Principal Evaluation Profile (PEP) was adapted from the TEP with the permission of the Northwest Regional Educational Laboratory. The items in the PEP were designed to parallel the items in the TEP as closely as possible and were field tested before the instrument was used in this study. A copy of the PEP can be found in Appendix B.

The Teacher Evaluation Profile

Instructions. This section of the questionnaire asked teachers to respond to questions about their current teaching assignment. They were informed that all responses would remain anonymous and they were asked to write in a code number for their school district and for their school. The code number had been assigned to each school within the district. In the last part of this section they were asked three questions about their training in ITIP.

Criterion questions. In this section teachers were asked to describe their most recent evaluation experience in terms of three key questions. The first question asked them to rate the quality of their most recent evaluation experience. The second question inquired about the overall impact of the evaluation upon their teaching practices. A third question, added to the

original TEP for the purposes of this study, asked teachers to rate the value they place upon evaluation as a source of professional growth.

General instructions. This section described the procedure for marking the questionnaire and displayed a sample question.

Attributes of the teacher. The nine questions in this section asked teachers to rate themselves on their instructional competence, personal expectations, openness to change, subject knowledge, and experience.

Attributes of the evaluator. Teachers were asked to rate the evaluator on items related to credibility, persuasiveness, patience, and trust. There were twelve items in this section.

Procedures of the evaluation. The nine items in this section asked the teacher's perceptions of the procedures used in the most recent evaluation. These items described performance criteria, standards, and data collection procedures.

Attributes of the feedback. The nine items in this section described characteristics of the feedback that teachers might receive in the evaluation conference.

Attributes of the evaluation context. The amount of time spent on the evaluation process by both the teacher and the evaluator and the availability of resources for growth were the two key topics examined in this section.

ITIP experience. The five items in this section were added to the original TEP for the purposes of this study. These items asked teachers to describe their experience with ITIP.

The Design of the Instrument

The Teacher Evaluation Profile (TEP) was developed from a series of case studies on effective teacher evaluation practices conducted by Duke and Stiggins (1986). The instrument has been used in a wide variety of school districts across the nation. The most recent technical analysis of the TEP was conducted by the Northwest Regional Education Laboratory during the 1987-88 school year. This study, conducted by Stiggins and Nickel (1988), involved over 3,000 teachers in 27 school districts from the states of Connecticut, Montana, Ohio, Oregon, and Washington.

The first part of the analysis examined questionnaire items and subscale intercorrelations. Instrument and subscale internal consistency reliability estimates were computed, as were subscale intercorrelations. In addition, the item intercorrelation matrix was factor analyzed to examine the factor structure of the TEP. For additional information see Stiggins and Nickel (1989).

Correlations among questionnaire items were also examined. The technical analysis reviewed coefficient alpha estimates of the internal reliability of the five subscales and estimates of the intercorrelations among scales. Reliabilities tended to be consistently high and scale intercorrelations were moderate. The exception was the teacher scale, which was less reliable and appeared to be statistically independent of the other scales. However, the internal reliability of the total instrument was .93.

A simple bivariate analysis of the relationships between all items and all criterion ratings was calculated. Only the items that were most highly correlated with the quality and impact ratings were retained. They were the

same items for each criterion, and therefore, for the combined rating. Attributes of the evaluator (items 10, 17, 20, and 21) and attributes of the feedback (31, 34, 35, and 36) consistently were more highly correlated with the teachers' perceptions of the quality and impact of an evaluation experience. However, every TEP item reflecting attributes of the evaluator, procedures, feedback, and context yielded a statistically significant correlation with the quality and impact ratings. For additional information see Stiggins and Nickel (1989).

In terms of technical characteristics, the TEP meets the standards for validity and reliability required to measure the climate for teacher evaluation within a school district. It is valid in that it provides data on attributes of teacher evaluation that have been shown in the research literature to be related to teacher growth and development. It is reliable in that it produces internally consistent data on those attributes. In addition to the technical characteristics discussed, the TEP is efficient to administer and to score. The questionnaire takes about 15 minutes for the teacher to complete, and it can be administered in large-group contexts.

Modification of the TEP

The original TEP had two criterion questions about the quality and impact of the evaluation. A third question about the value that the teachers placed on the evaluation experience was added for the purpose of clarifying the differences between quality, impact, and value. Because this study examined teacher attitudes toward ITIP, five questions which asked about this model of instruction were added to the questionnaire. In addition, items were added to the TEP which describe the current assignment of the

teachers and their training and background in ITIP. A parallel instrument based on the modified version of the TEP was developed and administered to the administrators in the study.

The items added to the TEP were field tested with a sample of teachers who were not involved in the study and modified based on the results of the field test. The items then were reviewed and approved by representatives of the Northwest Regional Educational Laboratory and by Dr. John Lind and Dr. Loyde Hales of Portland State University.

Permission to conduct this study was obtained from the Portland State University Human Subjects Research Review Committee (HSRRC).

Coefficient alphas for the modified TEP were calculated. The results of the analysis can are presented in Chapter IV.

The Principal Evaluation Profile

Instructions. This section asks the principals to describe themselves as evaluators of teachers. They were asked to code their questionnaire with a school number and were told that their responses would remain anonymous. They were asked to describe their training in teacher evaluation and in ITIP.

Criterion questions. In this section principals were asked to rate the quality of the evaluations they conducted in the current year and to also rate the impact and value of the evaluation process.

General instructions. This section described the procedures for marking the questionnaire and displayed a sample question.

Attributes as a teacher evaluator. Principals were asked to rate themselves on their skills in helping teachers to grow through evaluation. These question were parallel to the questions on the TEP.

Attributes as an evaluator as seen by teachers. Principals were asked to give their perceptions of how teachers would rate them in their skills as an evaluator. These questions were the same as those in Section B of the TEP.

Procedure of the evaluation. Principals were asked to describe the procedures used during the teacher evaluation process.

Attributes of the feedback. The items in this section were the same as the items on the TEP except for item 35, which asked if the principal sent the teachers written communication after informal observations.

Attributes of the evaluation context. The items in this section were the same as the items for this section in the TEP.

ITIP experience. The items in this section were parallel to those in the TEP and asked principals to respond from their perspective.

Design of the Principal Evaluation Profile

The PEP was designed to parallel the TEP and to collect data on the same items from the perspective of the principal. The PEP was field tested and modified based on the results of the field test. The items were reviewed and approved by representatives of the Northwest Regional Educational Laboratory and by Dr. John Lind and Dr. Loyde Hales of Portland State University. The PEP was also reviewed by the Human Subjects Research Review Committee (HSRRC) at Portland State University. The instrument was also reviewed and approved for use by central office administrators in the

school districts in the study. Coefficient alphas for the scales in the PEP were calculated. The results of this analysis are presented in Chapter IV.

DATA COLLECTION

The purpose of the study and the modified version of the TEP was reviewed with the curriculum directors, staff development administrators, and superintendents of the two districts involved in the study. Permission to conduct the study was granted by the superintendents of the districts in April of 1988. To ensure that both individual and school responses would remain anonymous, principals in each district were asked to select one principal to assign a school code number for each building. The principal agreed not to disclose the code number and to distribute and collect the envelopes containing the TEP.

During the field test it was found that teachers had difficulty in using the Scantron marking sheets. Therefore, the teachers and principals in the study were asked to mark their responses directly on the TEP booklet. The responses were later transferred to Scantron sheets for computer analysis.

The TEP was administered during one faculty meeting. The teachers who were absent were asked to send the questionnaire to the school secretary in a plain envelope. Questionnaires were collected from each building by the designated principal and were returned to the researcher.

DATA ANALYSIS

The data from the two questionnaires were transferred from the TEP booklets and coded on Scantron sheets for computer analysis. Mean scores and standard deviations were calculated for the responses of teachers and principals to each item of the TEP and the PEP. In addition, mean scores and standard deviations were calculated for the criterion questions (Quality, Impact, and Value) and for the scale scores for the TEP and the PEP. Correlations between teacher scores on the TEP and the scores on the criterion questions were also calculated. The mean scores of principals and teachers on each scale of the TEP were compared and significant differences ($p < .05$) were identified using ANOVA. Item scores were reported for significant scales.

TEP scale scores were correlated with criterion questions from the TEP which described the Quality, Impact and Value of the evaluation. In addition, differences in the amount of training in ITIP, and differences between the responses of probationary and tenured teachers were analyzed. Principals' scores on the PEP were compared to teachers' scores on the TEP scales and items using ANOVA and significant differences were identified at the $p < .05$ level. Item scores for significant scales were reported. Principal responses were compared with teacher responses on scale scores using ANOVA and significant differences at the $p < .05$ level were reported.

RESEARCH QUESTIONS

Data from the teacher and administrator questionnaires were analyzed to investigate the following questions:

1. Is the value that teachers place on teacher evaluation related to the characteristics of the evaluation?

Teacher responses to the three criterion questions on impact, quality, and value were examined in the first phase of the analysis of this question. The second phase of the analysis examined the relationships between teachers' responses to the forty-nine TEP items and their rating of the criterion questions. This item/criterion relationship was explored by examining the correlations between TEP items and the criterion ratings of quality, impact, and value.

2. Does the amount of training that teachers receive in ITIP affect the value they place on teacher evaluation?

ANOVA technique was used in the first phase of the analysis of this question to explore the relationship between the number of ITIP courses taken by teachers and their rating of the criterion questions. In the second phase of the analysis ANOVA was used to examine the relationship between the teachers' rating of the six scale scores on the TEP and the number of ITIP courses completed by teachers. Item scores for each of the significant scales were reported. Mean scores for significant items were compared with the number of ITIP classes taken.

3. Do probationary and tenured teacher have different perceptions about the nature and value of teacher evaluation?

An analysis of variance technique was used to examine differences between the mean scores of probationary and tenured teachers on each of

the six scales of the TEP. Item scores for significant scales were reported. Criterion questions of quality, impact, and value were compared with scale scores using ANOVA. Mean scores and standard deviations for probationary and tenured teachers were compared.

4. Do principals and teachers have different perceptions about the nature and value of teacher evaluation?

Differences in the scores of principals and teachers on the the criterion questions about quality, impact, and value were compared using the ANOVA technique. Differences in mean scale scores for principals and teachers was also examined using ANOVA technique. Individual items for significant scales were reported.

In addition to the the four main research questions, a supplemental research question was explored. This question was stated as follows: How similar are the research findings from the two school districts with the national findings compiled by the Northwest Regional Educational Laboratory?

CHAPTER IV

RESULTS

INTRODUCTION

The purpose of this study was to examine the perceptions of elementary teachers and principals concerning teacher evaluation within two school districts which use ITIP as their basic instructional format. The instruments used in this study were a modified version of the Teacher Evaluation Profile (TEP) and a parallel instrument for principals the Principal Evaluation Profile (PEP). The instruments were used to examine the following factors in the evaluation process: the teacher, the evaluator, the procedures, the feedback, the context, and the teachers' experience with the ITIP model of instruction. These attributes were compared with the subjects' criterion ratings of the quality, impact, and value of the evaluation process. In addition, the perceptions of probationary teachers, tenured teachers, and principals were examined.

The following sections are included in this chapter: (1) Teacher Respondent Information; (2) Total Teacher Responses; (3) Principal Respondent Information; (4) Total Principal Responses; (5) Analysis of the Research Questions; and (6) Summary of Chapter IV.

TEACHER RESPONDENT INFORMATION

Eleven of the 271 teachers in the original sample were excluded from the study because they worked less than half time or were specialists. Of the 260 elementary teachers remaining in the sample , 233 responded. The response ratio for teachers is 233/260, or 90%.

TOTAL TEACHER RESPONSES

In response to the general research question on teacher attitudes toward teacher evaluation, this section will examine the results of the responses of all the elementary teachers in the study. Included in this examination is an analysis of ITIP teaching experience and training, and the sample size, mean, and standard deviation for each of the 49 items in the six sections (A-F) of the modified TEP: (A) Attributes of the Teacher; (B) Perceptions of the Evaluator; (C) Attributes of the Procedures; (D) Attributes of the Feedback; (E) Attributes of the Evaluation Context; and (F) ITIP Experiences.

Table III shows that approximately 60% of the teachers in the study had over five years of experience with ITIP, 70% had taken two or more classes in ITIP and 55% had taken seven hours or more in ITIP workshops.

TABLE III
TEACHER EXPERIENCE WITH ITIP

Experience	Frequency	Percent	Cum. Percent
Years Using ITIP			
0 -1	19	8.2	8.2
2 - 5	75	32.2	40.3
6 -10	102	43.8	84.1
11 - 15	22	9.4	93.6
More then 15	15	6.4	100.0
Number of ITIP Classes			
No Classes	9	3.9	3.9
One Class	57	24.5	28.3
Two Classes	84	36.1	64.4
Three Classes	48	20.6	85.0
More Than three	35	15.0	100.0
Hrs. ITIP Workshops			
0	37	15.9	15.9
1 - 6	70	30.0	45.9
7 - 12	38	16.3	62.2
13 -18	18	7.7	70.0
More than 18	70	30.0	100.0

Teacher Attributes

Table IV shows the results of the teachers' rating of the attributes they bring to the evaluation process. Teachers in this study tended to score themselves highest on their level of professional expectations (Item 1 - 4.7). They also rated their knowledge of subject matter in the higher range. (Item 7 - 4.3). Additionally, orientation to change (Item 3 -4.3), orientation to experimentation in the classroom (Item 4 - 4.2), and knowledge of technical aspects of teaching (Item 6 - 4.1), were rated above 4.0 on a five-point scale. The lowest mean score was a 3.5 on Item 9 which asks the teachers to describe the usefulness of their teacher evaluations prior to their most recent evaluation experience.

Item 8 asks teachers to describe their years of teaching experience. This item does not fit into a scale that measures teacher attributes in terms of knowledge of subject matter and orientation to change. Item 8 should appear in the first section of the profile which describes demographic data. The items in Section A should constitute a psychological construct which measures the attributes of the teacher. It should also be noted that the mean and the standard deviation for Item 8 represent values for the scale, not the number of years of teaching experience. The median years of experience is nine years.

The low standard deviations indicate that teachers varied little among themselves in their responses to the items.

TABLE IV
RESPONSE OF TEACHERS TO THE ITEMS DESCRIBING
THE ATTRIBUTES OF THE TEACHER

Item	N	Mean	S. D.
1. Rate the strength of your professional expectations of yourself.	233	4.700	0.521
2. Orientation to risk taking	233	3.923	0.832
3. Orientation to change	233	4.300	0.751
4. Orientation to experimentation in classroom	233	4.240	0.739
5. Openness to criticism	232	3.823	0.914
6. Knowledge of technical aspects of teaching	233	4.082	0.793
7. Knowledge of subject matter	229	4.380	0.635
8. Years of teaching experience	233	3.755	1.244
9. Experience with teacher evaluation prior to most recent experience	232	3.466	1.023

Attributes of the Evaluator

Table V shows the teachers' perceptions of the attributes of the evaluator. Teachers rated the evaluator, (in this study the elementary principal), high on credibility as a source of feedback (Item 10 - 4.1). In addition, they reported that their working relationship with the principal tended to be a helping relationship (Item 11 - 4.4). They also reported that the principal was viewed as trustworthy (Item 12 - 4.3).

The interpersonal manner (Item 13 - 4.4), temperament (Item 14 - 4.3) and flexibility (Item 15 - 4.1) of the principal were rated at 4.0 or above on a five-point scale.

The principals' knowledge of technical aspects of teaching had the highest mean score (Item 16 - 4.4). Teachers reported that the principals displayed the capacity to demonstrate or model needed improvement (Item 17 - 3.8), were familiar with the teachers' classrooms (Item 18 - 3.7), and had experience with classrooms in general (Item - 4.1).

The teachers rated the principal's ability to offer practical suggestions for improvement in the positive range (Item 20 - 3.9). Teachers were also persuaded by the principals' rationale for suggestions for improvement (Item 21 -3.9).

TABLE V
RESPONSE OF TEACHERS TO THE ITEMS DESCRIBING THE
ATTRIBUTES OF THE EVALUATOR

	N	Mean	S. D.
10. Credibility as a source of feedback	231	4.104	0.959
11. Working relationship with you	233	4.369	0.891
12. Level of trust	233	4.330	1.008
13. Interpersonal manner	232	4.358	0.947
14. Temperament	231	4.303	0.944
15. Flexibility	231	4.082	1.007
16. Knowledge of technical aspects of teaching	233	4.425	0.751
17. Capacity to demonstrate or model needed improvements	231	3.814	1.015
18. Familiarity with our classroom	233	3.712	1.066
19. Experience with classrooms	233	4.086	0.961
20. Usefulness of suggestions	230	3.887	0.974
21. Persuasiveness of rationale for suggestions	227	3.811	0.989

Attributes of the Procedures

Table VI shows the response of teachers to the items in Section C dealing with the procedures of teacher evaluation. Teachers reported that standards for evaluation were communicated to them (Item 22 - 3.6), were clear (Item 23 - 3.8) and were generally endorsed by them (Item 24 - 3.8). They were less certain that the standards were the same for all teachers and fifteen teachers did not respond to this question (Item 25 - 3.3). Observation of classroom performance was the main source of information collected for the purposes of teacher evaluation (Item 26 - 4.2).

School records (Item 27 - 2.2) and reports describing student achievement (Item 28 - 2.7) were not often used as sources of information. Teachers reported receiving less than two formal observations per school year (Item 29 - 2.7) and less than one informal evaluation observation per month (Item 30 - 2.6).

Classroom observation was the primary source of data collection for teacher evaluation (Item 26), and the goal of classroom observation was to obtain a representative sample of teacher performance in order to draw conclusions about a teacher's general level of competency.

TABLE VI
 RESPONSE OF TEACHERS TO THE ITEMS DESCRIBING
 THE ATTRIBUTES OF THE PROCEDURES

Item	N	Mean	S. D.
22. Were standards communicated to you?	231	3.576	1.056
23. Were standards clear to you?	229	3.760	1.100
24. Were standards endorsed by you as appropriate for your classroom?	224	3.799	1.080
25. Were the standards the same for all teachers.	214	3.346	1.089
26. Observation of your teaching performance.	231	4.203	0.945
27. Examination of school records.	228	2.241	1.141
28. Examination of student achievement.	229	2.672	1.233
29. Number of formal observations.	231	2.684	0.813
30. Number of informal observations.	231	2.649	0.966

Attributes of the Feedback

Teacher responses shown in Table VII demonstrate that teachers were generally satisfied with amount of information received in the feedback (Item 31 -3.6). However, the frequency of formal and informal feedback received the lowest ratings of any of the items in this section (Items 32 - 3.1 and 33 - 2.8).

The depth of information provided in the feedback was ranked by teachers in the upper range of the five-point scale (Item 34 - 3.5). Teachers in this study ranked the quality of ideas in the upper range of the scale (Item 35 - 3.7). The specificity of the information provided in the feedback conference was also ranked in the upper range (Item 36 - 3.8).

Teachers' perceptions concerning the nature of the information provided in the feedback, (Item 37 - 3.8) were consistent with their rankings of the quality and specificity of the feedback. Timing of the feedback was also ranked in that same range (Item 38 - 3.8).

The item which received the highest ranking in this section described the extent to which the feedback focused on the district teaching standards (Item 39 - 4.0).

TABLE VII
RESPONSE OF TEACHERS TO THE ITEMS DESCRIBING
THE ATTRIBUTES OF THE FEEDBACK

Item	N	Mean	S. D.
31. Amount of information received	232	3.629	.0963
32. Frequency of formal feedback	232	3.060	1.194
33. Frequency of informal feedback	231	2.848	1.254
34. Depth of information provided	231	3.476	1.008
35. Quality of the ideas and suggestions contained in the feedback	231	3.714	1.002
36. Specificity of information provided	231	3.765	1.056
37. Nature of information provided	231	3.818	1.039
38. Timing of the feedback	228	3.754	1.150
39. Feedback focused on district teaching standards	228	3.991	.984

Attributes of the Context

The teachers indicated that the amount of time spent on the evaluation process by themselves and by all the other participants in the process was substantial (Item 40 - 3.4). However, teachers in the study ranked the time allotted during the teaching day for professional growth much lower than any other item in this section (Item 41 - 2.2).

They indicated that training programs and models of good practice were generally available (Item 42 - 3.6). In addition, the clarity of policy statements regarding the purpose for evaluation (Item 43 - 3.7) and the intended role of evaluation (Item 44 - 3.8) were both ranked in the upper range as shown in Table VIII.

TABLE VIII
RESPONSE OF TEACHERS TO THE ITEMS DESCRIBING THE
ATTRIBUTES OF THE EVALUATION CONTEXT

Item	N	Mean	S. D.
40. Amount of time spent on the evaluation process including your time and that of all other participants	233	3.412	0.882
41. Time allotted during the teaching day for professional development	232	2.168	1.140
42. Availability of training programs and models of good practice	232	3.569	1.432
43. Clarity of policy statements regarding purpose for evaluation	233	3.657	1.100
44. Intended role of evaluation	233	3.790	1.104

Attributes of the ITIP Experience

Table IX shows that teachers in the study usually followed ITIP procedures in their classes (Item 45 - 3.8). Teachers also indicated that their ITIP skills were evaluated (Item 46 - 4.0) and that they valued having a common vocabulary for discussing instruction (Item 47 - 3.9). They reported a high level of commitment to ITIP (Item 48 - 3.7) and that they placed value on ITIP as a means for improving instruction (Items 49 - 3.8).

TABLE IX
RESPONSE OF TEACHER TO THE ITEMS DESCRIBING
THEIR EXPERIENCE WITH ITIP

Item	N	Mean	S. D.
45. The extent to which I follow ITIP procedures in my classes.	233	3.773	0.673
46. The extent to which my ITIP skills are evaluated by my principal.	233	3.966	0.978
47. The value of having a common vocabulary in ITIP (e.g. guided practice) for discussions in my evaluation conferences.	232	3.940	1.070
48. My level of commitment to ITIP	233	3.717	0.859
49. The value I place on ITIP for improving instruction.	232	3.806	0.926

Criterion Questions

The criterion questions form an important benchmark for examining the relationship between the attributes of a teacher evaluation and its perceived effect on teachers. Teachers were asked to rate the Quality, Impact, and Value of the evaluation experience on a ten-point scale. As Table X shows, they rated Value at a mean of 7.1 and Quality at 7.0. However, the mean score for Impact was only 5.4.

TABLE X
RESPONSES OF TEACHERS TO THE CRITERION QUESTIONS ON
QUALITY,IMPACT AND VALUE OF TEACHER EVALUATION

Criterion Question	N	Mean	S. D.
Regard the entire evaluation process, including planning for evaluation, classroom observations and feedback. How would you rate the overall quality of the evaluation?	229	6.978	1.893
Rate the overall impact of your last evaluation experience on your professional practices.	232	5.444	2.366
Rate the value you place upon the process of teacher evaluation as a source of growth.	233	7.112	2.151

Scale Characteristics of Sections A-F

Teacher responses were also analyzed in order to seek information related to the dependability of the items and scales in the questionnaire. Coefficient alphas were computed to estimate the internal consistency of the items in each of the six sections of the modified TEP.

The mean, standard deviation, and scale reliabilities of Sections A-F of the modified TEP are shown in Table XI. The scale reliabilities ranged from .61 in Section A to .92 in Section B. The standard deviation ranged from .064 in Section B to .470 in Section C. The means ranged from 4.104 on Section B to 3.211 in Section C.

TABLE XI
MEANS, STANDARD DEVIATIONS, AND RELIABILITY COEFFICIENTS FOR
THE TOTAL TEACHER SAMPLE ON THE SECTIONS OF THE TEP

Section	Mean	S. D.	Coefficient Alpha
A. Attributes of the teacher	4.072	.142	.605
B. Attributes of the evaluator	4.105	.064	.915
C. Attributes of the procedures	3.211	.469	.749
D. Attributes of the feedback	3.561	.133	.889
E. Attributes of the context	3.322	.433	.622
F. Attributes of ITIP	3.838	.012	.808

PRINCIPAL RESPONDENT INFORMATION

A modified version of the TEP was designed, field tested, and administered to the principals. The questions in Section A were modified to reflect the role of the principal in supporting teacher growth and, therefore, responses of teachers and principals are not compared for Section A in the later analysis of the data. Question 35 was also modified to gain information about principals' communication to teachers regarding informal classroom observations. This question was excluded from comparative analysis with the teacher data.

All of the principals in the fourteen schools participated in the study for a response rate of 100 percent. The principals' experience with evaluation and ITIP is reported in Table XII. Principals in the study tended to have extensive training in the teacher evaluation process with five of them reporting more than 160 contact hours of coursework in this subject.

Only three of the principals reported having less than five years of experience in working with ITIP. Two of the principals had over ten years of experience with the ITIP model.

Twelve of the fourteen principals had taken three or more classes in ITIP training through district inservice or college classes. In addition, seven of the principals reported more than 18 hours of ITIP workshops.

TABLE XII

PRINCIPALS' EXPERIENCE WITH EVALUATION AND ITIP

Experience	Frequency	Percent	Cum. Percent
Hrs. Evaluation Training			
0 - 10	0	0.0	0.0
11 - 40	1	7.1	7.1
41 - 80	2	14.3	21.4
81 - 160	6	42.9	64.3
More than 160	5	35.7	100.0
Years Using ITIP			
0 - 1	1	7.1	7.1
2 - 5	2	14.3	21.4
6 - 10	9	64.3	85.7
11-15	2	14.3	100.0
More than 15	0	0.0	0.0
Number of ITIP Classes			
No Training	0	0.0	0.0
One Class	1	7.1	7.1
Two Classes	1	7.1	14.3
Three	2	14.3	28.6
More than Three	10	71.4	100.0
Hrs. ITIP Workshops			
0	0	0.0	0.0
1 to 6	2	14.3	14.7
7 to 12	2	14.3	28.6
13 to 18	3	21.4	50.0
More than 18	7	50.0	100.0

TOTAL PRINCIPAL RESPONSE

This section presents the sample size, mean, and standard deviation for the responses of all the elementary principals in the fourteen schools in the study. Included in the discussion is an analysis of the sample size, mean, and standard deviation for each of the 49 items in the six sections of the TEP and the three criterion questions on quality, impact and value. The mean scale score, the standard deviation and the coefficient alpha for each of the six scales is also presented.

Principals' Support of Teacher Growth

Table XIII shows the principals' perceptions of their ability to help teachers grow in the nine teacher characteristics identified in the TEP. The questions in this section ask the principals to describe their ability to support teachers in their professional growth. Because of the difference in the roles between principals and teachers, the questions in this section are not as comparable as they are in the other sections of the TEP.

Principals rated themselves highest on their ability to help teachers experiment in the classroom (Item 4 - 4.4), take risks (Item 2 - 4.1), and change (Item Item 3 - 4.0) . They rated themselves the lowest on their ability to help teacher increase their knowledge of subject matter (Item 7 - 3.3) The figures listed for Item 8 are the mean and standard deviation of the item score, not the years of teaching experience.

TABLE XIII
 RESPONSE OF PRINCIPALS TO THE
 ITEMS IN SECTION A OF THE TEP

Item	N	Mean	S. D.
1. Ability to help teachers to set high professional expectations	14	3.929	0.730
2. Ability to encourage teachers to take risks	14	4.071	0.730
3. Ability to help teachers change	14	4.000	0.784
4. Ability to help teachers to experiment in the classroom	14	4.375	0.663
5. Ability to help teachers receive criticism	14	3.643	0.842
6. Ability to help teachers increase their knowledge of technical aspects of teaching	14	3.786	0.975
7. Ability to help teachers increase their knowledge of subject matter	14	3.286	0.726
8. Years of experience as a principal	14	2.929	0.917
9. Experience with teacher evaluation prior to most recent experience	14	3.643	1.008

Principals As Teacher Evaluators

In this section the principals described their perceptions of how the teachers they evaluated would assess the principal's skills as a teacher evaluator. As shown in Table XIV, they gave themselves a high marks on their credibility as a source of feedback (Item 10 - 4.0) Their working relationship with the teachers was rated the highest in Section B (Item 11 - 4.5). The principals predicted that teachers would also rate them high on the level of trust (Item 12 - 4.4).

They rated themselves above 4.0 on their interpersonal manner (Item 13 - 4.1), temperament (Item 14 - 4.2), and flexibility (Item 15 - 4.1). Knowledge of technical aspects of teaching (Item 16 - 3.9) and capacity to demonstrate or model needed improvements (Item 17 - 3.6) were ranked somewhat lower by principals. Familiarity with teachers' classrooms (Item - 18 - 3.9) and experience with classroom in general (Item - 20 - 3.8) were in the upper range.

Principals also ranked themselves somewhat lower on the usefulness of suggestions for improvements (Item 20 - 3.8 and their ability to be persuasive in giving teachers a rationale for needed improvements (Item 21 - 3.6)

TABLE XIV
RESPONSE OF PRINCIPALS TO THE
ITEMS IN SECTION B OF THE TEP

Item	N	Mean	S. D.
10. Credibility as a source of feedback	14	4.000	0.679
11. Working relationship	14	4.500	0.519
12. Level of trust	14	4.429	0.514
13. Interpersonal manner	14	4.143	0.770
14. Temperament	14	4.214	0.579
15. Flexibility	14	4.143	0.770
16. Knowledge of technical aspects of teaching	14	3.857	1.027
17. Capacity to demonstrate or model needed improvements	14	3.571	1.158
18. Familiarity with teachers' classrooms	14	3.857	0.864
19. Experience with classrooms in general	14	4.000	0.555
20. Usefulness of your suggestions for improvements	14	3.786	0.579
21. Persuasiveness of your rationale for improvements	14	3.571	0.514

Principals and Procedures

As shown in Table XV, the principals identified the observation of classroom performance as the primary source of performance information used in the evaluation process (Item 26 - 4.4). They also reported that the examination of student achievement was not used extensively (Item 28 - 3.1). They also reported that the examination of classroom records such as lesson plans and other records was not used to the same extent as classroom observation or the examination of student achievement. (Item 27 - 2.4).

Principals reported that fewer than two formal observations per teacher were conducted each year (Item 29 - 2.6 - 1.6 observations) and that teachers received informal observations more than once each month (Item 30 - 3.4 - about once each month). The reports of teachers in this study on the number of formal observations matched the principals' reports. However, teachers in the study reported receiving fewer informal observations.

The standards to be used in the evaluation process were communicated to teachers (Item 22 - 4.1) and they were clearly understood by the teachers (Item 23 - 4.0). Principals also reported that the standards were generally endorsed by the teachers as appropriate (Item 24 - 3.6) and that they were generally the same for all teachers (Item 25 - 3.4).

TABLE XV
RESPONSE OF PRINCIPALS TO THE
ITEMS IN SECTION C OF THE TEP

Item	N	Mean	S. D.
22. Were standards communicated to teachers?	14	4.071	0.730
23. Were the standards clearly understood by the teachers?	14	4.000	0.748
24. Were the standards endorsed by the teachers as appropriate for their classrooms?	14	3.643	0.745
25. Were the standards the same for all teachers?	14	3.385	1.502
26. Observation of classroom performance	14	4.353	0.633
27. Examination of classroom or school records (lesson plans, etc.)	14	2.429	0.938
28. Examination of student achievement	14	3.071	1.141
29. Approximate number of formal observations per teacher this year	14	2.571	0.852
30. Approximate number of informal observations per teacher this year	14	3.357	0.633

Principals and Feedback

As shown in Table XVI the extent to which the feedback focused on district teaching standards was one of the two top items rated by principals (Item 39 - 3.9). Principals also reported that the nature of the information provided to the teacher in the evaluation conference was descriptive rather than judgmental (Item 37 - 3.9).

It should be noted that Item 35 on the administrator questionnaire does not match Item 35 on the teacher questionnaire. Precautions were taken during the analysis of the data not to compare the two items, and Item 35 was excluded from the summaries of data used in the study. Item 35 on the administrator questionnaire asked principals how often written communication was sent to the teacher after an informal observation. The principals reported that they seldom sent written communications to teachers after an informal observation (Item 35 -2.7)

The timing of the feedback was reported by principals to be more immediate than delayed (Item 38 - 3.8). They also reported that they provided teachers with a great deal of information (Item 31 - 3.7) and that the information was specific (Item 36 - 3.7).

TABLE XVI
RESPONSE OF PRINCIPALS TO THE
ITEMS IN SECTION D OF THE TEP

Item	N	Mean	S. D.
31. Amount of information given to teachers	14	3.714	0.611
32. Frequency of formal feedback	14	3.429	1.089
33. Frequency of informal feedback	14	3.429	1.016
34. Depth of information provided	14	3.643	0.745
35. Written communication sent to teacher after informal observation	14	2.714	1.326
36. Specificity of information provided	14	3.714	0.914
37. Nature of information provided	14	3.929	0.730
38. Timing of the feedback	14	3.786	0.975
39. Feedback focused on district teaching standards	14	3.929	0.917

Principals and the Evaluation Context

Principals' ratings on items related to evaluation context as shown in Table XVII indicate that the clarity of the district policy statement regarding the purposes for evaluation was rated the highest of all items in this section (Item 43 - 4.1). The amount of time spent on the evaluation process was also rated high by principals (Item 40 - 4.0).

The availability of training programs and models of good practice for principals (Item 42 - 3.4) was in the moderate range. It should be noted that this question is not parallel with Item 42 in the TEP. The clarity of district policy statements regarding purposes for evaluation (Item - 43 - 4.1) was relatively high.

Principals reported that teacher growth was emphasized more than teacher accountability as the intended role for evaluation (Item 44 - 3.6). Time allotted during the teaching day for professional development was rated the lowest by principals (Item 41 - 3.3).

TABLE XVII
RESPONSE OF PRINCIPALS TO THE ITEMS
IN SECTION E OF THE TEP

Item	N	Mean	S. D.
40. Amount of time spent on the evaluation process including your time and that of all other participants	14	4.000	0.679
41. Time allotted during the teaching day for professional development	14	3.286	0.914
42. Availability of training programs and models of good practice for principals	14	3.429	0.938
43. Clarity of district policy statement regarding purposes for evaluation	14	4.143	1.027
44. Intended role for evaluation	14	3.643	0.745

Principals and ITIP

The value of having a common vocabulary in ITIP for use in discussions with teachers in evaluation conferences was rated the highest by principals (Item 47 - 4.6). Principals also reported that they encouraged the use of ITIP procedures (Item 45 - 4.4) and that they evaluated the ITIP skills of the teachers (Item 46 - 4.2) as shown on Table XVIII.

TABLE XVIII
RESPONSE OF PRINCIPALS TO THE
ITEMS IN SECTION F OF THE TEP

Item	N	Mean	S. D.
45. The extent to which I encourage the use of ITIP procedures	14	4.429	0.514
46. The extent to which I evaluate the ITIP skills of teachers	14	4.214	0.802
47. The value of having a common vocabulary in ITIP (e.g. guided practice) for discussions with teachers in evaluation conferences	14	4.571	0.938
48. My level of commitment to ITIP	14	4.357	0.745
49. The value I place on ITIP for improving instruction	14	4.357	0.745

Criterion Questions

Principals rated the criterion questions on Quality, Impact, and Value on a ten-point scale. As Table XIX shows, the mean score for Value was the highest at 7.1 and the mean score for Quality was the next highest at 6.8. However, they rated the Impact of the evaluation the lowest at 5.8.

TABLE XIX
RESPONSE OF PRINCIPALS TO THE CRITERION QUESTIONS ON QUALITY,
IMPACT, AND VALUE OF TEACHER EVALUATION

Criterion Question	N	Mean	S. D.
Regard the entire evaluation process, including planning for evaluation, classroom observations and feedback. How would you rate the overall quality of the teacher evaluations you did this year?	14	6.786	1.626
Rate the overall impact of the evaluation experience on the professional practices of the teachers you evaluated this year.	14	5.786	1.616
Rate the value you place on teacher evaluation as a source of growth for teachers	14	7.143	2.248

Scale Characteristics of Sections A-F

Scale scores were calculated for each of the six scales within the principals' survey. In addition, these responses were analyzed to determine the dependability of the items and scale in the questionnaire. Table XX shows the coefficient alphas for each of the scale scores.

The mean, variance, and scale reliabilities for sections A -F are also shown on Table XX. The scale reliabilities ranged from 0.78 in Section D to 0.28 in Section E. Items in Section E tend to be situation specific to what is happening in the district in contrast to items in other sections that are more personal perceptions. The variance ranged from 0.45 in Section C to 0.08 in Section B. The means ranged from 4.4 in Section F to 3.5 in Section C.

TABLE XX
MEANS, STANDARD DEVIATIONS, AND RELIABILITY COEFFICIENTS
FOR THE PRINCIPAL SAMPLE ON THE SECTIONS OF THE TEP

Section	Mean	S. D.	Coefficient Alpha
A. Teacher	3.738	0.185	0.738
B. Evaluator	4.006	0.089	0.757
C. Procedures	3.462	0.453	0.691
D. Feedback	3.587	0.140	0.777
E. Context	3.700	0.134	0.281
F. ITIP	4.386	0.017	0.622

ANALYSIS OF THE RESEARCH QUESTIONS

This section presents the results of the study relative to the specific research questions listed in Chapter III. Included in this section are reports which describe the following: an analysis of the relationship between the teachers response to the items on the six sections of the TEP and the criterion questions, an analysis of the relationship of ITIP training and the criterion questions, the response of tenured and probationary teachers to the TEP, and the response of principals to the TEP.

Characteristics of Teacher Evaluation

1. Is the value that teachers place on teacher evaluation related to the characteristics of the evaluation?

Evidence of the value that teachers place on teacher evaluation can be obtained by examining the teachers' responses to the three criterion questions that ask them to rate the quality, impact, and value of the evaluation experience. Table X in this study describes the responses of teachers to the criterion questions as rated on a ten-point scale. The responses of teachers in the study indicate that they rate the quality and the value of the evaluation higher than the impact. The means for Quality and Value are 7.1 and 7.0 respectively; however, the mean score for Impact is only 5.4.

The analyses relating to this research question involved the relationship between teachers' responses to TEP items and their rating of the criterion items of quality, impact, and value. This item/criterion relationship

was explored by examining the correlations between TEP items within each of the six sections of the TEP and the criterion ratings for quality, impact and value.

Section A of the TEP describes the attributes of the teacher. As shown in Table XXI, the experience of teachers with teacher evaluation prior to their most recent evaluation (Item 9) appears to influence their perceptions of the quality, impact, and value more than any other factor in this section.

One of the assumptions underlying the development of a summated rating scale is that the scale is made up of items that are of an approximately equal attitude value. There is a problem within Section A in that not all items are of an approximately equal attitude value. For example, Item 8 asks teachers to describe their years of teaching experience. This item is not equal with the other attitude items within Scale A. Therefore, it might be argued that Section A does not form a psychological construct that can be used as a summated scale.

TABLE XXI
CORRELATION FOR TEACHERS' RESPONSES ON TEP ITEMS IN SECTION A
AND CRITERION RATINGS FOR QUALITY, IMPACT, AND VALUE

Item	Quality	Impact	Value
1. Rate the strength of your professional expectations of yourself.	.0332	.0388	.0380
2. Orientation to risk taking	-.0509	.0307	-.0096
3. Orientation to change	.0047	.1529*	.1368*
4. Orientation to experimentation in the classroom.	.0412	.0399	.0890
5. Openness to criticism	.2048**	.1892**	.2256**
6. Knowledge of technical aspects of teaching	.0167	-.0205	-.0190
7. Knowledge of subject matter	.0149	.0256	.0352
8. Years of teaching experience	-.0115	.0709	-.1172
9. Experience with teacher evaluation prior to most recent experience.	.2999***	.3085***	.4684***

* $p < .05$

** $p < .01$

*** $p < .001$

Table XXII presents the correlations between the TEP items in Section B and the criterion items. Credibility (Item 10) was most highly correlated with their perceptions of the quality of the evaluation. The usefulness of suggestions for improvement was also correlated with their criterion ranking of quality (Item 20).

Persuasiveness of the rationale for suggestions was most highly correlated with the teachers' perceptions of the impact of the evaluation (Item 21). The usefulness of suggestions for improvements was showed the next highest correlation with impact (Item 20).

Of the three criterion items, value was less strongly correlated with items in Section B than were the questions dealing with quality and impact.

TABLE XXII
CORRELATION FOR TEACHER RESPONSES ON TEP ITEMS IN SECTION B
AND CRITERION RATINGS FOR QUALITY, IMPACT, AND VALUE

Item	Quality	Impact	Value
10. Credibility as a source of feedback	.5228***	.3889***	.2411***
11. Working relationship with you	.3449***	.2660***	.1293*
12. Level of trust	.3307***	.2642***	.0665
13. Interpersonal manner	.3241***	.2228**	.0926
14. Temperament	.2041**	.1526*	.1110*
15. Flexibility	.1871**	.1927**	.1202*
16. Knowledge of technical aspects	.2552***	.2251***	.1927*
17. Capacity to demonstrate or model	.4108***	.3953***	.2549***
18. Familiarity with your classroom	.3959***	.3782***	.2014**
19. Experience with classrooms	.2193**	.2681***	.1332*
20. Usefulness of suggestions	.4424***	.4163***	.3263***
21. Persuasiveness of rationale	.4293***	.4403***	.2676***

* $p < .05$

** $p < .01$

*** $p < .001$

Items in Section C deal with the attributes of the procedures used during the evaluation. Table XXIII reports the correlations between these items and the teacher ratings of quality, impact, and value. The teachers' endorsement of the standards as appropriate for their classrooms was the item most highly correlated with the criterion of quality (Item 24). This finding provides new information not found in the review of the literature on the research on the TEP. The clarity of the standards was the next item most highly correlated with quality (Item 23).

The clarity of the standards (Item 23) was most highly correlated with the teachers' perceptions of the impact of the evaluation. The items in Section C were not as highly correlated with the criterion question of value as they were with the questions of quality and impact.

In this section of the TEP the teachers' perceptions of quality, impact, and value are most closely correlated with issues related to their endorsement of the performance criteria defined in the standards and the clarity of the performance standards communicated to them in the evaluation process. The number of formal and informal observations by the principal was not highly correlated with the teachers' perceptions of quality, impact and values.

TABLE XXIII
CORRELATION FOR TEACHER RESPONSES ON TEP ITEMS IN SECTION C
AND CRITERION RATINGS FOR QUALITY, IMPACT, AND VALUE

Item	Quality	Impact	Value
22. Were standards communicated	.4544***	.4509***	.3736***
23. Were standards clear to you	.4939***	.4619***	.3459**
24. Were standards endorsed by you as appropriate	.5035***	.4291***	.3709***
25. Were the standards the same for all teachers	.2131**	.2221**	.1771
26. Observation of your classroom performance	.2868***	.1548*	.1620*
27. Examination of classroom or school records	.2407***	.2952***	.1416**
28. Examination of student achievement	.2241**	.2701***	.2103**
29. Number of formal observation per year	-.0071	.1666**	.1885**
30. Approximate number of informal observations	.1613*	.1835**	.0999

* $p < .05$

** $p < .01$

*** $p < .001$

The attributes of the feedback received by teachers in the evaluation conference are examined in Section D of the TEP. The three items in this section that were most highly correlated with the criterion ratings of quality and impact were Items 31, 34, and 35, as shown in Table XXIV.

The depth of the information provided in the conference was the most highly correlated with the teachers' perceptions of the quality of the evaluation (Item 34). The amount of information received in the conference was also strongly related to the criterion of quality (Item 31).

The depth of information was most highly correlated with the impact of the evaluation experience (Item 34). Item 35, on quality of the ideas and suggestions contained in the feedback, was also strongly correlated with the teachers' perceptions of the impact of the evaluation.

There was a weaker correlation between items in Section D of the TEP and the criterion question of value. The frequency of informal feedback and the depth of information had the the highest correlation with the criterion of value (Items 33 and 34).

The depth of information, amount of information, and quality of the ideas in the feedback were most highly correlated with the criterion items of quality and impact (Items 31, 34, and 35).

TABLE XXIV
CORRELATION FOR TEACHER RESPONSES ON TEP ITEMS IN SECTION D
AND CRITERION RATING FOR QUALITY, IMPACT, AND VALUE

Item	Quality	Impact	Value
31. Amount of information received	.5221***	.3995***	.2786***
32. Frequency of formal feedback	.3179***	.3660***	.2995***
33. Frequency of informal feedback	.3515***	.3538***	.3720***
34. Depth of information provided	.5370***	.5084***	.3277**
35. Quality of the ideas and suggestions contained in the feedback	.4908***	.4144***	.2730***
36. Specificity of information provided	.4235***	.3580***	.2675***
37. Nature of information provided	.3692***	.2846***	.1175**
38. Timing of the feedback	.4104***	.3170***	.3229***
39. Feedback focused on district teaching standards	.4170***	.2365***	.1556*

* $p < .05$

** $p < .01$

*** $p < .001$

Five items in Section E examine the attributes of the evaluation context. Two of these, items 40 and 44, had a moderate correlation with the criterion ratings in of quality, impact, and value, as shown in Table XXV. However, the items describing the context of the evaluation are not as strongly correlated with the criterion rating as are the items describing the attributes of the evaluator and the feedback.

TABLE XXV
CORRELATION FOR TEACHER RESPONSES ON TEP ITEMS IN SECTION E
AND CRITERION RATING FOR QUALITY, IMPACT, AND VALUE

Item	Quality	Impact	Value
40. Amount of time spent on the evaluation process including your time and that of all other participants	.2627***	.3404***	.2488***
41. Time allotted during the teaching day for professional development	.1542*	.3114**	.3123***
42. Availability of training programs and models of good practice	.1930**	.1668*	.1709**
43. Clarity of policy statements regarding purpose of evaluation	.2563***	.2296***	.1922**
44. Intended role of evaluation	.2913***	.2410***	.2548***

*p < .05

**p < .01

*** p < .001

Table XXVI reports correlations involving the teachers' experiences with the ITIP model of instruction. The item in this section that was most highly correlated with all three of the criterion ratings dealt with the value that teachers placed on ITIP for improving instruction (Item 49). However, none of the items in this section were as strongly correlated with the teachers' perception of quality, impact, and value as were the items that deal with the evaluator, the procedures, the feedback, and context.

TABLE XXVI
CORRELATION FOR TEACHER RESPONSES ON TEP ITEMS IN SECTION F
AND CRITERION RATINGS FOR QUALITY, IMPACT, AND VALUE

Item	Quality	Impact	Value
45. The extent to which I follow ITIP procedures in my class	.2498***	.1935**	.1192*
46. The extent to which my ITIP skills are evaluated by my principal	.2326***	.1911**	.1663*
47. The value of having a common vocabulary in ITIP(e.g. guided practice) for discussions in my evaluation conferences	.2646***	.1473*	.2554***
48. My level of commitment to ITIP	.1864**	.2614***	.2558***
49. The value I place on ITIP	.2170**	.2658***	.2694***

* $p < .05$

** $p < .01$

*** $p < .001$

The results of this analysis indicate that there is a relationship between the value that teachers place on teacher evaluation and the characteristics of the evaluation. There appears to be a strong relationship between the teachers' perception of the quality, impact, and value of the evaluation and certain key attributes of the evaluation process. The prior experience of the teacher with teacher evaluation and the skills of the evaluator appear to be two of the critical factors which influence teacher perceptions of the evaluation process. The quality of the feedback given to the teacher in the evaluation conference also appears to influence the teacher's perception of the quality, impact and value.

ITIP Training and Teacher Attitudes Toward Evaluation

2. Does the amount of training that teachers receive in ITIP affect the value they place on teacher evaluation?

The relationship between the value that teachers place on the evaluation process and the amount of ITIP training was investigated first by examining the teacher ratings on the criterion questions and the number of ITIP courses taken by teachers. ANOVA was used to examine this relationship and the results of the analysis are displayed on Table XXVII. No significant differences were found in teacher ratings of the criterion questions and the amount of training taken in ITIP courses.

TABLE XXVII
ANALYSIS OF VARIANCE TABLE FOR TEACHER RATING OF
CRITERION QUESTION AND NUMBER OF ITIP COURSES

Variable	Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of F
Quality	Between	2.399	3	0.800	0.217	.884
	Within	794.560	216	3.679		
	Total	796.959	219	3.639		
Impact	Between	3.061	3	1.020	0.178	.911
	Within	1257.988	219	5.744		
	Total	1261.049	222	5.680		
Value	Between	3.816	3	1.272	0.270	.874
	Within	1033.412	219	4.719		
	Total	1037.229	222	4.672		

The second phase of the analysis examined the relationship between the number of ITIP courses taken by teachers and their ratings of the scale scores. Item scores were then examined for significant scales. The ANOVA technique was used in this phase of the investigation. Significant differences at the .01 level were found for Scale A and at the .05 for Scale F, as shown on Table XXVIII. No other significant differences were found.

TABLE XXVIII
ANALYSIS OF VARIANCE TABLE FOR
SCALE SCORES AND ITIP COURSES

Variable	Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of F
Teacher Scale	Between	4.353	3	1.451	9.410	.001
	Within	33.926	220	0.154		
	Total	38.280	223	0.172		
Evaluator Scale	Between	0.192	3	0.064	0.130	.942
	Within	108.310	220	0.492		
	Total	108.501	223	0.482		
Procedures Scale	Between	1.658	3	0.553	1.392	.246
	Within	86.951	219	0.397		
	Total	88.610	222	0.399		
Feedback Scale	Between	0.373	3	0.124	0.201	.896
	Within	136.351	220	0.620		
	Total	136.724	223	0.613		
Context Scale	Between	0.895	3	0.298	0.615	.606
	Within	106.768	220	0.485		
	Total	107.663	223	0.483		
ITIP Scale	Between	4.168	3	1.389	3.202	.024
	Within	95.466	220	0.434		
	Total	99.634	223	0.447		

Table XXIX shows the results of the ANOVA comparing item scores for the Teacher Scale (Part A) and the number of ITIP courses taken by teachers. Teachers who had taken more ITIP course work tended to rate themselves higher in several important areas.

There was a significant difference at the $p < .05$ level for Item three which asks teachers to describe their orientation to change. In addition, teachers who had taken a greater number of ITIP courses tended to rate themselves higher on their openness to criticism (Item 5 - $p < .05$).

There was a significant difference at the $p < .001$ level between teachers who had taken more work in ITIP and the other teachers. This item asks teachers to rate themselves on their knowledge of technical aspects of teaching. This response should not be surprising because the ITIP training program is designed strengthen teachers' knowledge and skills in both theory and technical expertise. Interestingly, teachers with more training in ITIP also rated themselves much higher on their knowledge of subject matter (Item 7).

As noted earlier in this study, Item eight does not contribute to the unity of Scale A which is designed to describe the attributes of the teacher. The level of significance for this item may indicate that teachers with more experience in teaching have taken more courses in ITIP.

TABLE XXIX
ANALYSIS OF VARIANCE TABLE FOR SCALE A AND
NUMBER OF ITIP COURSES TAKEN BY TEACHERS

Variable	Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of F
Item 1	Between	0.862	3	.287	1.121	.341
	Within	56.419	220	.256		
	Total	57.419	223	.257		
Item 2	Between	3.823	3	1.274	1.860	.137
	Within	150.731	220	0.685		
	Total	154.554	223	0.693		
Item 3	Between	4.699	3	1.566	2.800	.041
	Within	123.047	220	0.559		
	Total	127.746	223	0.573		
Item 4	Between	2.871	3	0.957	1.777	.153
	Within	118.517	220	0.539		
	Total	121.388	223	0.544		
Item 5	Between	7.047	3	2.349	2.820	.040
	Within	178.256	214	0.833		
	Total	183.303	217	0.854		
Item 6	Between	15.239	3	5.080	10.399	.001
	Within	104.540	214	0.489		
	Total	119.780	217	0.552		
Item 7	Between	6.375	3	2.125	5.612	.001
	Within	81.024	214	0.379		
	Total	87.399	217	0.403		
Item 8	Between	38.635	3	12.878	9.054	.001
	Within	304.397	214	1.422		
	Total	343.032	217	1.581		
Item 9	Between	0.898	3	0.299	0.287	.835
	Within	223.437	214	1.044		
	Total	224.335	217	1.034		

Table XXX shows a clear trend in the response of teachers to the items in the Teacher Scale. The trend shows that the more classes teachers take the higher their mean score on the significant items. This trend continues even after teachers have taken more than three classes.

TABLE XXX
SIGNIFICANT ITEM MEANS AND STANDARD DEVIATIONS
ON SCALE A FOR TEACHERS GROUPED BY ITIP COURSES

Item	Classes Taken	N	Mean	S. D.
3. Orientation to change	No Classes	9	4.111	0.600
	One Class	57	4.315	0.827
	Two Classes	84	4.202	0.803
	Three Classes	48	4.250	0.636
	More Than Three	43	4.628	0.598
5. Openness to criticism	No Classes	9	3.556	0.726
	One Class	57	3.824	0.889
	Two Classes	84	3.738	0.933
	Three Classes	48	3.724	0.962
	More Than Three	34	4.265	0.790
6. Knowledge of technical aspects of teaching	No Classes	9	3.444	0.882
	One Class	57	3.772	0.732
	Two Classes	84	4.036	0.735
	Three Classes	48	4.333	0.630
	More Than Three	5	4.514	0.612
7. Knowledge of subject matter	No Classes	9	4.222	0.667
	One Class	57	4.158	0.676
	Two Classes	83	4.349	0.652
	Three Classes	46	4.522	0.549
	More Than Three	34	4.677	0.479

Table XXXI shows the ANOVA for number of ITIP courses taken and item scores for Scale F. Teachers who had taken more ITIP classes tended to have a greater commitment to the program (Item 48). There was also a significant difference between teachers in the extent to which they used ITIP procedures in their classroom (Item 45). Teachers with more course work in ITIP rated the value of a common vocabulary higher (Item 47).

TABLE XXXI
ANALYSIS OF VARIANCE TABLE FOR SCALE F AND
NUMBER OF ITIP COURSES TAKEN BY TEACHERS

Variable	Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of F
Item 45	Between	5.679	3	1.893	4.906	.003
	Within	89.520	230	0.386		
	Total	95.199	233	0.405		
Item 46	Between	4.512	3	1.384	1.537	.206
	Within	208.844	230	0.900		
	Total	212.996	233	0.906		
Item 47	Between	11.965	3	3.988	3.615	.014
	Within	255.968	230	1.103		
	Total	267.932	233	1.140		
Item 48	Between	5.607	3	1.869	2.694	.047
	Within	160.935	230	0.694		
	Total	166.542	233	0.709		
Item 49	Between	4.949	3	1.650	2.028	.111
	Within	188.712	232	0.813		
	Total	193.661	235	0.824		

Table XXXII shows the same trend that was found in Scale A. The more classes that teachers have taken in ITIP the higher their mean score on the items in Section F. The trend continues even after they have taken three or more classes. It might be predicted that there would be a ceiling effect after two or more classes but that assumption is not supported by the data.

Table XXXII
SIGNIFICANT ITEM MEANS AND STANDARD DEVIATIONS
ON SCALE A FOR TEACHERS GROUPED BY ITIP COURSES

Item	Classes Taken	N	Mean	S. D.
45. The extent to which I follow ITIP procedures in my class	No Classes	9	2.889	1.166
	One Class	57	3.719	0.559
	Two Classes	84	3.762	0.669
	Three Classes	48	3.792	0.582
	More Than Three	35	4.086	0.612
47. The value of having a common vocabulary in ITIP	No Classes	9	2.889	0.601
	One Class	57	3.842	1.192
	Two Classes	84	3.845	1.035
	Three Classes	47	4.043	0.954
	More Than Three	35	4.457	0.950
48. My level of committment to ITIP	No Classes	9	3.000	1.224
	One Class	57	3.649	0.767
	Two Classes	84	3.655	0.843
	Three Classes	48	3.854	0.825
	More Than Three	35	3.971	0.785

Probationary and Tenured Teachers

3. Do probationary and tenured teachers have different perceptions about the nature and value of teacher evaluation?

ANOVA technique was used in the analysis of this question by first examining the teacher responses to the criterion questions of quality, impact, and value. As Table XXXIII shows, no significant difference was found.

TABLE XXXIII
ANALYSIS OF VARIANCE TABLE FOR CRITERION QUESTIONS
FOR PROBATIONARY AND TENURED TEACHERS

Criterion Variable	Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of F
Quality	Between	0.447	1	0.447	0.124	.725
	Within	807.110	224	3.602		
	Total	807.557	225	3.589		
Impact	Between	0.011	1	0.011	0.002	.937
	Within	1285.900	227	5.664		
	Total	1285.911	228	5.639		
Value	Between	13.460	1	13.460	2.927	.088
	Within	1043.613	227	4.597		
	Total	1057.073	228	4.636		

In addition this question was explored by first examining the scale scores for the TEP and then examining the item scores for significant scales. Significant differences in perceptions of probationary and tenured teachers were found in Scales A, Teacher Attributes, and Scale E, Attributes of the Evaluation Context, as shown in Table XXXIV.

TABLE XXXIV
ANALYSIS OF VARIANCE TABLE FOR SCALE SCORES
FOR PROBATIONARY AND TENURED TEACHERS

Variable	Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of F
Teacher Scale	Between	2.112	1	2.112	12.866	.001
	Within	37.442	228	0.162		
	Total	39.555	229	0.172		
Evaluator Scale	Between	0.244	1	0.244	0.501	.479
	Within	110.941	228	0.486		
	Total	111.185	229	0.485		
Procedures Scale	Between	0.666	1	0.066	0.164	.685
	Within	91.648	227	0.404		
	Total	92.314	228	0.404		
Feedback Scale	Between	0.526	1	0.526	0.852	.356
	Within	140.808	228	0.617		
	Total	141.335	229	0.617		
Context Scale	Between	2.159	1	2.154	4.615	.032
	Within	106.487	228	0.467		
	Total	108.642	229	0.081		
ITIP Scale	Between	0.012	1	0.012	0.025	.873
	Within	108.308	228	0.475		
	Total	108.320	229	0.473		

Table XXXV shows that significant differences between probationary and tenured teachers were found on Items 5, 6, 7, and 9 of Section A, which describes the attributes which the teacher brings to the evaluation process. Significant difference between probationary and tenured teachers also was reported for Item 8. However, this item asks teachers to report their years of teaching experience.

Probationary teachers reported a greater degree of openness to criticism (Item 5). Probationary teachers also reported that their previous experience had been more helpful than did tenured teachers (Item 9).

Probationary teachers reported significantly lower scores for knowledge of subject matter and knowledge of technical aspects of teaching than did tenured teachers (Items 6 and 7).

No significant differences were found in the scores for probationary and tenured teachers for Items 1, 2, 3, and 4. Item 1 asks teacher to rate the strength of the professional expectations they set for themselves and Item 2 asks them to rate themselves on their orientation to risk taking. Item 3 asks teacher to rate themselves on their orientation to change. Although the difference in the scores between probationary and tenured teachers on Item 3 was not statistically significant, the probationary teachers mean score was 4.5 as compared to 4.2 for tenured teachers, it does appear that probationary teachers tend to be more open to change.

TABLE XXXV

ANOVA TABLE FOR SCALE A FOR PROBATIONARY AND TENURED TEACHERS

Variable	Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of F
Item 1	Between	0.051	1	0.051	0.189	.664
	Within	62.248	228	0.273		
	Total	62.300	229	0.272		
Item 2	Between	0.474	1	0.474	0.683	.409
	Within	158.117	228	0.693		
	Total	158.591	229	0.692		
Item 3	Between	2.042	1	2.042	3.643	.057
	Within	127.852	228	0.560		
	Total	129.895	229	0.567		
Item 4	Between	1.1049	1	1.104	2.0195	.156
	Within	124.742	228	0.547		
	Total	125.847	229	0.549		
Item 5	Between	7.075	1	7.075	8.732	.003
	Within	183.937	227	0.810		
	Total	191.013	228	0.837		
Item 6	Between	8.497	1	8.497	16.020	.000
	Within	120.933	228	0.530		
	Total	129.430	229	0.565		
Item 7	Between	8.277	1	8.277	22.763	.000
	Within	81.456	224	0.363		
	Total	89.734	225	0.398		
Item 8	Between	159.251	1	159.251	200.261	.000
	Within	181.309	228	0.795		
	Total	340.560	229	1.487		
Item 9	Between	7.609	1	7.609	7.701	.006
	Within	227.246	227	1.001		
	Total	234.855	228	1.030		

Table XXXVI describes the differences in mean scores and standard deviations for probationary and tenured teachers on Scale A. Approximately 26% of the teachers in the study were probationary teachers who had less than three years of teaching experience in the districts in the study.

TABLE XXXVI
SIGNIFICANT ITEM MEANS AND STANDARD DEVIATIONS ON
SCALE A FOR PROBATIONARY AND TENURED TEACHERS

Item	Tenure Status	N	Mean	Stand. Dev.
5. Openness to criticism	Probationary	48	4.479	0.652
	Tenured	182	4.247	0.772
6. Knowledge of technical aspects of teaching	Probationary	48	4.104	0.778
	Tenured	182	4.275	0.729
7. Knowledge of subject matter	Probationary	48	4.021	0.668
	Tenured	178	4.489	0.585
8. Years of teaching experience	Probationary	48	2.167	0.808
	Tenured	182	4.212	0.912
9. Experience with teacher evaluation prior to most recent experience	Probationary	48	3.813	0.982
	Tenured	181	3.365	1.005

In Section E, which describes the evaluation context, the forty-eight probationary teachers had a higher mean score on all items than did the tenured teachers. Probationary teachers in the study received additional training during the school day under a new teacher orientation program, which accounts for the difference in their response to Item 41. As Table XXXVII shows, no single item was significant.

TABLE XXXVII
ANALYSIS OF VARIANCE TABLE FOR SCALE E
FOR PROBATIONARY AND TENURED TEACHERS

Variable	Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of F
Item 40	Between	0.823	1	0.823	1.050	.306
	Within	178.572	228	0.783		
	Total	179.395	229	0.783		
Item 41	Between	4.497	1	4.497	3.520	.061
	Within	291.224	228	1.277		
	Total	295.721	229	1.291		
Item 42	Between	2.116	1	2.116	1.472	.226
	Within	326.215	227	1.437		
	Total	328.331	228	1.440		
Item 43	Between	1.559	1	1.559	1.285	.258
	Within	276.614	228	1.213		
	Total	278.173	229	1.215		
Item 44	Between	2.448	1	2.448	2.013	.157
	Within	277.242	228	1.216		
	Total	279.691	229	1.221		

Table XXXVIII shows the differences in mean scores and standard deviations for probationary and tenured teachers on Scale E. All items in Scale E were displayed because while no single item showed a significant difference, the trend was for probationary teachers to rate the items higher than tenured teachers.

TABLE XXXVIII
ITEM MEANS AND STANDARD DEVIATIONS ON SCALE E
FOR PROBATIONARY AND TENURED TEACHERS

Item	Tenure Status	N	Mean	Stand. Dev.
40. Amount of time spent on the evaluation process including your time and that of all other participants	Probationary Tenured	48 182	3.521 3.3736	0.899 0.881
41. Time allotted during the teaching day for professional development	Probationary Tenured	48 182	2.438 2.093	1.023 1.155
42. Availability of training programs and models of good practice	Probationary Tenured	48 181	3.750 3.514	1.082 1.228
43. Clarity of policy statements regarding purpose for evaluation	Probationary Tenured	48 182	3.8125 3.6099	0.915 1.145
44. Intended role of evaluation	Probationary Tenured	48 182	3.979 3.725	0.785 1.171

Teachers and Principals

4. Do principals and teachers have different perceptions about the nature and value of evaluation?

The first step in exploring this question was to examine the differences in responses to the criterion questions on quality, impact and value. ANOVA technique was used and, as shown on Table XXXIX, no significant differences were found in the responses of teachers and principals.

TABLE XXXIX
ANALYSIS OF VARIANCE TABLE FOR TEACHER AND PRINCIPAL
RATINGS OF THE CRITERION QUESTIONS

Criterion Variable	Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
Quality	Between	.488	1	0.488	1.383	.710
	Within	851.240	241	3.522		
	Total	851.689	245	3.476		
Impact	Between	1.542	1	1.542	0.283	.595
	Within	1327.629	244	5.441		
	Total	1328.171	245	4.6549		
Value	Between	.012	1	0.013	0.003	.959
	Within	1134.800	244	4.651		
	Total	1134.812	245	4.632		

Significant differences at the $p < .05$ level were found on the Teacher Scale. However, questions on the two questionnaires were not equivalent. Therefore, the item scores for this section are not reported. Significant differences were found for the Context Scale and the ITIP Scale, as reported on Table XXXX.

TABLE XXXX
ANALYSIS OF VARIANCE TABLE FOR SCALE SCORES
FOR PRINCIPALS AND TEACHERS

Variable	Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of F
Teacher Scale	Between	1.498	1	1.498	8.495	.003
	Within	43.211	245	0.176		
	Total	44.709	246	0.181		
Evaluator Scale	Between	0.138	1	0.138	0.29	.584
	Within	113.297	245	0.462		
	Total	113.436	246	0.461		
Procedures Scale	Between	0.609	1	0.609	1.560	.212
	Within	95.245	244	0.390		
	Total	95.854	245	0.391		
Feedback Scale	Between	0.339	1	0.339	0.570	.450
	Within	145.703	245	0.594		
	Total	146.043	246	0.593		
Context Scale	Between	1.913	1	1.913	4.175	.042
	Within	112.308	245	0.458		
	Total	114.222	246	0.464		
ITIP Scale	Between	3.924	1	3.924	8.593	.003
	Within	111.896	245	0.456		
	Total	115.821	246	0.471		

As Table XXXXI shows, principals indicated that a greater amount of time was spent on evaluation, (Item 40) and that more time was allotted for professional growth during the school day (Item 41). A follow-up interview with principals indicated they counted released time for workshops and committee as professional growth time for teachers. Item means for teachers and principals can be found on Table III and Table XII in this document.

TABLE XXXXI
ANALYSIS OF VARIANCE TABLE FOR SCALE E
FOR PRINCIPALS AND TEACHERS

Variable	Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of F
Item 40	Between	4.643	1	4.643	6.074	.014
	Within	185.749	243	0.764		
	Total	190.392	244	0.780		
Item 41	Between	16.339	1	16.339	12.81	.001
	Within	309.931	243	1.275		
	Total	326.269	244	1.337		
Item 42	Between	0.269	1	0.269	0.191	.662
	Within	342.001	243	1.407		
	Total	342.270	244	1.403		
Item 43	Between	3.048	1	3.048	2.525	.113
	Within	293.377	243	1.207		
	Total	296.425	244	1.215		
Item 44	Between	0.312	1	0.312	0.270	.604
	Within	280.652	243	1.155		
	Total	280.964	244	1.151		

In the ITIP Section the wording of Item 45 is not parallel on the principals' and teachers' questionnaires and, therefore, this difference was not examined. As Table XXXXII shows, the principals rated the value of a common vocabulary in ITIP (Item 47), the value of ITIP for improving instruction Item 49), and their level of commitment to ITIP (Item 48) significantly higher than did teachers. Item means for teachers and principals can be found on Table IX, and Table XVIII in this document.

TABLE XXXXII
ANALYSIS OF VARIANCE TABLE FOR SCALE F
FOR PRINCIPALS AND TEACHERS

Variable	Source of Variation	Sum of Squares	DF	Mean Square	F	Signif of F
Item 45	Between	5.715	1	5.715	12.827	.001
	Within	108.268	243	0.446		
	Total	113.984	244	0.467		
Item 46	Between	0.847	1	0.847	0.898	.344
	Within	229.006	243	0.942		
	Total	229.853	244	0.943		
Item 47	Between	5.273	1	5.273	4.633	.032
	Within	276.580	243	1.138		
	Total	281.853	244	1.155		
Item 48	Between	5.455	1	5.455	7.432	.007
	Within	178.357	243	0.734		
	Total	183.812	244	0.753		
Item 49	Between	4.021	1	4.021	4.756	.030
	Within	205.448	243	0.845		
	Total	209.469	244	0.858		

SUMMARY OF CHAPTER IV

Summary of Teachers' Responses to the TEP

The majority of the teachers in this study reported a great deal of experience with ITIP. About 60 % of them have used ITIP in their classroom for six years or more and about 70 % have taken two or more ITIP classes. Teachers in this study rated themselves higher than four points on a five-point scale in their strength of professional expectations , knowledge of subject matter, and orientation to change. The teachers in this study rated the knowledge, skills, attitudes of the principals relatively high. The level of trust between principals and teachers and the credibility of the principals, as reported by teachers, was above four on a five-point scale.

Teachers reported classroom observation was the main source of information used in the evaluation process and that the examination of classroom records and student achievement played only a minor role in the evaluation process. The teachers' perceptions were that the frequency of informal feedback from principals was relatively low. However, formal feedback was seen as focused on the district standards, and the nature of the feedback was viewed as helpful.

The amount of time available for professional growth during the teaching day received the lowest score of any item in this study. However, the clarity of policy statements regarding the purpose of evaluation and the intended role of evaluation received high marks. Teachers in this study reported that their ITIP skills are regularly evaluated by the principals and that they place a high value on the use of a common vocabulary in ITIP to describe instructional practices in the evaluation conferences.

Summary of Principals' Responses

The response of the principals in this study was much more positive than the current research literature on teacher evaluation would predict. In many areas principal responses tended to be very similar to those of the teachers. For example, principals rated the criterion question in the same order and with very similar scores. The criterion question on value received the highest rating followed by the quality impact questions. It is interesting to note that principals agree with the teachers in giving the impact criterion the lowest rating.

Principals rated themselves high on their ability to help teachers grow through evaluation. They also gave themselves high marks on their evaluation skills. There was consensus between teachers and principals on the point that classroom observation is the primary source of information on teacher performance. They also agreed that student achievement and classroom records were rarely used in the evaluation process. One of the consistent themes revealed in this study is that both teachers and principals agree that standards for evaluation are clearly communicated and accepted by teachers.

In examining questions related to feedback, principals in this study noted that they usually did not leave a written communication following an informal observation. The clarity of the district policy on teacher evaluation and the amount of time spent on the process were rated high by principals. The principals strongly supported the ITIP model and identified the use of a common vocabulary to describe instruction as the most important aspect of the ITIP model.

Summary of the Research Questions

1. Is the value that teachers place on teacher evaluation related to the characteristics of the evaluation?

There was a strong positive correlation between the teacher ratings of the criterion questions on quality, impact, and value and certain attributes of the evaluation. While correlation does not indicate causality, the strength of the correlations does indicate a relationship. The prior experience of the teacher with the evaluation process was one of the key attributes that influenced the teachers perceptions of quality, impact, and value. Teachers also reported that the skills of the evaluator and the quality of the feedback were critical attributes which influenced their perceptions of evaluation.

2. Does the amount of training that teacher receive in ITIP affect the value they place on teacher evaluation?

No significant differences were found in the teacher ratings of the criterion questions and the training levels of the teachers. Significant differences were found in the Teacher Attribute Scale and in the ITIP Experience Scale. In an analysis of the Teacher Attribute Scale it was found that the greater the number of ITIP courses taken by teachers the higher they rated themselves on openness to criticism, orientation to change, and knowledge of technical aspects of teaching. In examining items within the ITIP Experience Scale it was found that the greater the number of ITIP courses taken the higher the teachers rated themselves on their commitment and use of ITIP.

3. Do probationary and tenured teachers have different perceptions about the nature and value of teacher evaluation?

No significant difference was found in the response of probationary and tenured teachers to the criterion questions of quality, impact, and value.

However, a significant difference was found in scale scores for the Teacher Scale and the Context Scale. On the Teacher Scale it was found that probationary teachers reported their previous experience with teacher evaluation was more helpful than did tenured teachers and they also reported a greater degree of openness to criticism. On the Context Scale probationary teachers reported more time available during the teaching day for professional development.

4. Do teachers and principals have different perceptions about the nature and value of evaluation?

No significant differences were found in the responses of teachers and principals on the criterion questions of quality, impact, and value.

Significant differences were found in the items of the Context Scale and the ITIP Experience scale. On the Context Scale principals reported more time available for teachers for professional development during the school day than did teachers. Principals also reported that more time was spent on the evaluation process. On the ITIP Scale principals reported a greater commitment to the use of ITIP and the value they place on the process for improving instruction.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

INTRODUCTION

Chapter V contains summaries, conclusions, and recommendations derived from this study on the attitudes of elementary teachers and principals toward the teacher evaluation process. The study was conducted in two suburban school districts near Portland, Oregon. The districts were selected because they were recognized regionally and nationally for their programs of teacher evaluation. Both districts had several years of experience in the ITIP model.

The major purpose of the study was to investigate the general perceptions of teachers and principals regarding teacher evaluation in two districts that had a major commitment to implementing and maintaining the ITIP model for instruction and teacher evaluation. Four research questions were developed to examine specific aspects of the teacher evaluation process and to explore differences in attitudes between groups. The research analysis involved the investigation of correlational data and the use of the ANOVA technique to examine significant differences in group means. Coefficients Alpha were calculated to determine the dependability of the items in the research instrument.

This chapter contains the following sections: (1) Summary of the Study; (2) Conclusions; (3) Recommendations; and (4) Summary of Chapter V.

SUMMARY OF THE STUDY

Teacher evaluation is seen as central to the purposes of school reform and to the goal of creating effective schools. The ITIP model of instruction and teacher evaluation has been adopted statewide by Texas, South Carolina, and Pennsylvania as the means to achieve state-wide school improvement goals.

An examination of the research literature revealed that little information was available on the attitudes of teachers and principals toward the ITIP model. One of the fundamental assumptions that guided this study was that the attitudes of teachers and principals toward teacher evaluation is an accurate indicator of the effectiveness of the process. The organizational factors that support the evaluation process and encourage mutual trust, credibility, and shared understandings were of particular interest in the study.

The instrument used in this study was the Teacher Evaluation Profile (TEP), developed and by the Northwest Regional Educational Laboratory. This instrument was designed to measure the characteristics of the teacher, the evaluator, the evaluation procedures, the feedback, and the context of the evaluation process. The TEP measures these attributes against the teachers' perceptions of the quality and impact of their most recent evaluation experience. This instrument was modified for the purposes of this study to include an additional set of questions on ITIP and an additional criterion question on value. In addition, a parallel version of the TEP was developed for use with principals involved in this study.

The findings of the study indicated that teachers and principals were positive in their perceptions of the quality, impact, and value of the evaluation process. Both groups, however, rated the impact of the evaluation lower than the quality and the value. Teachers rated themselves high on the professional attributes that they bring to the evaluation process. They were also positive about their working relationship with the principals in the evaluation process and gave the principals a high rating on their evaluation skills. Classroom observation was the primary means of collecting information on teacher performance, and teachers tended to endorse the district standards for evaluation. Teachers reported that feedback from principals was specific and that the ideas and suggestions in the feedback were generally helpful. Teachers indicated that there was little time allotted during the teaching day for professional development. Principals and teachers supported ITIP and reported that there was value in having a common vocabulary in ITIP for discussions in the evaluation conference.

The results of the study showed a relationship between the value that teachers place on the process and certain attributes of the evaluation. It did not appear that the amount of training in ITIP influenced the value that teachers placed on the process. However, all but nine of the 233 teachers in the study had taken at least one ITIP class. Probationary teachers tended to find the evaluation process more valuable than did tenured teachers. Principals and teachers gave strong support to the ITIP model as a means of improving instruction, but principals were more supportive of the model. Overall, responses of the teachers and principals were similar.

CONCLUSIONS

Comparison With a National Study

In an attempt to add another dimension to this study, the researcher compared the findings of this investigation with the results of a national study on teacher evaluation. During 1987-88 the Northwest Regional Educational Laboratory used the TEP to collect data from an independent sample of over 4,500 teachers from 27 school districts from the states of Connecticut, Montana, Ohio, Oregon, and Washington.

A word of caution is appropriate at this point. Studies which are based on the collection of data from self-report questionnaires have limitations. Also it should be recognized that any comparison of this kind has limited value because of the differences in the populations and settings investigated in the two studies. For example, the national sample contains data from both elementary and secondary schools, whereas the current study was conducted only in elementary schools. The national study contains data from urban, suburban, and rural schools while the current study was done in suburban schools. The national sample contains schools which employ a variety of instructional and evaluation formats, but this study was conducted in schools which used ITIP over a period of several years. However, this kind of general comparison can be of value in stimulating ideas about the evaluation process, and in suggesting areas for further research. A report comparing the results of the two studies can be found in Appendix C.

An examination of the teacher ratings of the two criterion questions on quality and impact showed that teachers in both the local and the national studies rated the quality of the evaluation higher than the impact. About 70 %of the teachers in the present study rated quality as a seven or above on the ten-point scale whereas only about 50% of the teachers in the national study rated quality at seven or above. About 40% of the teachers in the present study rated the impact at seven or above as compared to about 20% of the national study teachers. A third criterion question on the value of the evaluation process as a source of professional growth, unique to the present study, had a mean score of about seven on a ten-point scale. A reasonable interpretation of the findings from the current study on the criterion questions is that the evaluation process was of relatively high quality and value, but had only a moderate impact upon the professional practices of the teachers.

Obviously, a critical factor in the teacher evaluation process is the set of attributes that the teacher brings to the process. In comparing item scores for the Teacher Section of the TEP, it was found that teachers in the present study rated themselves as high or higher in all categories than did teachers in the national study. Teachers in the present study also rated the knowledge, skills, and attitudes of the evaluator much higher than did teachers in the national study. For example, in ranking the evaluator's persuasiveness and the capacity to model needed improvements, about 65% of the teachers in the current study gave the evaluator scores of four or above on a five-point scale as compared to about 35% of the teachers in the national study.

In describing the number of formal observations, 99% of the teachers in the current study reported receiving at least one or more formal observation per year as opposed to 83% for the national study. About 53% of the teachers in this study and 31% of the teachers in the national study reported receiving at least one informal observation each month.

In comparison with the national study, teachers in this study reported consistently higher scores on all items in the feedback section. The clarity of policies regarding evaluation, the intended role of evaluation, and the availability of training programs and models of good practice were rated much higher by the teachers in the current study. Section F, which examined teachers' experience with ITIP, was unique to this study. Teachers in the current study indicated a strong level of support for ITIP with the mean score for all questions approaching four point on a five-point scale.

Overall, the teachers in the present study were consistently higher in their ratings of the evaluation experience than the teachers in the national study. Furthermore, the response of teachers in the present study indicates that there is a relationship between specific attributes of the evaluation and the teachers' perceptions of the quality and impact of the evaluation. Clearly, the teachers' perceptions of the knowledge, skills, and attitudes of the evaluator are critical factors. The existence of clearly defined standards for evaluation and consensus by teachers and principals on the major purpose of evaluation are also important factors.

Conclusions From the Current Study

There is a clear trend in the responses of teachers and principals in this study to the three criterion questions on quality, impact, and value. Both groups rate the quality and value of the evaluation higher than the impact. They report that the evaluation process is a quality experience that has value, but has only a moderate impact on changing professional practices. National commissions, state departments of education, and local school boards have all identified teacher evaluation as the primary means of changing instructional practices. This belief that teacher evaluation should serve as the key instrument of school reform is an excellent example of Wise's description of hyperrationalization. This emphasis on teacher evaluation originated because the first wave of school reform in the early 1980s was based on a top-down philosophy of reform. Teacher evaluation was viewed by the reform movement as an instrument for improving instructional practice. However, the task of creating profound changes in teachers' attitudes, understandings, and teaching practices is a multifaceted issue; and to place the major emphasis on teacher evaluation as the primary instrument to bring about instructional change greatly underestimates the complexity of the issue. The more exciting finding is that teachers and principals in this study, in contrast to the findings of other studies, reported that the quality and value of the evaluation experience was generally high. This suggests that the information gained from this study on the procedures and contexts of the evaluation has important implications for educational practice.

The teachers in this study saw themselves as highly competent professionals who were open to change and growth. It is critical that

teachers bring to the evaluation experience not only a sense of competence, but also an openness to learning new information. This finding is also an important reflection of the impact of the district staff development programs upon the self-confidence of the teachers. It is significant that the teachers in both districts participated actively in the design of the evaluation procedures.

Clearly, the principal is the most important factor in the success of an evaluation program. Teachers in this study rated the knowledge, skills, and attitudes of the principals exceptionally high. Both districts in this study made major commitments to training all administrators in ITIP by first sending them to the UCLA Laboratory School for training with Madeline Hunter and then providing periodic inservice training. In addition the districts paid consultants, including Hunter, to come to the district and provide joint teacher/administrator training. Central office administrators and other principals sat in on teacher evaluation conferences and coached the principals in the development of their conference skills. The reports of the teachers in rating their evaluators show a high level of trust and mutual respect between teachers and principals in this study. These conditions are the enabling factors that are required for an effective program of teacher evaluation.

The findings indicate that principals in this study were well trained in the technical skills of clinical supervision and ITIP. However, the information collected through this process was limited to observational data on teaching performance. While observational data represent an important source of information, it is too limited a source. Classroom records, lesson plans, and student achievement data were not widely used for purposes of

teacher evaluation by the school districts in this study. While it would be inappropriate to use standardized test data for the purpose of evaluating teachers, a review of student papers, homework, quizzes, and other information could provide valuable insights which would increase teacher effectiveness. This broadened scope of evaluation should take place in an atmosphere of mutual trust and support. Adequate training should be provided for teachers and principals to use these important sources of information.

Teachers in the study reported that fewer informal observations occurred than did principals. In addition, data from the study indicated that principals rarely left notes for teachers following an informal observation. It would be interesting to see if these two factors are related. If the principal left a brief note after an informal observation teachers would receive some immediate feedback on their teaching.

Tenured teachers in this study did not have time during the teaching day for professional development. There are at least two ways that this problem could be solved, but both would require additional funding. The length of the teaching day could be extended, or the number of days in the teaching contract could be extended.

The commitment of these districts in this study to teacher growth through evaluation and inservice programs is clearly understood by both teachers and principals. The level of training of both groups is very high. In the case of the teachers, the results show that commitment to ITIP increased as more classes were taken and continued to increase even after three classes were completed. Perhaps ITIP is not the right instructional model for all schools, but it is clear from the results of this study that the long term

support of a single instructional model, coordinated with the use of teacher evaluation strategies which are congruent with that model, resulted in a climate of shared values, mutual trust, and common goals.

One of the advantages of a narrowed focus on teaching is the high level of agreement on a clearly articulated definition of good teaching practice. The use of a common vocabulary to describe instruction is also helpful for both teachers and principals and results in clear and specific feedback in evaluation conferences. The availability of resources to support professional development is central to a successful evaluation program.

There was broad support of the ITIP model by teachers and principals in the study. Principals ranked the program significantly higher than teachers, but the mean score for teachers, 3.8 on a five-point scale, shows a high level of support. If a program is to be successful in the classroom it is critical that the building principal strongly endorse it. It is also critical that there is a high level of teacher support for the program. The findings of this study show that both conditions exist in these districts.

The findings of this study support the premise that the value that teachers place on the evaluation process is related to certain characteristics of the evaluation. The level of trust between teachers and principals, agreement on the standards for evaluation, and the quality of the feedback were strongly correlated with the teachers' perceptions of quality, impact, and value of the evaluation. These findings suggest that teacher evaluation is both an organizational issue and a technical issue. While it is important for principals to have technical skills in evaluation, it is equally important for them to be able to establish a climate of trust and shared understanding.

The amount of training that teachers received in ITIP did positively influence their perceptions of their orientation to change, their knowledge of the technical aspects of teaching, and their commitment to ITIP. However, there was not much evidence to support the notion that amount of training was the key factor because all but nine of the 233 teachers in the study had taken at least one class in ITIP. The finding that teachers' commitment to the program continued to increase even after they had taken more than three classes suggests that the inservice program has a strong influence on teachers' attitudes.

There should be a difference in the ways that probationary and tenured teachers are evaluated. Data collected in this study did not indicate that there were substantive differences in the evaluation procedures for probationary and tenured teachers. However, the researcher believes that tenured teachers should be encouraged to move toward a program of self-evaluation or peer evaluation and should be formally evaluated only once every two years. This would allow principals to spend more time working with probationary teachers. In addition, an induction program for probationary teachers should be a part of every staff development program.

Information collected in this study indicated that the differences in the perceptions of teachers and principals about the nature and value of evaluation were minor. The differences that were found were predictable in that principals do spend more time on evaluation and are usually somewhat more committed to the district instructional format than teachers. The responses of teacher and principals to the criterion questions of quality, impact, and value were almost congruent. Their agreement on the critical points included: an endorsement of the district standards for evaluation,

agreement on the clarity of district policy on teacher evaluation, and agreement of the intended role of evaluation. The effectiveness of any teacher evaluation system depends on the shared values and goals of the teachers and principals who must implement the system. Consensus of this sort is essential for a successful teacher evaluation program. Legislation, national mandates, and board policy will have little effect upon evaluation unless teachers and principals agree upon the critical aspects of teacher evaluation.

As Hunter's critics have indicated, the adoption and implementation of a single instructional format, as in the case of ITIP, can also create barriers to growth. Growth will not continue if the teachers and principals see this model as the only acceptable way to teach or if they fail to expand their repertoire of teaching strategies to include new models. A follow-up interview with the staff development directors of the districts in this study indicated that they were using ITIP as a basis for expanding the teaching skills of the teachers. One district was working to expand the use of a variety of teaching strategies and the other district was working on Cooperative Team Learning. Both directors said that they believed that ITIP training helped teachers to move more easily into learning new teaching strategies.

It is unfortunate that there is a lack of college and university participation in the design, implementation, and evaluation of these staff development and evaluation programs. Tapping the talents of the college and university community could add an important dimension to the process of teacher evaluation.

The positive attitudes of the teachers and principals toward teacher evaluation may be related to several factors not directly investigated in this study. The researcher believes that a key factor is that both districts in the study used the ITIP model, as Hunter has advocated, as a decision-making model for teachers. In contrast to this approach, many districts have used the model as a lock-step design for instruction. Studies of teacher and principal attitudes toward teacher evaluation in districts that have used that approach have tended to be negative; (Garman 1987; Mandeville and Rivers 1989). The districts in this study have used the decision making aspects of the ITIP model to expand the skills of teachers and principals to include other teaching strategies such as Cooperative Team Learning.

A second factor is that well-trained principals and staff development teams are available to coach, model good practices, and encourage teachers to expand their skills and to take the risks inherent in trying new teaching strategies. Both districts have made a longterm financial commitment to maintaining and expanding the skills of teachers and principals. Hunter (1989) has indicated that approximately two years of coaching are necessary for teachers to achieve a level of mastery in ITIP. The districts in this study have used the ITIP model for over 12 years, and the model has become an important part of the culture of the organizations.

RECOMMENDATIONS

Recommendations for Improving Practice

The following recommendations describe actions that the researcher believes can be implemented to improve the effectiveness of teacher evaluation programs.

1. Teachers and administrators must work together to design, implement, and review programs of evaluation which are consistent with the instructional format and the fundamental values and beliefs of the school district.
2. Teachers and principals ought to be jointly trained in both the instructional strategies used in the district and in the evaluation process used to evaluate those strategies.
3. Teachers and principals also ought to be trained in data collection techniques for evaluation that go beyond just classroom observation of teaching and include data on student achievement.
4. School districts could form partnerships with colleges and universities and draw upon the expertise of the partners to improve staff development programs, teacher evaluation, field based research, and college programs in teacher and administrator preparation.
5. Administrators should be trained in both the technical aspects of evaluation and in the organizational strategies which promote a collaborative culture within the school.
6. School districts must make long-term commitments to teacher evaluation and staff development. Whether the instructional format

is ITIP, Cooperative Learning or some other model, the district should maintain staff training support in that model. When staff members have mastered a model, they can expand their repertoire of teaching strategies to include other models.

7. Master teachers ought to be trained in strategies of self-evaluation and peer review and then released from the once-a-year formal evaluation cycle.

8. Induction programs must be designed to support teachers entering the school system, and time should be set aside during the school day for additional training.

9. The superintendent must give active support and consistent attention to implementing and sustaining a successful teacher evaluation program.

10. Master teachers and master principals should serve as mentors to probationary staff members.

Recommendations for Further Research

1. Other studies are needed that give sharply focused attention to the issue of adult development within school systems. Research on the developmental characteristics of teachers at various stages in their careers can guide the organization of learning experiences for school staff.

2. Case studies of successful programs in principal training should be conducted and replicated.

3. There is a need to conduct studies of successful staff development programs within different types of school districts in order to build effective models for replication.
4. Studies should be conducted to examine the effectiveness of various formats for providing time during the school day for professional growth. For example, studies of cost-effective staff development programs that are held during the teaching day would be helpful.
5. The current study should be replicated in similar school districts.

SUMMARY OF CHAPTER V

It has been said that knowledge is never complete, and this is particularly true in the case of teacher evaluation. The results of this study have contributed to our knowledge of teacher evaluation in two school districts, but there is a need for continued study of the process in a variety of settings. The results of this study have illustrated the value of integrating teacher evaluation into the major instructional programs of the school district.

Perhaps the greatest contribution this study can make to knowledge about teacher evaluation is to reaffirm the concept that teacher evaluation is of greatest value when it is conducted in a climate of trust, support, and open communication.

REFERENCES

- Acheson, K., & Gall, M. (1980). Techniques in the clinical supervision of teachers: preservice and inservice applications. New York: Longman.
- Acheson, K., & Gall, M. (1987). Techniques in the clinical supervision of teachers: preservice and inservice applications. New York: Longman.
- Argyris, C. (1964) Integrating the individual and the organization. New York: John Wiley.
- Argyris, C. & Schon, D. (1976). Theories of action. In W. Bennis and R. Berne (Eds.) The planning of change. New York: Holt, Rinehart and Winston.
- Barber, L. (1985). Improving teacher performance: Formative evaluation. Bloomington, IN: Phi Delta Kappa.
- Barber, L. (1987). Teacher evaluation: A report submitted to the teacher quality program. Bloomington IN: Indiana Department of Education.
- Barr, S., Burton, W., & Brueckner, L., (1955) Supervision. New York: Appleton-Century-Crofts.
- Blumberg, A., & Amidon, E., (1965). Teacher perceptions of supervisor teacher interaction. Administrator Notebook, 14, (1) 165- 170.
- Blumberg, A., (1974) Supervisors and teachers: A private cold war. Berkeley, CA.: McCutchan.
- Brandt, R., (1987) On teacher evaluation; A conversation with Tom McGreal. Educational Leadership, 44, (2) 20.
- Brandt, R., (1988) On research and school organization: A conversation with Bob Slavin, Educational Leadership, 46 (2) 22.

- Brookover, W. & Lezotte, L., (1982) Creating effective schools: Learning climate and achievement. Holmes Beach, FL. Learning Publications.
- Bridges, E., (1984) Managing the incompetent teacher. Stanford, CA. Institute for Research on Educational Finance and Governance, School of Education, Stanford University.
- Brophy, J., & Evertson, C., (1975) Learning from teaching: A developmental perspective. Boston: Allyn and Bacon.
- Burns, J., (1978) Leadership. New York: Harper and Row.
- Callahan, R., (1962) Education and the cult of efficiency. Chicago: University of Chicago Press.
- Cogan, M., (1973). Clinical supervision. Boston: Houghton Mifflin.
- Cogan M., (1976) Rationale for clinical supervision, Journal of Research and Development in Education, 9(2).
- Costa, A., (1984) A reaction to Hunter's knowing, teaching and supervising. in P. Hosford (Ed.) Using what we know about teaching. Alexandria, VA. Association for Supervision and Curriculum Development.
- Darling-Hammond, L., Wise, A., & Pease, S. (1983) Teacher evaluation in the organizational context: A review of the literature. Review of Educational Research, 53(3) 285-328.
- Davis, H., (1964) Evolution of current practice in evaluating teacher competency. In B. Biddle and W. Ellena, (Eds.). Contemporary research on teacher effectiveness. New York: Holt, Rinehart and Winston.
- Drucker, P. (1988) The coming of the new organization, Harvard Business Review, 88(1) 45-53.
- Edmonds, R., (1979) Some schools work and more can, Social Policy, 9(5) 28-32.

- Edmonds, R. (1981) A report on the research project: Search for effective schools and certain of the designs for school improvement that are associated with the project. East Lansing, Mich: The Institute for Research on Teaching, College of Education, Michigan State University, Unpublished report prepared for N.I. E., 1981.
- Emmer, E., Evertson, C. & Anderson, L., (1981) Effective classroom management at the beginning of the school year, Elementary School Journal, 80, 219-231.
- Evertson, C. et al., (1979) Organizing and managing the junior high classroom. Austin, TX. R and D Center for Teacher Education, University of Texas.
- Evertson, C., Anderson, C., & Anderson, L. (1980) Relationships between classroom behavior and student outcomes in junior high math and English classes. American Elementary Research Journal, 17, 43-60.
- Fayol, H., (1916) Administration Industrielle, Et Generale. Translated into English in 1929 by Constance Storrs as General and Industrial Management. London: Pitman & Sons.
- Flanders, N., (1967) Introduction to interaction analysis: in Amidon, E. and Hough, J. (Eds.) Theory, application and research. Reading MA. Addison-Wesley.
- Flanders N., (1970) Analyzing teaching behavior. Reading, MA.
- Garman, N., & Hazi, H., (1988) Teachers ask: Is there life after Madeline Hunter? Phi Delta Kappan, May.
- Gibboney, R., (1987) A critique of Madeline Hunter's teaching model from Dewey's perspective. Educational Leadership, 44, 46-50.
- Glickman, C., (1981) Developmental supervision. Alexandria, VA: Association for Supervision and Curriculum Development.
- Goldhammer, R., (1969) Clinical supervision: Special methods for the supervision of teachers. New York: Holt, Rinehart and Winston.

- Goldhammer, R., Anderson R., & Krajewski R., (1980) Clinical Supervision: Special methods for the supervision of teachers. Second Edition, New York: Holt, Rinehart and Winston.
- Goldsberry, L., (1984) Implication of the survey of supervisory practice findings. Paper delivered at the meeting of the American Educational Research Association, San Francisco.
- Good, T., and Grouws, D., (1979) The Missouri mathematics effectiveness project: An experimental study in fourth grade classrooms. Journal of Educational Psychology, 71, 355-362.
- Good, T., & Brophy, J., Looking in classrooms. (3rd. ed.) New York: Harper & Row.
- Hales, L., and Pratton J., (1986) The effects of active participation on student learning, Journal of Educational Research, 79 (4) 210-215.
- Hunter, M., (1966) When the teacher diagnoses learning. Educational Leadership, 23 (7) 545-549.
- Hunter, M. (1973) Appraising teaching performance: One approach. National Elementary Principal, 52 (5), 60-62.
- Hunter, M. (1979) Teaching is decision making. Educational Leadership, 37 (1) 62-64, 67.
- Hunter, M., (1980) Six types of supervisory conferences. Educational Leadership 37 (5) 408-410, 412.
- Hunter, M., (1984) Knowing, teaching, and supervising. In P. Hosford (Ed.) Using what we know about teaching. Alexandria, VA: Association for Supervision and Curriculum Development. 169-192.
- Hunter, M. (1985) What's wrong with Madeline Hunter? Educational Leadership, 42 (5), 57-60.
- Hunter, M., (1988) Well acquainted is not enough: A response to Mandeville and Rivers, Educational Leadership, 46 (4) 67-68.

- Kirst, M., (1983) State education policy in an era of transition. Institute for Research on Educational Finance and Governance, School of Education, Stanford University, Policy Paper No. 83-C7.
- Kirst, M., (1988) Who should control our schools? NEA Today, 74-79.
- Lewin, K. (1958) Group decisions and social change. In E. Maccoby, T. Newcomb, & E. Hartley, (Eds) Readings in social psychology. New York: Holt Reinhart and Winston.
- Likert, R., (1977) New patterns of management. New York: Mc Graw-Hill.
- Manatt, R., Palmer, K. & Hidlebaugh, E., (1976) Evaluating teacher performance with improved rating scales. NASSP Bulletin 60 (401) 21-23.
- Mandeville, G., & Rivers, J., (1988) Effects of South Carolina's PEP program. Educational Leadership 46 (4) 63-66.
- Mandeville, G. & Rivers, J., (1989) Is the Hunter model a recipe for supervision?. Educational Leadership. 46 (8) 39-43.
- Maslow, A., (1962) Toward a psychology of being. Princeton: D. Van Nostrand Co.
- Mayo, E., (1945) The Social Problems of an Industrial Civilization. Boston: Harvard University School of Business.
- McCarty, D., Kaufman, J. & Stafford, J. (1986) Supervision and evaluation; two irreconcilable processes? The Clearinghouse. 59 351-353.
- McGreal, T. (1983) Successful teacher evaluation. Alexandria, VA.; Association for Supervision and Curriculum Development.
- McGreal, T., (1986) Developing a teacher evaluation system: Commonalties of those systems that function effectively. Unpublished paper. University of Illinois, Campaign.
- McGregor, D. (1960) The Human Side of Enterprise. New York: McGraw-Hill.

- McIntyre, J., & Morris, W. (1984) Time management and instructional supervision. In Mcnaughton, R., Tracy, S. and Rogus, J. NASSP Bulletin. 68 475 1.
- McLaughlin, M. (1984) Teacher evaluation and school improvement. Teachers College Record. 8 (1), 193-205.
- McLaughlin, M., Wise, A., Darling-Hammond, L., and Bernstein, H., (1984) Teacher Evaluation: A study of effect practices. Santa Monica, CA. : Rand Corporation.
- McLaughlin, M., & Pfeifer, S. (1986) Teacher evaluation: Learning for improvement and accountability - case studies. Stanford Education Policy Institute, Stanford University.
- Millman, J. (1981) Handbook of teacher evaluation. Beverly Hills, CA: Sage Publications.
- Mosher, R., & Purpel, D. (1972) Supervision: The reluctant profession. New York: Houghton Mifflin.
- National Commission on Excellence in Education (1983). A nation at risk: The imperative for educational reform. Washington, U. S. Department of Education.
- Popham, J., (1988) Judgement based teacher evaluation. In Stanley S, & Popham J. (Eds) Teacher evaluation: Six prescriptions for success. Alexandria, VA. Association for Supervision and Curriculum Development.
- Retallick, J. (1986) Clinical supervision: Technical, collaborative and critical approaches. In Smyth, J., (Ed) Learning about teaching through clinical supervision. Sydney: Crom Helm. 85-110.
- Ritz, W. & Cashell, J. (1980) Cold war between supervisors and teachers? Educational Leadership. 38, (1) 77-78.
- Robbins, P., (1986) The Napa-Vacaville follow-through project: Evaluative outcomes, related procedures, and implications for practice. Elementary School Journal. 87, 139-157.

- Robbins, P., & Wolfe, P., (1987) Reflections on a Hunter-based staff development project, Educational Leadership, 44 (5) 56-61.
- Rosenshine, B., 1982) Teaching functions and instructional programs. Paper presented at NIE conference on implications of research on teaching for practice. Airlie House, VA:
- Rosenshine, B., (1983) Teaching functions in instructional programs. Elementary School Journal 83.
- Rosenshine, B., (1985) Synthesis of research on explicit teaching. Educational Leadership, 45 (8) 60-69.
- Ruck, C. (1984) Cold war revisited: The use of performance goals in the teacher evaluation process in Oregon, unpublished doctoral dissertation, University of Oregon.
- Sahling, N., (1981) An observational study of clinical supervision in selected elementary, junior, and senior high schools. Unpublished doctoral dissertation, Columbia University Teachers' College, DAI 41 (12) 4965
- Shinn, J. (1976) Teacher perceptions of ideal and actual supervisory procedures used by California elementary principals: The effects of supervisory programs sponsored by the Association of California School Administrators. Unpublished doctoral dissertation.
- Slavin R., (1987) The Hunterization of America's schools. Instructor, 96 56-59.
- Stallings, J., & Krassavage, E. (1986) Effects of instruction based on the Madeline Hunter model on student achievement; Findings from a Follow Through Project. Elementary School Journal 86, 5 571-588.
- Stiggins, R. & Duke, D., (1986) Teacher evaluation: Five keys to growth. Washington, National Education Association.
- Stiggins, R. & Duke, D. (1988) The case for commitment to teacher growth: Research on teacher evaluation. Albany: State University of New York.

- Stiggins, R. & Nickel, P. (1989) The Teacher Evaluation Profile: A technical analysis, Journal of Personnel Evaluation in Education, 1 (5).
- Taylor, F., (1912) The principles of scientific management. New York: Harpers.
- Tyack, D. (1975) The one best system. Cambridge, MA. Harvard University Press.
- Tyack, D. & Hansot, E. (1982) Managers of Virtue. New York: Basic Books.
- Weber, M. (1947) The theory of social and economic organization. Edited by Talcott Parsons and translated by Henderson, P. and Parsons, T. New York: Free Press of Glencoe.
- Wiles, K. (1955) Supervision for better schools. 2nd ed. Englewood Cliffs, N. J. Prentice-Hall.
- Wise, A. (1979) Legislated learning: The bureaucratization of the American classroom. Berkeley, CA.: University of California Press.
- Wise, A., Darling-Hammond, L., McLaughlin, M. & Bernsein, H. (1984) Teacher evaluation: A study of effective practices. Santa Monica, CA. Rand Corporation.
- Wolfe, P. (1984) Implementation of the Hunter instructional model: A staff development study paper presented at the National Staff Development Conference, Willimasburg, Virginia, Dec. 1984.
- Wolfe, P. (1987) What the seven-step lesson plan isn't. Educational Leadership, 44 Feb. 70-71.
- Zonca, P. (1972) A case study exploring the effects on an intern teacher of the conditions of openness in a clinical supervisory relationship. Unpublished doctoral dissertation, University of Pittsburg, DAI 33 658 A.

APPENDIX A

QUESTIONNAIRE: TEACHER EVALUATION PROFILE

PLEASE NOTE

**Copyrighted materials in this document have
not been filmed at the request of the author.
They are available for consultation, however,
in the author's university library.**

**144-150
152-160
162-166**

University Microfilms International

APPENDIX B

QUESTIONNAIRE: PRINCIPAL EVALUATION PROFILE

APPENDIX C

REPORT: COMPARISON WITH NATIONAL TEP STUDY