Teacher performance evaluation: a nationwide status report of type, content and duration of training for public school teachers

Joseph Michael Petrone

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Educational Administration and Supervision Commons

Recommended Citation


https://lib.dr.iastate.edu/rtd/9406
INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600
Teacher performance evaluation: A nationwide status report of type, content and duration of training for public school teachers

Petrone, Joseph Michael, Ph.D.
Iowa State University, 1990
Teacher performance evaluation: A nationwide status report of type, content and duration of training for public school teachers

by

Joseph Michael Petrone

A Dissertation Submitted to the Graduate Faculty in Partial Fulfillment of the Requirements for the Degree of DOCTOR OF PHILOSOPHY

Department: Professional Studies in Education
Major: Education (Educational Administration)

Approved:
Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa
1990
# TABLE OF CONTENTS

## CHAPTER I. INTRODUCTION
- Teacher Accountability and Growth 1
- Training for Teacher Evaluation 1
- Training: Key Factor in Quality Evaluation 3
- Training for Teachers Not Just Administrators 4
- Training for Teacher Evaluation: A Clear Staff Development Mandate 5
- Problem Statement 7
- Purpose 9
- Objectives 9
- Hypotheses 10
- Basic Assumptions 10
- Scope of the Investigation 11
- Delimitations 11

## CHAPTER II. REVIEW OF THE LITERATURE
- Introduction 13
- Background of the Problem 15
- History of Teacher Performance Evaluation 17
- External Factors 21
- Internal Factors 22
- Introduction of McGreal's Taxonomy of Teacher Evaluation Models 27
- McGreal's Taxonomy of Teacher Evaluation Models 29
CHAPTER IV. FINDINGS

Introduction

Questionnaire Return Rate

State return rate
Geographic region return rate
District size return rate
District locale return rate

Findings

Findings reported by weighted frequencies and percentages
Research questions one through four
The first research question

Awareness training by evaluation model
Modeling training by evaluation model
Practice training by evaluation model
Feedback training by evaluation model

The second research question

Effective teaching training content by evaluation model
Interpersonal communications training content by evaluation model
Observational data analysis skills training content by evaluation model
Overall evaluation training content by product and artistic models
Common law model by duration of teacher evaluation training
Goal-setting model by duration of teacher evaluation training
Clinical supervision model by duration of teacher evaluation training
Artistic and product models by duration of teacher evaluation training

The third research question

Awareness type training by district size
Modeling type training by district size
Practice type training by district size
Feedback type training by district size 88
Effective teaching training content by district size 89
Interpersonal communications training content by district size 90
Observational data analysis skills training content by district size 91
Duration of teacher evaluation training by district size 93

The fourth research question 93

Common law evaluation model by district size 93
Goal-setting evaluation model by district size 96
Clinical supervision evaluation model by district size 96
Artistic and product evaluation models by district size 96

Research questions five and six 97

Teacher evaluation training is independent of teacher evaluation models 98
Training for teacher evaluation is independent of district size 100

CHAPTER V. SUMMARY, CONCLUSIONS, LIMITATIONS, DISCUSSION AND RECOMMENDATIONS 102

Summary—Overview of the Study 102
Conclusions—Overview of the Results 103

The first research question 104

Awareness training described by evaluation models 104
Modeling training described by evaluation models 104
Practice training described by evaluation models 104
Feedback training described by evaluation model 105

The second research question 105

Effective teaching content described by evaluation model 105
Interpersonal communications content described by evaluation model 105
Observation and analysis skills content described by evaluation model 106
Duration of training described by evaluation model 106

The third research question

Effective teaching content described by district size 107
Interpersonal communications content described by district size 107
Observation and analysis skills content described by district size 107
Awareness training described by district size 107
Modeling training described by district size 108
Practice training described by district size 108
Feedback training described by district size 108
Duration of training described by district size 108

The fourth research question

Evaluation models described by district size 108

The fifth research question

Training independent of models 109

The sixth research question

Training independent of district size 110

Limitations of the Study 110
Discussion of the Study 112
Recommendations for Practice 115
Recommendations for Further Research 117
Suggestions for Replication and Refinement 118

BIBLIOGRAPHY 120
ACKNOWLEDGMENTS 133
APPENDIX A. QUESTIONNAIRE 134
APPENDIX B. HUMAN SUBJECTS APPROVAL FORM 139
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Type of public school district categorized by size</td>
<td>60</td>
</tr>
<tr>
<td>Table 2</td>
<td>Sample districts' expansion weights listed by size and United States geographic areas</td>
<td>62</td>
</tr>
<tr>
<td>Table 3</td>
<td>Questionnaire return rate categorized by state and district size</td>
<td>70</td>
</tr>
<tr>
<td>Table 4</td>
<td>Questionnaire return rate by United States geographic area</td>
<td>72</td>
</tr>
<tr>
<td>Table 5</td>
<td>District response rate by district size category</td>
<td>73</td>
</tr>
<tr>
<td>Table 6</td>
<td>Return rate categorized by locale of districts</td>
<td>74</td>
</tr>
<tr>
<td>Table 7</td>
<td>Types of teacher evaluation training by evaluation models (first research question)</td>
<td>78</td>
</tr>
<tr>
<td>Table 8</td>
<td>Effective teaching training content by evaluation models (second research question)</td>
<td>81</td>
</tr>
<tr>
<td>Table 9</td>
<td>Interpersonal communications training content by evaluation models (second research question)</td>
<td>82</td>
</tr>
<tr>
<td>Table 10</td>
<td>Observational data analysis skills training content by evaluation models (second research question)</td>
<td>83</td>
</tr>
<tr>
<td>Table 11</td>
<td>Duration of teacher evaluation training by evaluation models (second research question)</td>
<td>85</td>
</tr>
<tr>
<td>Table 12</td>
<td>Types of teacher evaluation training by district size (third research question)</td>
<td>87</td>
</tr>
<tr>
<td>Table 13</td>
<td>Effective teaching training content by district size (third research question)</td>
<td>89</td>
</tr>
<tr>
<td>Table 14</td>
<td>Interpersonal communications training content by district size (third research question)</td>
<td>91</td>
</tr>
<tr>
<td>Table 15</td>
<td>Observational data analysis training content by district size (third research question)</td>
<td>92</td>
</tr>
<tr>
<td>Table 16</td>
<td>Duration of teacher evaluation training by district size (third research question)</td>
<td>94</td>
</tr>
</tbody>
</table>
Table 17. Teacher evaluation models by district size (fourth research question) 95

Table 18. The fifth research question: Is teacher performance evaluation training provided to teachers independent of teacher evaluation models 99

Table 19. The sixth research question: Is teacher performance evaluation training independent of district size 101
CHAPTER I. INTRODUCTION

Teacher Accountability and Growth

The most recent educational reform pendulum swing has produced a national impetus for accountability and growth. The National Commission on Excellence in Education in 1981 and their report, A Nation at Risk (1983), started this steady movement and subsequent reports have maintained a cadence call for national school improvement. This decade's educational reform movement has lost little strength in the years since the "first" report and has firmly established itself as a national and state priority. So persistent are the demands to improve the outcomes of schooling that "the public has come to believe that the key to educational improvement lies in upgrading the quality of teachers rather than in changing school structure or curriculum" (Darling-Hammond, Wise, & Pease, 1983). For, it is believed, through quality evaluations public school districts can and must address both issues of accountability and growth. This study investigates training for teachers to help them meet the standards of performance used in their evaluation process, an issue central to the development of quality teacher behaviors.

Training for Teacher Evaluation

The key variables in this national investigation are teacher performance evaluation and training; therefore, this review of literature has focused on these two constructs. For the purpose of this study, definitions are as follows: (1) training - any staff development activity that attempts to help teachers improve instruction (Sparks,
1983) and designed to advance the knowledge, skills, and understanding of teachers in ways that lead to changes in their thinking and classroom behavior (Fenstermacher & Berliner, 1983); and (2) teacher performance evaluation - a means of improving teaching performance (Bolton, 1973). This definition of performance evaluation differs from pro forma, perfunctory summary rating of a teacher intended to meet legal obligations. Teacher performance evaluation, for the purpose of this research, is a planned, cooperative assessment of a teacher's performance made through the use of multiple classroom observations, collections of pertinent artifacts, and resultant goal attainment, all predicated on mutually determined teacher performance criteria. The climate in which the appraisal is conducted is collegial and communicative, characterized by a shared growth orientation and a positive trusting evaluator-evaluatee relationship.

Sources of information for this first stage review of literature include, but have not been limited to: ERIC CD ROM Service, Educational Administration Abstracts, Psychological Abstracts, and Dissertation Abstracts.

Quality teacher performance evaluation requires the linkage of staff development to credible, evaluative feedback which will increase the likelihood that teachers will act on the results for the purpose of improvement (Pfeifer, 1986). This point is clearly articulated; however, highly developed and current teacher evaluation systems across the country generally lack integration between teacher evaluation and staff development (McLaughlin, 1982). Also, Weber (1987), who more recently
completed a synthesis of the literature on teacher evaluation, noted that evaluation requires as much clarity about objectives and methods as teaching itself does, and fully as many interpersonal skills.

**Training: Key Factor in Quality Evaluation**

The intensive national scrutiny and study currently being performed and used in the quest for better schools has resulted in a call for more accountability and a focus on growth. However, there is such suspicion and sentiment that these supervisory objectives may not be reached jointly with the broader goal of school improvement (Tracy & McNaughton, 1989). Researchers like Darling-Hammond et al. (1983) explain there are four minimal conditions for successful teacher evaluation:

- All actors in the system have a shared understanding of the criteria and processes for teacher evaluation;
- All actors understand how these criteria and processes relate to the dominant symbols of the organization; that is, there is 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;
- 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;
- All actors in the system perceive that the evaluation procedures allow them to strike a balance . . . that is, that the procedure achieves a balance between control and autonomy for the various actors in the system. (p. 320)

Implicit in each of these success traits is the central focus of this investigation - training for teachers in their own evaluation processes. A closer look exposes phrases such as: "shared understanding of criteria and processes," "understand how these processes and criteria relate," "evaluation procedure enables and motivates them to improve
their performance," and "procedure achieves a balance between control and autonomy." Each phrase confirms the need to train all "actors" in the processes of evaluation. Stiggins and Bridgeford (1984) further support this claim as they state:

All evaluators and staff must be thoroughly trained. Everyone involved in the evaluation should know how to use the evaluation instruments to acquire useful, objective data, interpret results, and use those results to advantage. (p. 33)

Again, Stiggins along with Duke (1988) "stress that both supervisors and teachers should receive evaluation training" (p. 137). Furthermore, these researchers state:

We believe it is insufficient to do as many states and districts have done and require training only for supervisors. Such a course of action perpetuates the idea that evaluation is something done to teachers. If teacher evaluation is to contribute to growth, teachers as well as supervisors must be well versed in the process. Both must understand the components of good teaching, master the skills of interpersonal communication, and know how to make sense of data collected on teaching. Such knowledge increases the likelihood that teachers eventually will become the agents of their own professional development, rather than remaining dependent on others. (p. 137)

Training for Teachers Not Just Administrators

Weber (1987), in his recent synthesis of the literature on teacher evaluation, emphasizes the importance of the confidence, trust and specificity in the evaluation process. His research review clearly places a significant burden on administrators to become better trained coaches for teacher improvement rather than detectives of deficiencies. McGreal (1983) foreshadows this later call by Weber as he explains that "with the exception of additional time spent with supervisors on the
responsibilities in the goal-setting conference, on observation
techniques, and on conferencing and feedback skills, administrators and
teachers should receive the same training" (p. 144). Clark, Lotto and
Astuto (1984) might call McGreal's approach "mandated participation" because such directive staff development activities or top-down,
authoritarian training seems to oppose popular beliefs about lasting
change strategies. However, this same researcher and, later, Glickman
(1987, 1988) explain the mixture of highly participatory (bottom-up) with
authoritarian (top-down) staff development movements are generally the
most effective and lasting.

Practitioners and researchers agree that the ideal evaluation system
first teaches teachers about the evaluation process (Pfeiffer, 1986).
"Teacher performance evaluation (TPE) is a skill (or a series of skills)
and like skiing, tennis, or winning at Atari, TPE can be enhanced by
training" (Manatt, 1982, p. 1).

Training for Teacher Evaluation: A Clear
Staff Development Mandate

Other research reports by Stiggins and Bridgeford (1985) have
related the perceptions of teaching professionals with regard to the
formulation of a teacher performance evaluation system, and they
conclude: "Teachers want at the very least, an evaluation system that
provides accurate information on classroom needs, opportunity to acquire
and master new learning approaches, and collegial support when
instigating needed changes" (p. 92). Rothberg (1984) implies
accountability needs are being met by teacher evaluation processes;
however, growth and improvement are infrequent outcomes. This is suggested as he explains that current evaluation models provide for specific performance assessment and only marginally assist individuals in analyzing their classroom performance. Pfeifer (1986) supports this claim as he explains that if training is utilized to enhance skill areas which usually impede effective evaluation, then the performance appraisal process moves beyond the traditional notions of accountability to a loftier status of improved classroom performance. The inclusion of enhanced training practices in the evaluation process which suggest efficiency of staff development linkage to evaluation outcomes must be, at least described, if not confirmed, at the national level.

Specifically, the goal of this investigation is to provide a clearer picture of just what type, content, and duration of training provisions are used in school districts nation-wide. It appears that researchers agree the potential exists to meet both growth and accountability needs through quality teacher evaluation; however, at this time "educators generally concur that even highly developed evaluation systems are not helping teachers either individually or collectively to improve their skills" (Stiggins & Bridgeford, 1985, p. 87). In order for teacher evaluation "to assure successful and lasting instructional enhancement . . . core training should be provided to all teachers and administration in a district" (McGreal, 1988, p. 4). The results of this study provide a heretofore nonexistent status of national training offerings provided to teachers in their own evaluation systems.
Problem Statement

The research base demonstrates the call for training provisions for teachers in support of their evaluation systems. McGreal (1983), a strong proponent for training of all "actors" in the evaluation system, categorizes predominant characteristics of evaluation systems into five models: Common Law, Goal-Setting, Product, Clinical Supervision and Artistic or Naturalistic. The themes of accountability and growth appear consistently throughout these models. The issues of accountability and growth, believed by some to be competing or almost mutually exclusive goals of evaluation, have diverted attention from a common, mutually inclusive theme, that of training.

Training expectations from growth (neo-progressives) advocates differ from those of accountability (neo-traditionalists) proponents (Tracy & McNaughton, 1989). Neo-traditionalists training participants often acquire skills through several intense workshops and sometimes falsely expect that training transfers such skills to the level of application (Tracy & McNaughton, 1989), whereas Garman (1986b), representing the neo-progressive point of view, believes one needs extensive and practical, supervised training before a supervisor approximates expert status. Regardless which side one takes, certainty exists on the importance of training. Administrators and teachers need the skills necessary for the clear understanding of organizational expectations of teacher behavior and, most importantly, improved classroom instruction resulting in increased student learning, an outcome in which all model proponents believe. This implicit, yet untested,
agreement among the five models of evaluation is the problem and focus of this investigation.

Little research exists describing the type, duration, and content of teacher training for improved instruction and the dependence of these factors on teacher performance expectations listed as criteria in teacher performance evaluation processes. This study, more specifically, serves to establish whether a commonality of training or need for training among the five models of teacher evaluation exists in national public school practice. This study partially solves the problem of the shallow research base in the area of training teachers in the performance expectations listed as criteria in their evaluation processes. As a solution to the problem, this investigation describes the status of content, type, and duration of training which exists for teachers in teacher evaluation processes, thus identifying the heretofore implicit training demand made by both accountability and growth evaluation advocates.

Staff development for teacher evaluation need not be an "either-or" proposition but, rather, an effective combination of accountability and growth-oriented training. That is to say, this study attempts to substantiate a commonality which appears to exist among the five widely accepted models of teacher evaluation, with training being used as the common ground.

The following questions more specifically address the problem:

a. What are the types of teacher evaluation training among public school districts using specified evaluation models?
b. What are the content and duration of teacher evaluation training utilized in public schools categorized by models of teacher evaluation?

c. What types, content and duration of evaluation training are most frequently used by public school districts of various sizes?

d. What evaluation models are utilized most frequently by public school districts of various sizes?

Purpose

The purposes of this study are:

a. To determine the type, duration, and content of training provided to teachers in their evaluation systems.

b. To determine whether teacher evaluation training was either dependent or independent of evaluation models and district size?

Objectives

The overarching goal of this study is to develop a survey which poses questions ultimately serving to collect information on training type, duration, and content and to provide information on which level of the McGreal's Taxonomy of teacher evaluation model is used by public schools included in the population. Specific objectives are:

a. To collect and categorize data from national public school districts in order to identify the type, estimate the duration, and report the content of training offered teachers in their evaluation processes.

b. To report the utilization of the five McGreal models of teacher
evaluation.
c. To select and use appropriate statistical tests for each hypothesis.
d. To report findings and suggest further research questions as a result of conclusions from the data collected.

Hypotheses

a. Teacher performance evaluation training provided to teachers is independent of teacher evaluation models.
b. Teacher performance evaluation training is independent of district size.

Basic Assumptions

a. Persons completing individual questionnaires are knowledgeable in the district's teacher evaluation system.
b. Districts responding to this survey are representative of a sample which will be used to estimate the prescribed national population.
c. Respondents to the questionnaire are expected to provide complete and accurate information.
d. Orientation level training, at a minimum, is essential when attempting to train teachers in their own teacher evaluation system.
e. Districts are categorized as using one of the five models listed in the McGreal Taxonomy of teacher evaluation and employing a certain type, duration, and content of teacher evaluation
training.


g. The generalization of the findings is restricted to the research (in teacher evaluation models, content, type, and duration of training) from which the individual items on the questionnaire have been based.

Scope of the Investigation

Delimitations

a. This sample provides for separate questionnaire mailings made to 700 public school districts.

b. The questionnaire respondents for this study will be superintendents or superintendents' designees.

c. This study is limited to a stratified, nonproportional and randomly selected school district sample representing the prescribed national population.

d. The population from which the random sample is made includes those districts with 20 or more teachers.

e. Districts within the nine (9) states which have mandated appraisal procedures, criteria, and subsequent teacher evaluation training requirements have been excluded from this study.

f. The results of this investigation can represent school year 1989/90, the year in which the data were collected.
g. The district size (number of teachers) is based on the data gathered in the Common Core of Data (CCD) surveys collected March 1988 by the National Center for Educational Statistics (NCES).
CHAPTER II. REVIEW OF THE LITERATURE

Introduction

The school restructuring movement of the 1990s has its root in "equal access to knowledge" and "public demonstration of results" (Glickman, 1990, p. 40), with teacher evaluation playing a starring role on the stage of accountability. However, many questions and confusing rhetoric have dimmed this role justifiably assigned to teacher evaluation. Darling-Hammond et al. (1983) report that political pressures on schools, climate and cultural considerations all contribute to a school's readiness and willingness to design an effective, reliable and legally-discriminating teacher performance evaluation system. They explain that "successful teacher evaluation requires consistent and shared views of the teaching-learning process and of the organizational context in which teacher evaluation takes place" (p. 317).

Similar "enabling conditions" were described by McLaughlin and Pfeifer (1986) when they linked an effective evaluation to a process dependent upon a district's organizational environment. These researchers believe the success of a teacher performance evaluation system hinges on the level to which a district's organizational environment displays:

- trust between teachers and administrators,
- open communications throughout the district,
- commitment to organizational and professional improvement,
- visibility for evaluation. (p. 49)

More recently, practitioners such as Gainey (1990) clearly support these research conclusions as he cites "mutual respect, trust, and shared
responsibility" as factors associated with successful teacher evaluation (p. 17).

Agreement exists between teachers and administrators with regard to the importance of shared responsibility, open communication and commitment to learning; it is essential to strengthen these teacher performance evaluation success factors. Considerably more attention must be given to these conditions through training for teacher performance evaluation. Such training represents a complex, multifaceted and opportune adult learning forum. Training is a necessary feature of successful teacher performance evaluation, but is seldom provided to the degree necessary to transfer learning and promote professional development.

Training of professional educators in their evaluation process and the subsequent linkage to teacher performance evaluation has been suggested by many researchers (McGreal, 1983; Stiggins & Duke, 1988; Pfeifer, 1986; Manatt, 1982). On face value, this connection is inherently and inextricably linked to standards of professional behavior or those observable standards listed as performance criteria established by most school districts and used as a benchmark to judge performance. Equally important, however, teacher performance evaluation has the potential to establish meaningful professional development directions and commitments for teachers which could serve as catalysts fueling improvement in classroom instruction and, consequently, student achievement.

Assuming this training to be truly essential, it is certainly
appropriate and, most importantly, professionally essential to provide a brief history of the external and internal factors associated with teacher performance evaluation. Also, the most widely accepted research-based models of this complex process is discussed. This literature review is also intended to sketch the history of adult learning, explain effective components of educator training and define the training content most often associated with successful teacher performance evaluation.

Background of the Problem

In the past, according to Darling-Hammond (1990), public school K-12 teacher evaluation "has generally not been a high-stakes activity, in part because improving the quality of teachers has not been seen as critical for improving the quality of education..." (p. 17). Thus, teacher evaluation, where practiced, was often an exercise to which few resources and little organization and attention was devoted. As a consequence, teacher evaluation has often had little influence on decisions about personnel, staff development, or the structure of teaching. The current restructuring movement has placed added demands and attention on improving student learning, and teachers are seen as having the greatest impact on such student achievement outcomes.

A problem exists when a description of nation-wide status of training which is intended to support personnel decisions, professional development, and the actual structure of teaching is unavailable in the literature. While the literature is replete with case studies supporting teacher growth (Stiggins & Duke, 1988), it presents little evidence
suggesting practices employed by public schools to train teachers in the essential elements of quality evaluation: two-way communication, effective teacher techniques and observation skills (McLaughlin & Pfeifer, 1988). These elements are not only central to quality evaluation, but also, if properly applied, would serve to provide a staff development infrastructure sure to advance the external political demands of accountability and internal organizational goals of increased student learning. Furthermore, findings in the literature on staff development and teacher evaluation have failed to consistently reflect the common purposes of teacher accountability and improvement (Pfeifer, 1986).

A paragraph from a recent publication (Joyce, 1990) serves to place into perspective the nature of the problem and the need to pay heed to the call for definitive investigation.

The entire field of clinical support of teachers, whether by peers, supervisors, or principals, badly needs study, particularly because it is by far the largest component of staff development in most districts, and its theoretical structure is attractive to district policy makers. To provide teachers with information about effective teaching behavior and with mirrors reflecting the extent to which their practice includes those behaviors appears eminently sensible. Not to study how to do it well, however, makes much less sense. It should not be difficult to locate some teachers who have not been exposed to the content and who do not manifest it in their normal classroom behavior, engineer a really solid treatment, and find out what it takes to make a difference. We believe it can and probably does in many districts, albeit undocumentably, but we can scarcely believe that there is no better research in this area. The newer variations on the theme are on no better ground than the older ones. (pp. 30-31)

Most specifically, Alfonso and Firth (1990) suggest a research agenda for supervision practices. They succinctly capitate much of what the problems are as they ask five questions for researchers to study:
(1) What personal and professional characteristics should instructional supervisors have? 
(2) What conditions are necessary for effective instructional supervision? 
(3) What organizational structures permit the most effective instructional supervision? 
(4) What professional development (preparation programs and certification) would be most effective and desirable? 
(5) What are the dimensions of instructional supervision? (p. 183)

There is an absence of a direct and consistent link between personnel evaluation training, improved instruction and school improvement (Fielding & Shalock, 1985). Therefore, it becomes evident that research needs to be completed in order to answer questions such as: What is the most widely used type of training models for teachers in their own evaluation? What training content is currently being used across the country to support teacher performance evaluation training? What duration of training is usually provided to public school teachers in their performance evaluation processes?

History of Teacher Performance Evaluation

Issues central to teacher evaluation and effectiveness did not appear relevant until students were expected and later required to go to school. For it seemed early in recorded history, quality teachers such as Socrates, Aristotle, Mohammed and Jesus Christ were recognized based upon their ability to captivate student followers or, as in the case of Jesus Christ and Mohammed, disciples. Also, it must be noted that the focus of teaching during antiquity was directed toward adults. Not until the end of the Middle Ages were younger members of the "community" compelled to attend school. These very early "attendance requirements"
generally involved "private institutions and probably did not have to make any effort to attract students. The demand for education far exceeded the supply. In such a situation, there was no public pressure to evaluate teachers, although some of the head masters of these schools attempted to develop criteria by which masters and ushers, or assistant teachers could judge themselves" (Travers, 1981, pp. 14-15).

Throughout the Middle Ages and continuing during the United States Colonial Period, teachers believed that the responsibility of learning was the learners'. During this period, which was characterized by education made up of common schools, teaching was characterized as organization, management and the control of children and efficient use of materials (Smith, 1984). Not until early in this century did the practice of assessing teachers based upon effective teaching strategies emerge. This occurred as crude rating systems and instruments were designated to evaluate a teacher's ability to teach (Peterson, 1982).

Teaching was beginning to surface as having a credible role to play in assuring student learning rather than serving a purely managerial function. Research on the management aspects of teaching continued well into the middle of this century (Jackson, 1966). This author explained the significance of designing effective lesson plans. However, the perception was changing from one of management/control to teaching/learning. The criteria established for teacher accountability were beginning to alter as early as the 1950s (Barr, Bechdolt, Coxe, Gage, Orleans, Renners & Ryans, 1952). Other teacher effectiveness research was conducted throughout the 1960s and 1970s to support this
shifting trend (Gage, 1963; Flanders & Simon, 1969; Rosenshine & Furst, 1973; Coker & Coker, 1979). Hunter (1979) offers compelling evidence that effective teaching elements are definable as she observes that, "Now adequate preparation parallels that of medicine, for it requires the professional to learn, internalize and implement the contributions of science to increased productive human functioning" (p. 62).

Even before the knowledge base of teacher effectiveness became strikingly visible, the ability to perform better evaluation seemed near at hand, and the need for sound evaluation became clearly focused. In fact, Rose (1963) suggested that teachers would embrace evaluation if the reason for such inspection is to promote growth rather than find fault. This atmosphere of contingent acceptance appeared encouraging; however, the largest of two teacher organizations, the National Education Association, continued to strongly echo this evaluation for growth sentiment 14 years later. They would favor evaluation of teachers for instructional growth and improvement, but not for accountability or control (Soar & Soar, 1977).

Such tension in public school teacher's evaluation has marked the later portion of this century and, to some extent, exists today. However, recent and repeated calls for accountability and restructuring has left little time to debate whether teacher evaluation shall be done. The public call for school reform in the 1980s has bridged to the 1990s as public resolve for restructuring. These public outcries leave little doubt that emphasis is on improved teaching and teacher evaluation (Weber, 1987).
These current reform and restructuring forces have contributed to the increased emphasis on teacher performance evaluation. Such public, political activity has promoted a critical investigation of teacher performance evaluation, a very significant school improvement component. Many issues, external and internal, have necessitated this long over-due inspection of teacher effectiveness. On the national scene, the widely read report *A Nation at Risk* (1983) called for a way to differentiate competency levels among the novice teacher, the veteran, seasoned instructor and the highly effective or master teacher. In the report from the Education Commission of the States, *Action for Excellence* (1983), leaders faced with the prospect of having to make hard decisions on state budgets, governors and state legislators alike, confronted the issue squarely as they directed strong sentiments toward teacher evaluation and held the conviction that evaluation is the fuel which would ignite other school improvement measures. Bennett (1986), then Secretary of Education, expressed words intended to challenge the educational community to improve classroom performance. He demanded that schools look carefully at the prospect of paying teachers on the basis of effective and productive teaching rather than longevity and contract status. These strong sentiments were fully endorsed as Finn (1986) proposed the quality development of effective education leaders and the efficient measurement of such effectiveness as essential ingredients if public schools were to achieve excellence.
External Factors

The words of these influential national leaders and reform advocates are not the only external events pushing teacher performance evaluation into the national political spotlight. Activities in a number of states and large school systems across the country reinforced the pressing need for valid performance evaluation systems (Malen & Hart, 1987). A sizable number of large school districts are implementing merit/incentive programs. While hundreds of millions of dollars are being spent on performance-based pay schemes, there is cause for concern. Natrlello and Cohn (1983) explains the absolute necessity of research-linked collaboratively developed and reliable evaluation systems when developing such merit pay plans. Without these prerequisites, they argued, merit pay is sure to fail. Teacher evaluation systems possessing such prescriptions are the exception rather than the rule.

Differing views regarding the purpose of teacher performance evaluation has led to varying effects and differing success levels which contribute to the sensitive and difficult nature surrounding the design and implementation of evaluation systems (Natriello, 1990; Natriello & Dornbusch, 1981). For example, to acquire and maintain federal and state funding programs, local school districts implementing evaluation activities have placed more emphasis on accountability (McLaughlin & Pfeifer, 1988). Subsequently, some districts place less emphasis on evaluation activities that focus on teacher improvement. At the same time, local school districts in many states are forced by legislative decree to provide performance appraisal (McCarthy & Cambron-McCabe, 1987). In nine
states, all districts are expected to use state-mandated teacher performance evaluative criteria and procedures (Wiederhold, 1989).

Holley and Fields (1977) reported the existence of tension between public school districts and professional organizations. More recently, Darling-Hammond et al. (1983) support this early claim as they explain not only the inevitability of similar tensions, but the new competing demands of public accountability and teacher growth. Such competing demands can be and should be merged to serve similar interests. However, at the present time, both purposes lack quality assurances and, thus, fail to capture teacher support—a fundamental obstacle to the betterment of evaluation systems and practices.

In summary, external factors have placed teacher performance evaluation systems under close national and local scrutiny. Teacher performance evaluation has been transformed, as a result of many external demands, from a ceremonial activity to an essential element of effective schooling. Its ambiguity and complexity discourage simple explanation; however, through ongoing study, its dynamic nature can be better understood in hopes of serving both accountability purposes and improvement efforts.

Internal Factors

Studies have focused on the internal school organizational factors instrumental for successful school improvement (Felt, 1985). Central to long-lasting internal school improvement is the permanent acceptance and capacity for change (Hall & Hord, 1987). Important factors, among the
many ingredients associated with a successful internal organizational climate, are the beliefs possessed by teachers that teacher effectiveness can be improved through quality teacher performance evaluation (Johnston, 1985).

Researchers have described the internal school environment using various terminology. Common among the language appears to be readiness—the stability of the institutional climate to support and encourage the development of a teacher performance evaluation system (Darling-Hammond et al., 1983; Knapp, 1982). The organizational components of the school in which the evaluation system will be constructed must be receptive to change and committed to the exhaustive work which will ensure during the design phase of the system development. It is explained in the literature that districts beginning the design of a teacher performance evaluation system appear to share common elements which strengthen the climate of the district and pave the way for a meaningful teacher evaluation program (Knight, 1984). The elements central to a more conducive district climate are: decision for change, environmental readiness, strong leadership committed to teacher performance evaluation and to strategies of open, two-way communication, and active teacher involvement in developing evaluation activities (Darling-Hammond et al., 1983).

Other researchers such as Stiggins and Bridgeford (1985) have studied the perceptions of teaching professionals with regard to the formulation of a teacher performance evaluation system, and they conclude: "Teachers want, at the very least, an evaluation system that
provides accurate information on classroom needs, opportunity to acquire and master new learning approaches, and collegial support when instigating needed changes" (p. 92).

Weber (1987) in his synthesis of the literature on teacher evaluation emphasizes the importance of the confidence, trust and specificity in the evaluation process. His research review clearly places a significant burden on administrators to become better trained coaches for teacher improvement rather than discoverers of deficits. This training theme is essential to foster innovative beliefs and understanding about the purpose of teacher evaluation. Training sessions most effective in promoting and establishing a district-wide culture for evaluation involve not only administrators, but teachers (McLaughlin, 1990). An equally revealing observation is made with regard to internal conditions for successful evaluation by McGreal (1983): "With the exception of additional time spent with supervisors on the responsibilities in the goal setting conference, on observation techniques, and on conferencing and feedback skills, administrators and teachers should receive the same training" (p. 144).

Training provided to teacher and administrator is of vital importance for a stable and conducive culture as well as instructional improvement. However, the secondary products of common language and expectations—components essential to an effective evaluation culture—may be of equal value. Weber (1987) calls this the "requirement of reciprocity—the feeling of mutual respect between teacher and evaluator . . . the heart of the evaluation process" (p. 39).
This respect and feeling of trust, in large measure, results from formal involvement of teachers in the actual design of a total personnel evaluation system and should not be reserved for only a few select, administrative participants. The use of stakeholder committees when attempting to implement new practice results in a sense of ownership, especially in the development of a "new" teacher performance evaluation system (Manatt, 1987; Harris, 1986; Fullan, 1982). Such integral members of the prototype design and review team are usually composed of teachers, administrators, board members, parents, patrons, and, sometimes, students. They serve to determine standards of performance, instrumentation, and procedures for the teacher performance evaluation. These direct and meaningful involvements promote an environment of trust, ownership and substantive collaboration between teacher and administrator—a first step in the establishment of a lasting culture of cooperation, accountability and growth. Manatt (1988) further assures reciprocity of respect between administrator and teacher in his "mutual benefit appraisal system" by challenging stakeholders (including both teachers and administrators) to "create an administrator performance evaluation system and a teacher performance evaluation system simultaneously" (p. 81).

A variety of evaluation models and functions are described in the research (Popham, 1988; Darling-Hammond et al., 1983; McGreal, 1983; Manatt, Palmer & Hidlebaugh, 1976), and it is little wonder why variation exists among training provisions to support the acquisition of skills that enhance the performance of teachers.
Wise et al. (1984), writing for the Rand Corporation, completed a thorough investigation into teacher evaluation practices and revealed that a myriad of promises exist, but just as many problems are present. After studying performance evaluation practices in more than 30 districts, overwhelming evidence suggested that districts where teacher evaluation had better teacher/administrator trust relationships and improved communication were characterized as having more effective and growth-oriented evaluation systems. Rand researchers reported the problems which most often appeared to surface in the literature and in practice as follows:

(1) The most effective teacher performance evaluation design process has not been determined.

(2) A specific external and internal organizational culture promotes readiness for an effective teacher evaluation system; however, researchers hold little promise that a clear consensus on key climatic elements will emerge.

(3) A variety of strategies aimed at developing teacher performance evaluation processes exist.

(4) Environmental requirements, both internal and external, need equal treatment to assure teacher growth and accountability.

(5) The system of teacher performance evaluation has not been adequately researched, nor has sufficient attention been rendered to fully develop effective consistent practice.

Wise et al. (1984) in this Rand report concludes by stating that the importance of internal organizational consistency of district mission and
norms distinguishes the effective from ineffective teacher performance evaluation systems.

The preceding research points to a reasonably well-established, reliable and valid data base from which teacher evaluation design and implementation decisions have been made and will continue to be made. However, essential elements of the teacher performance quality assurance equation appear missing. This study serves to investigate one of the many variables contributing to quality teacher performance evaluation—training for teachers in their own evaluation system. This is done by providing the status of training content, type and duration. A reasonable approach is to clearly determine the status of teacher performance evaluation training and its dependence on certain demographic variables such as district size, and the more overarching variable—teacher performance evaluation model.

An understanding of teacher performance evaluation models existing in United States public school districts which do not have legislated/mandated evaluation procedures and criteria is fundamental when describing training content, type and duration. A reasonable first step is to identify the most well-documented research-based models of teacher performance evaluation and explain each

**Introduction of McGreal's Taxonomy of Teacher Evaluation Models**

The strength of teacher performance evaluation process rests on the quality of the performance expectations or standard defined by the district, the measurement of these standards through the use of data, and
the effectiveness of the communications designed to report findings. The
criteria of performance, technical and interpersonal skills reflect the
essential elements necessary for the success and growth of teaching,
ultimately leading to increased student learning. For the most part, all
models of teacher evaluation have as a goal the growth of teachers and
increased student learning; however, school districts have adopted many
different approaches to the actual process of teacher evaluation.

McGreal (1983) presents five models most often associated with K-12
public school organizations. These models have as common themes the
professional growth of teachers. All models have some expectations of
competencies for the teacher, not all of which present specific classroom
and professional performance criteria. However, all are characterized by
expecting some behavior from the teacher if nothing more than involvement
in passive process orientation which may or may not stress self-
evaluation and goal setting.

Next, it shall be seen that the way one defines the evaluation of
teachers greatly influences the type, duration, and content of training
for the actual evaluation process.

The Association for Supervision and Curriculum Development published
a book in 1983 by Thomas L. McGreal entitled Successful Teacher
Evaluation. This book provides a research-based and practitioner-tested
taxonomy of evaluation models. He lists five distinct models of teacher
evaluation. McGreal carefully establishes the fact that every evaluation
process utilized in public education could be classified as one of the
following five models: (1) Clinical supervision; (2) Artistic; (3) Goal-
setting; (4) Common law; and (5) Product.

Since this study intends to report the status of training to support teacher performance evaluation processes, McGreal's (1983) taxonomy of teacher evaluation models is used to classify the returns from this study. Each of the five models is briefly described.

McGreal's Taxonomy of Teacher Evaluation Models

Clinical supervision model

The perception of many supervisors is that this is the process utilized in their evaluation system. However, the extremely high need for quality two-way supervisory communications between the teacher and supervisor is oftentimes less than adequate in actual practice. A relationship of collegiality and collaboration characterizes this model. "The focus is expected to be on teacher motivation and improvement rather on quality control" (McGreal, 1983, p. 26).

Garmon (1986a), who credits Morris Cogan with stewardship over the clinical approach, explains that the clinical supervisor demands intense, prolonged periods of time in communicating about the classroom observations and analyzing classroom events. The collegiality and collaboration aspects are but two of many clinical supervision assumptions according to Sergiovanni (1982). He offers five additional tenets of clinical supervision: (1) teaching is a complex enterprise; (2) supervision is a team process; (3) teaching will be modified in ways the teacher desires; (4) the supervisor assists the teacher to select improvement goals; and (5) effective supervision provides conditions
which motivate the teacher to self-improvement and enhanced acquisition.

**Artistic model**

This model provides for a variety of outcomes that can be anticipated and unanticipated from a teaching episode; in fact, because of the flexibility of this approach, it may provide the most complete view of teaching and learning. The model's appeal appears to stem from the one of its foundational beliefs that teaching is art and the teacher could be equated with the conductor or the orchestrator of an aesthetic experience. Central to supervision of such a definition of teaching is the complete understanding of the supervisor with regard to the artistic and naturalistic abilities of the teacher. In the artistic model, the general "appreciative and intuitive nature of the observer" is more important than the practice of prescribing a specific set of activities to follow and then observing whether or not they took place.

The strength of this model lies in the sensitivity, perceptivity, and knowledge of the supervisor (Eisner, 1982). It demands a great deal of time and commitment on the part of the supervisor to become very familiar with the expressive character of what the teacher and students were doing. This model holds that teaching is an artistic process; the more aesthetically pleasant and gratifying the experience (of students in the classroom), the more students will learn.

**Goal-setting model**

The goal-setting model is characterized by individualized appraisal of performance leading to decisions for self-improvement. These self-
evaluations result in clear ideas of what is to be accomplished and ultimately lead to growth targets or improvement commitments. As McGreal (1983) states: "The focus should be on showing continual growth and improvement and continually doing things better" (p. 15).

Various procedures have been attributed to goal-setting processes; in fact, early researchers incorporated goal-setting as component of teacher evaluation processes (Bolton, 1973; Redfern, 1980). More recently, to obtain the highest quality and most productive system of teacher performance evaluation, one researcher has integrated goal-setting or performance improvement commitments as a solid component of the evaluation process (Manatt, 1988).

Common law model

This model of teacher evaluation is characterized by high supervisor-low teacher involvement in that the evaluator is often seen as "doing something to the teacher," not with the teacher. In this model, evaluation is seen as observation; that is, there appears to be exclusive reliance on the observation of the classroom teaching episode as the sole data source. This model, most often, is typified by equal treatment of probationary and non-probationary teachers. Furthermore, a major emphasis is placed on summative judgment or the "bottom-line" statement of how one "stacks-up" when compared to colleagues. This component of common law is in sharp contract to a formative approach which is a collegial, collaborative data gathering process characterized by high supervisor-high teacher involvement intended to enhance teacher
performance and student learning. Finally, common law models routinely utilize standardized performance criteria which, most often, have been collected from other districts and approved by a local committee.

Common law evaluation relies on simplified definitions of evaluation and on procedures and processes that have remained virtually unchanged for years and, as McGreal points out, "the label 'common law' is used for such systems since most districts who employ this form of evaluation have done so for so long that they have finally married it by formalizing the procedures" (p. 9).

**Product model**

Student products as measured by criterion-referenced and norm-referenced tests are used in this model to evaluate teacher's productivity—ability to produce student learning. "This model assumes that an important function of teaching is to enhance student learning" (Millman, 1981, p. 146). This model is laden with complexity problems. Even though conceptually this evaluation method appears logical and functionally plausible, it is singularly the most controversial model of teacher evaluation. Proponents suggest that student performance is objective data and, therefore, is better than the subjective ratings most often the dominant source from which final judgments are made in the teacher evaluation process. Those critical of this method of evaluation cite numerous problems: inadequate tests, confounding influences on student growth and lack of reliable statistical measures.

Recent advances in statistical measures and heightened demands for
teacher accountability have placed additional emphasis on this model. However, researchers like Glass (1990) contend: "Student-achievement data cannot tell teachers how to teach; such data are not viewed as credible for distinguishing good teachers from bad ones; and data once gathered will tend to be used" (p. 238).

History of Adult Learning for Professional Development

Early definitions and recent outcomes of adult learning

In order to discuss training or the development of adult learners, a discussion of adult learning theory appears essential. The most concise and purposeful definition was offered by Cronbach (1963) when he stated: "Learning is shown by change in behavior as a result of experience" (p. 71). This definition becomes central to this study, for it connotes the importance of experiences provided to teachers through various training types, specific training content and training amounts. Furthermore, it is assumed that teachers who are more familiar and, ultimately, transfer skills from teacher performance evaluation training will grow to become more highly skilled performers in the classroom, thus more successful promoting student achievement. Growth defined as development of competencies and fulfillment of potential is also used by theories to define learning. For example, Bruner (1966) observes that it is perfectly logical to utilize a pertinent theoretical construct as a baseline from which to describe growth and, ultimately, explain resulting increased competency. He goes on to explain the multiple aspects of growth, and that careful theoretical selection can facilitate the explanation
of virtually any competency shift as growth.

Bruner's focus on cognition explains learning in the context of competency development and when combined with the humanistic approach espoused by Rogers (1969), the training outcomes of action-researchers such as Joyce and Showers (1988) appear to be supported. In fact, such Rogerian influence is clearly noted as one compares the Joyce and Showers (1988) training outcomes with those of Carl Rogers. For example, Rogers (1969) specifies that learning is characterized by personal involvement, self-initiated discovery, evaluation of the learner or trainer, and ultimately the meaning of learning is transferred into the experience of the learner. Analogous language is utilized in the Joyce and Showers (1988) model as they define their outcomes as knowledge or awareness, self-initiated change in attitudes, development of skills through feedback observation and transfer of training. The Rogers characteristics developed more than two decades ago continue to be quite compelling and have had apparent influence on Joyce and Showers (1988) when direct associations are made to their definitions of training outcomes.

Early perspectives on adult learning

As early as Lindeman (1926), adult learning curriculum has been said to have had its foundation in learner's experiences. Mann (1929) explains that experience is the adult learner's most valuable tool for learning and such experience accounts for as much as the teacher's formal knowledge base. Jackson (1931) gives an apology for the most widely used
adult learning methodology—the lecture—and suggests greater use of discussion groups and self-directed experiences. In the same year, Mackaye (1931) explains the need to provide a renewed delivery for adult education as he suggests that training provisions beyond traditional classroom methodology and content must be produced in the field rather than at teacher training colleges.

Teacher training at all times should exemplify and demonstrate the teaching methods found most effective with adult groups. Because of the variety of needs to be served, an adult training program will give opportunity to utilize many teaching methods such as group projects, observation, individual study, and lectures. If the use of each method is preceded by an examination of its potential value and is followed by an analysis of its effectiveness, every lesson will not only serve its own specific purpose, but will also demonstrate a technique of teaching (Wiese & Maxwell, 1939; Russell, 1938).

Accordingly, Knowles (1978), an adult learning theorist, claimed that in 1940 most of the key factors contributing to a unified theory of adult learning had been established. The joining of these essential elements into a unified theory was articulated ten years later by this same researcher (Knowles, 1950). Kahler, Barton, Holmes and Bundy (1985) reported that as of 1955 five stages in the process of adult learning had stabilized and emerged as a possible model. When analyzed, these early stages are seen as having remarkable similarities to current training strategies. These stages are as summarized in the following list:

(1) awareness—first learning of idea or practice,
(2) interest—seek more information about a topic,
(3) evaluation—mental application of the idea or practice,
(4) trial—the idea is actually practiced, and
(5) adoption—the practice is accepted or a transfer occurs.

Contemporary researchers have added the feedback process, as a viable sixth training strategy. This very effective element appears to be a highly desirable component of effective training and creates essential conditions for maintenance and, ultimately, transfer of training (Joyce and Showers, 1988).

Although overlapping in time, Tough's (1971) more comprehensive review of the 30 years of adult learning content research following World War II revealed that adult learning produced a wider scope of rewards. He suggests a range of results which included refined group and individual methods, increased insight, heightened self-awareness and expanded sensitivity. These outcomes to a large extent, paralleled earlier research findings shown by Houle (1972) which codified training settings into three general adult learning modes. His postulated modes are the delivery systems which appeared to foreshadow earlier research outcomes or rewards resulting from adult training. Houle names his delivery modes inquiry, instruction, and performance and defines each as:

**Inquiry**—The mode of inquiry is the process of creating some new synthesis, idea, technique, policy, or strategic action.
**Instruction**—Established objectives exist and an activity has been designated for the learners to achieve the objectives.
**Performance**—The mode of performance is the process of internalizing an idea or using a practice habitually, so that it becomes a fundamental portrait of the way in which a learner thinks about and undertakes his or her work. (p. 32)
Knowles (1978), continuing to articulate an adult learning theory, identified the elements as group work, interest of the learner, readiness to learn, evaluation of the learning and a quality group leader or facilitator. This adult learning researcher offered four assumptions of andragogy (adult learning) summarized below:


2. Role of experience—de-emphasis on transmittal techniques and an emphasis on experiential techniques.

3. Readiness to learn—adult learning expects that learners are ready to learn those things they need.

4. Orientation to learning—adults tend to have a problem-centered orientation to learning rather than a subject orientation; adults assume application of learning follows a learning opportunity.

Nadler (1984) further supports these conclusions as he responds with five assumptions of his own: (1) need for knowledge, (2) self-directed focus of the adult learner, (3) variety and quality of training opportunity, (4) maturity level of the adult learner, and (5) direct applicability of the learning experience. These and other findings define adults as autonomous learners with the need to define their own learning environment and set the rate at which they are introduced to new skills. Manatt (1990) summarizes the assumptions made by Nadler and other authors on the subject of adult learning as he observes: "adult learners need control of learning space, pace and place."

Less convinced in the unity of adult learning theory, Sork and
Buskey (1986) suggest adult education program planning lacks integration and does not appear to build upon, elaborate or otherwise improve and expand existing formulations and prescriptions. They maintain that models of program planning for adult training continue to use research published fifteen or twenty years ago. However, after a careful review of the literature, it appears that learning as defined by early theorists has been generally supported over the last two decades. Training outcomes have been renamed and refined but, overall, much remains consistent with past practice and research (Joyce and Showers, 1988).

It is with this fifty-year historical foundation that the significance of training or adult learning becomes essential for effective teacher performance evaluation. Joyce and Showers (1988) have applied the founding research on adult learning and education to professional classroom teaching and, thus, formulated a hierarchy of training components which represent a methodology widely used to train educators in the acquisition of new skills.

Adult Learning and Teacher Effectiveness

Smylie (1988), in his research on organizational implications of staff development, concluded that with the exception of interactions teachers have with their colleagues about instruction, school context variables made little difference in teacher acquisition of new skills. It was this researcher's conclusion that "teacher change seems rooted in individual perceptions of self as influenced by experiences within classrooms and with teaching colleagues" (p. 24).
Just as teacher evaluation is influenced by external and internal organizational factors, so too staff development is influenced, not only by the organizational context in which it is conducted, but the various psychological states of individuals who are intended to receive training interventions (Darling-Hammond et al., 1983; Fullan, 1982; Katz and Kahn, 1978). Therefore, contextual factors or enabling conditions of training either assist or retard the transfer of effective teaching skills to a level of "executive control." A variety of necessary entry level factors must be present. The most important of these factors is the teacher's perception and attitudes with regard to increased student learning. Guskey (1986) suggests that unless teachers truly experience increased student achievement or desired modifications in student behavior as a result of training, they will not change.

The need for effective and focused staff development (training) is clear. The above research is replete with calls for establishing training for teachers in the essential components of effective classroom instruction, otherwise listed as standards of performance on teacher evaluation instruments.

Essential Elements of Training

Edelfeldt (1975), in a review of literature conducted for the National Education Association, recognizes the inadequacy of inservice education and contends that a clearer focus particularly on the teaching and learning process is needed to strengthen staff development (adult learning) programs. He concludes that "in most (staff development)
programs, little attention is given to formulating a comprehensive concept of inservice education. Too often, objectives are narrow and unrelated to a larger purpose or rationale. The bulk of the programs are of short duration and attack a single topic. . . . The approach is piecemeal. And the result is patchwork" (p. 6). This NEA contention is supported by similar, yet nonaligned, sentiments such as those proposed by Gall, Hailsley, Baker and Perez (1984) as they report "current inservice education appears to consist largely of unintrusive, comfortable experiences that reinforce prevailing patterns of school work. Experiences that seek to improve school work against measured criteria are uncommon" (p. 122). Though harsh in their assessment of current status, much is known about effective staff development and the essential elements of training which facilitate such professional development.

Effective features of instructional skills training include clear statements of objectives and rationales, adequate demonstration, well-designed materials and opportunity for practice and feedback (Showers, 1983; Wolfe, 1984). Sparks (1984), a researcher in effective staff development, provides six steps for the improvement of inservice/staff development. She suggests the steps in a hierarchy:

(1) the need for the adult learners to be ready and committed to the training,

(2) the determination of learner needs,

(3) groupings of learners which focus the training to the ready audience,
(4) implementation of the training plan,
(5) evaluation of the training results, and
(6) reassessment, refinement and extension.

Later, this same researcher proposes exact implementation procedures
which fully explain step four or the implementation of the training plan
(Simmons & Sparks, 1985).

In a recent publication by Joyce and Showers (1988), training
objectives related to effective training for teacher growth are listed
below:

(1) The knowledge or awareness of educational theories and
practices, new curriculums or academic content.
(2) Changes in attitudes toward self (role perception
changes), children (minorities, handicapped, gifted), academic
content (attitudes toward science, English as a second
language, math).
(3) Development of skill (the ability to perform discrete
behaviors such as designing and delivering questions of various
cognitive levels or the ability to perform clusters of skills
in specific patterns as in a synectic exercise).
(4) Transfer of training and "executive control" (the
consistent and appropriate use of new skills and strategies for
classroom instruction). (p. 68)

These well-documented objectives are supported by training
components equally universal in the research. A multiplicity of training
components are available, but four categories most frequently appear in
the research. They are:

(1) introduction to applicable theoretical frameworks relevant to a
topic under discussion,
(2) verified case study or presenting a model segment of the skill
or strategy chosen for study,
(3) drill and rehearsal of the competency or technique,
feedback follow-up regarding approximation to or acquisition of expertness in the selected skill (Joyce & Showers, 1988; Baker & Showers, 1984; Sparks, 1983; Hall & Loucks, 1977; Borg, 1970).

Feedback, the fourth component supported by these researchers, is a subcomponent of coaching for transfer. This highly regarded element of adult learning has gained much credibility as a result of research conducted by Showers (1983, 1984, 1985). She explains that coaching must have a content to coach, and the greater the knowledge base of the professional being coached, the greater the benefits from any coaching effort. Furthermore, she claims that with thorough training—theory, demonstration, practice and feedback/coaching—most teachers will acquire skills and strategies previously absent.

Training and Teacher Performance Evaluation

Holley (1982) contends districts need to make better use of evaluation data. "When evaluator ratings are summarized across competencies or areas, the training needs of both evaluatees and evaluators emerge. The data should be captured and used for the improvement of both the evaluation process and instruction" (p. 7). Furthermore, when building and district aggregate totals of summative teacher performance evaluation are analyzed, identification of staff development needs surface and are useful in determining professional development directions (Stevenson, 1987). So the importance of joint training in the understanding of the evaluation process, criteria and observational data analysis enables such summative ratings to be truly
representative samples of teacher performance and strengthens its usefulness when planning staff development or individual growth plans.

Early in the literature, it was well-established that quality training for effective evaluation was a necessity. In fact, Kirchner and Reisberg (1962) succinctly state that because the basic objective of appraisal is to differentiate performance for either advancement (growth) or demotion (reassignment or termination), the trained supervisors are doing the more effective job. Such sentiments were echoed later in the mid-seventies as evaluation questions were posed by United States business. The importance and need for quality training was postulated as one part of the answer to questions such as:

Do the managers assembled know what the evaluation criteria presently are? There is evidence to indicate that managers within organizations are often confused about what the "real" criteria are that are being used to evaluate them. Confusion of this type leads to a serious deterioration in the value of performance appraisal by the line managers who use and are affected by the system (Kirchner & Reisberg, 1962, p. 37).

More recently, Casner-Lotto (1988) reinforces this private sector sentiment as she reports the essential nature of collaboration and communication between line and staff personnel. Casner-Lotto (1988) states:

When such [training] responsibilities are shared—and when the training department clearly communicates corporate strategy to managers—training efforts are better integrated into operations and planning and reflect both corporate priorities and the needs of employees and their managers. (p. 15)

Implicit in these remarks is the foundation essential for successful evaluation—understanding of the procedures, knowledge of criteria and clear expectations of the outcome.
Stiggins and Bridgeford (1984) explained that administrators believed there to be four major obstacles which limit the development of a more formative evaluation system:

1. teacher's lack of trust in the [evaluation] process,
2. insufficient time for evaluation,
3. the adversarial context of evaluation and
4. principal's skills as evaluators. (p. 24)

These same researchers provide further information on the perceptions of principals regarding teacher evaluation systems and noted "lack of training for evaluators to be a consistent problem among principals" (p. 24). The majority of public school principals surveyed in 1980 admitted that their primary professional development interest was to develop teacher evaluation skills (Gudridge, 1980).

Cogan (1973) discussed the cyclical nature of training for improved teacher performance as he suggested that teachers need a "more careful long-term planning for longer phases of their school-based efforts. They need [training] programs rather than fads and episodes" (p. 225).

Goldhammer et al. (1980) called for adequate and relevant training for evaluators; however, this clinical supervision advocate had an expanded view of supervision. His vision included not only proper training for administrators, but his emphasis on clinical conferencing (communication) skills also included teachers.

Wickert (1987) continued to spell out the importance of training for administrators responsible for teacher performance evaluation as he calls for consistency in training practices, promotes the requirement of a minimum level of evaluator competency and suggests inter-rater
reliability among administrators in a single district. The aforementioned research certainly points to the need to train, at a minimum, evaluators and even hints at the need to include teachers. However, many teacher performance evaluation experts have expanded their view of training beyond the purview of the evaluator and expect equal treatment for the evaluatee. Such is true with Weber (1987) as he reviewed the literature on teacher performance evaluation and strongly endorses the important call made four years earlier by McGreal (1983) that both administrators (evaluator) and teacher (evaluatee) must be well-versed in the objectives and methods of the teacher appraisal system. Furthermore, Weber (1987) lists the recurring theme that teacher participation in every phase of teacher evaluation is essential—from the establishment of the performance criteria to staff development options resulting from a completed evaluation cycle.

All evaluators and evaluatees must be thoroughly trained in the evaluation process (McGreal, 1983; Stiggins & Bridgeford, 1984; Stiggins & Duke, 1988). Everyone involved in the evaluation should know how to use the evaluation instruments to acquire useful, objective data (data collection), interpret and feedback the results (two-way communications), and use the results to improve instruction (effective teaching).

Teacher performance evaluation is far from being an exact science. Recent attempts are being made to lessen the gap between informed speculation and pure science, and this is why it is even more critical that we understand as much about it as possible so that the information can be translated into school improvement, increased student learning and
teacher growth. Thomas (1979) implores the educational profession to investigate the highly probable impact performance evaluation could have on teacher training programs if it were better understood. The results, he believes, would support improved schooling, student learning and teacher growth.

The supervision of teachers by principals and other administrators provides a ready-made opportunity for suggestions regarding professional growth and, ultimately, aggregate, district staff development. As obvious and natural as this might seem, this staff development indicator has received little attention as a general phenomenon (Moore & Hyde, 1981). McLaughlin (1982) also identifies this paradox, as he explains the "lack of integration between teacher evaluation and staff development or district curriculum guides" (p. 11).

Knapp (1982) notes strong teacher performance evaluation programs required both staff involvement and specified relationships between teacher development and evaluation. But despite the urging of researchers and educators themselves, not much has happened. He goes on to speculate that although effective evaluation of individual teachers can provide "a more accurate picture of an individual teacher's needs than, for example, the group needs assessments commonly used . . . systematize evaluation of individual teachers does not yet appear to be a standard part of staff development planning" (p. 8).

McLaughlin and Pfeifer (1988) reported that one of their case study schools "acknowledged the importance of integrating training and evaluation. The staff development office in the district coordinated all
of the training for the variety of roles which take part in career development and teacher evaluation. As both input to and consequences of the evaluation process, staff development is an integral part of the professional life in the district" (p. 39).

Content of Training for Teacher Performance Evaluation

Introduction

Even before educational researchers were discussing teacher performance evaluation as a discrete researchable process, educators such as Katz (1955) had postulated three critical skills necessary for effective supervision. In his estimation, these required supervisor skills were technical—ability to utilize knowledge, methods, and techniques; human—people skills; and conceptual—ability to visualize the outcomes. It takes little imagination to translate these skills into what teacher evaluation researchers are referring to as requisite skills for both evaluator and evaluatee. Repeatedly in the research, investigators referred to the importance of understanding the elements of effective, productive teaching, the necessity for efficient, genuine communications and the ability to make sense of the data collected during a classroom observation (Stiggins & Duke, 1988; McLaughlin & Pfeiffer, 1986; Duke & Stiggins, 1986; Medley et al., 1981; McGreal, 1986).

Equally telling are the three aspects stressed by Stow and Sweeney (1981) when they noted that an effective teacher performance evaluation system must contain three essential aspects: (1) the process must assist teachers in improving performance; (2) the actual data gathered
must be meaningful for the teachers; and (3) a quality supervisory conference.

Most recently, Stiggins and Duke (1988) discuss training provisions for teacher performance evaluation as it relates to policy and practice, presenting a clear case for the need for more information. A paragraph from these researchers' most recent book bears repeating for emphasis and clarity:

It is important to note that we have stressed that both supervisors and teachers should receive evaluation training. We believe it is insufficient to do as many states and districts have done and require training only for supervisors. Such a course of action perpetuates the idea that evaluation is something done to teachers. If teacher evaluation is to contribute to growth, teachers as well as supervisors must be well versed in the process. Both must understand the components of good teaching, master the skills of interpersonal communication and know how to make sense of data collected on teaching. Such knowledge increases the likelihood that teachers eventually will become the agents of their own professional development rather than remaining dependent on others. (p. 137)

These researchers clearly point to the threefold training content necessary for effective growth-oriented evaluation—knowledge of good teaching, interpersonal, two-way communications and data collection and analysis. It is from this research base that the following review of these elements is offered.

**Effective two-way communication**

The importance of understanding two-way communications, particularly how such communications relate to the supervisory or growth conference—an expected result from teacher evaluation—is extremely important. This
claim is supported by Sweeney (1982) who strongly suggests improvement of classroom performance is strengthened with the development of a positive trusting and nonauthoritarian relationship which takes place in clear, concise and meaningful supervisory conferences. It appears professional growth through training intervention is most successful when the growth facilitators are engaged with participating teachers in some fashion (Coladarci & Gage, 1984). Such engagement is characterized by effective two-way communication.

Baker (1971) leaves little doubt as to the importance of two-way communication as he explains the active nature of listening and the need to engage oneself in a posture for listening, both physical and mental. Inasmuch as supervisory conferences, from the supervisor's position, should be three parts listening compared to one part fact sharing, the importance of this communication competency is equally significant for both evaluator and evaluatee. Smyth (1981) asserts that the prime objective of clinical supervision is teacher growth and instructional improvement which is in large measure due to intensive, effective conferencing that occurs between teacher and evaluator, especially in the formative stages of supervision. Joyce and Showers (1988) note that only about two or three times each year the average United States teacher conferences with supervisors on teaching and schooling. Such condition prevails for either of two reasons, supervisors do not have the time or they lack essential skills; it is suggested that both reasons are factors and with additional training at least one of the contributors would be decreased and possibly the outcome would stimulate supervisors to make
Stiggins and Bridgeford (1984) offer the thoughts of principals from one of their district case studies when they were asked to prioritize the barrier to successful performance evaluation. These practitioner evaluators responded by implying that important skills needed to evaluate and the training needed to support such evaluation are frequently not available, not used, or ineffective. At least two sets of skills are lacking: (a) skills in evaluating teacher performance, and (b) skills in communicating with teachers about the evaluation process and results.

**Classroom observation and data analysis**

Classroom observation, coupled with non-evaluative feedback, conducted in teacher assistance groups becomes a valuable tool for training for instructional improvement. Brophy (1979) explained that teachers working as a group obtained useful, applicable feedback on their own teaching behaviors and created a collegial environment by breaking down the barriers of isolation that so frequently inhibit professional development.

Several other studies have shown that teaching can be improved by relating non-evaluative feedback (formative information) to teachers about their classroom teaching behavior. Furthermore, these research conclusions imply that such change in teacher behavior is a result of providing teachers with constructive, specific observational data related directly to a teaching episode which was measurable but, more importantly, relevant to student learning outcomes (Whitehall, 1956;
Research clearly spells out the need to train observers to reduce rating errors and increase the reliability of behavioral observations (Latham, Palmer, & Hiddlebaugh, 1975). After a review of the research, Edwards (1985) points out that most United States public school teachers expect to be observed for the purpose of supervision and evaluation; however, these same teachers are confused about the purposes of observation. Furthermore, this researcher suggests that not only are teachers confused, but lack an understanding of actual observation techniques and, possibly more important, teachers have concerns about the skills of the observers. Other researchers have noted that a wide variety and complexity of data gathering systems exist. Flanders (1970) immediately comes to mind when classroom observation is discussed. This researcher's methodology is but one of many data collection techniques (Simon & Boyer, 1970; Frudden, 1980; Stodolsky, 1984; Edwards, 1985). These researcher findings provide certain evidence that confusion and unfocused training are the norm for classroom observation and data analysis.

Faast (1982) speculates that the results of improved evaluation would be improved teacher performance and educational services for students, along with greater accuracy of teacher performance status and increased mindfulness of the evaluation process. This investigator goes on to call for an increased focus on training in the skills necessary to effectively gather data in the classroom.
Effective teaching

"The average teacher in the United States engages in the formal study of teaching and schooling, including new content and curriculums, for only about three days per year" (Howey et al., 1978, p. 29). These words serve as an indictment or call for the improvement of our training efforts in effective, essential teaching techniques. Edwards (1981) further establishes the fact that teacher behavior directly affects student outcomes and explains the need for researchers to consider aggregate effects within the reality of teaching practice. It is obvious that teachers should be taught the range and the appropriateness of these teaching strategies (Joyce & Showers, 1983). The actual strategies should represent those research-based standards that serve as performance criteria on the evaluation instrument. Manatt (1981) clearly supports such a system of teacher performance evaluation as he calls for the use of effective teaching research to define performance criteria. Faast and Stow (1984) place the burden of producing evidence regarding the skill of teachers on the evaluator and expect teacher evaluators: (1) to understand effective teaching behaviors; (2) to be capable of collecting and analyzing data; (3) to complete supervisory conferences; (4) coach for effective teaching; and (5) write complete summative evaluation reports.

Brophy and Good (1986) suggest that pre-service and in-service teacher education in both subject matter and pedagogy are essential. They continue to explain that "many teachers, even recently trained ones, are not aware of important concepts and findings from research on
teaching" (p. 370).

With well-defined measures (of effective teaching) used during and at the end of the evaluation process, more objective data are gathered and utilized by the teacher in self-evaluation by staff development personnel when formulating inservice programs, and by administrators when making personnel decisions (Beach & Reinhartz, 1984).

Private industry human resource developers advocate training for performance standards established by the organization intended to increase productivity. Nadler and Nadler (1989) concisely sum up the need to provide training to not only the manager, but also the performer as they state:

It is possible, however, that the performers have forgotten the desired behaviors. That happens in many situations where the performance is repetitive; the performer develops shortcuts or engages in practices that do not seem to be wrong—practices that are more comfortable for the performer than those previously learned. If their variations, nevertheless, produce the desired output, it is best to maintain the status quo. If the variations produce a different or lower level of output, or if they conflict with the performance of others, it may be necessary to provide training. (p. 50)

Summary

The extensive nature of the variables involved in this investigation necessitated a review of literature on teacher evaluation and adult learning and the interrelationships between each. Also, it has served to outline the importance of the type and content of training for teacher evaluation.

Many national political leaders and researchers have debated the essential internal and external organizational factors associated with
successful teacher performance evaluation. However, most of what makes a
difference is taking place in local districts as they experiment with a
multitude of promising practices. The literature continually confirms
the reality and expected outcomes of teacher evaluation as defined by the
local practitioners—teacher improvement and teacher accountability.
Overall, the authors and researchers cited in this review would agree
that accountability is the summative purpose of evaluation and would
accept Scriven's (1988) definition of accountability as demonstratable
responsibility. Teacher improvement, the formative result of evaluation,
is often promoted as the ultimate aim of teacher performance evaluation
and frequently appears at odds with the legal demand of and the public's
plea for accountability.

In the view of many commentators, these two goals are, at a minimum,
difficult to demonstrate in any simultaneous fashion and some have
characterized these competing demands as mutually exclusive. This review
has proposed elements which contribute to a joining of these enabling
imperative; namely: a shared knowledge base, mutual trust and respect
between evaluator and evaluatee, evaluation purpose seen as improvement
oriented and the necessity of accountability placed honestly in the
context of the overall mission of the organization.

The conclusion of this review establishes that both the internal and
external motives of the organization can be fostered and strengthened
through the use of quality teacher performance evaluation. This can be
accomplished as one views evaluation's purposes in the environment of
their basic results. Accountability is observed to promote effective
personnel status decisions while serving to provide an avenue for school certification and recognition. The improvement goal directs the professional development of teachers and cultivates school improvement efforts.

Through this literature review, the conspicuous relationship between teacher evaluation and staff development readily emerges. This relationship manifests itself as the public and professional objectives of teacher performance evaluation—improvement and accountability. Staff development for the most part expects instructional improvement while lawful evaluation predicts, with certainty, accountability. It was, in part, the purpose of this review to show that evaluation and staff development should not be seen as two separate functions, rather as complimentary professional activities conducted with a shared understanding of criteria and purpose.

Because one of staff development's primary functions is the improvement of instruction and is half the product of the expected outcome of teacher performance evaluation, it is little wonder why researcher and authors have investigated this linkage. This major predictor of successful teacher evaluation and indicator of purposeful staff development has prompted this investigation into the nationwide status of training for teacher performance evaluation. The writers agree that such a dependency exists; however, a knowledgeable void with regard to actual practice persists and mandates further information to bridge this gap. The preceding review of the literature has, and the subsequent study will provide needed information on the nationwide status of
training for teacher performance evaluation.

Through the delineation of teacher evaluation models into a widely accepted and research-based taxonomy and training categories, this study serves to provide information about this critical linkage by testing the relationship between training types and evaluation models as well as district size. Furthermore, this investigation describes possible commonalities among the actual content and amount of training for teachers in their own evaluation processes.
CHAPTER III. METHODS AND PROCEDURES

Development of the Questionnaire

A review of literature in both adult learning and teacher evaluation served to establish the basis for the major components of the survey instrument. Also, questions were listed on the survey to establish pertinent demographic data.

A specific taxonomy of teacher evaluation models (McGreal, 1983) was identified and used in the questionnaire. Respondents were presented definitions of the five models and asked to select the one most closely representing the teacher performance evaluation system utilized in their district.

Stiggins and Duke (1988) in their research review uncovered commonalities in content for effective teacher performance evaluation training. The preceding review of the literature conformed with their findings and also asserted that effective teaching, interpersonal communication and classroom data/observation analysis skill were essential elements of training for teacher performance evaluation. Respondents were asked either to identify that they did include such content in the training conducted in their district or that such content did not apply (or was absent from their training program).

The definition of training type as defined by Joyce and Showers (1988) was submitted for the respondents' consideration. A request was made in this section for the respondent to select the teacher performance evaluation training type which identified professional development
provisions most often associated with the training provided in their district. Finally, training lengths grouped in hour categories were listed and those answering the questionnaire were asked to select the appropriate range of hours devoted to teacher performance evaluation training.

Instrumentation

The questionnaire (see Appendix A) was divided into four main parts. The initial question served to establish whether or not a district had a teacher evaluation system. The second question established the position of the actual respondent with questions three and four determining the number of teachers in the school district. This was the first main section and constituted the demographic portion of the questionnaire.

An expert jury of practicing educational administrators, both principals and superintendents, professors of educational administration, educational research associates and a private industry consultant were asked to review, refine, and authorize piloting of the questionnaire. Also, an educational researcher was consulted and the portion of the survey on the McGreal taxonomy of models was similar to this researcher's questionnaire (Roettger, 1990). These review activities occurred during the time period beginning October 3, 1989, and ending December 22, 1989. Such reviews and consultations resulted in revisions and refinements. The final validation and authorization allowed transmittal of the initial questionnaire on December 29, 1989.
Sample Design

Much of the information found in the Common Core of Data Public School Universe computer tape (1988) was documented in the Directory of Public Elementary and Secondary Education Agencies (1988) from the United States Department of Education, Office of Educational Research and Improvement. These United States government school census documents served as the population data source from which the stratified random sample was drawn. These data sources contain names and addresses of 15,579 local public school districts in the United States.

The population from which the random sample was made includes those districts with 20 or more teachers and states (41 overall) that do not have legislated mandates for teacher performance evaluation criteria and procedures (Wiederhold, 1989).

The number of teachers was the aggregate number of classroom teachers reported for schools associated with the district on another Department of Education data set Common Core of Data Public School Universe (1988). The number of teachers was not reported by some 1,149 districts. For the purpose of this sample design and selection, the number of teachers was estimated for these districts using average student/teacher ratio for all districts with complete data.

After excluding the states of Alabama, Delaware, Georgia, Hawaii, North and South Carolina, Texas, Tennessee and Virginia (states with mandated criteria and procedures as well as specific training requirements for teacher performance evaluation), the sampling frame contained 9,760 school districts. These states and districts were
grouped into eight geographic regions: New England, Mideast, Southeast, Great Lakes, Great Plains, Southwest, Rocky Mountains, and Far West.

There were two types of public school districts in the sample. The district was either an independent, autonomous local educational agency or a local school district with a supervisory unit which shares a superintendent and administrative services with other local districts. Table 1 depicts the number of local independent and "shared" districts in each size stratum. It should be noted that 89.9 percent of the districts included in the sample are independent local educational agencies. The highest percentage of "shared" districts is found in the smallest size category.

Table 1. Type of public school district categorized by size

<table>
<thead>
<tr>
<th>Number of teachers in the district</th>
<th>Independent local educational agency</th>
<th>Shared support local educational agency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual number of districts(^a)</td>
<td>Estimated number of districts(^b)</td>
</tr>
<tr>
<td>20-119</td>
<td>121</td>
<td>5661</td>
</tr>
<tr>
<td>120-249</td>
<td>110</td>
<td>1762</td>
</tr>
<tr>
<td>250-599</td>
<td>101</td>
<td>960</td>
</tr>
<tr>
<td>600-1999</td>
<td>62</td>
<td>324</td>
</tr>
<tr>
<td>≥2000</td>
<td>52</td>
<td>63</td>
</tr>
<tr>
<td>Total</td>
<td>446</td>
<td>8769</td>
</tr>
</tbody>
</table>

\(^a\)Number of districts who completed the questionnaire.

\(^b\)Number of districts represented by the responding districts.
All school districts which had 2,000 or more teachers were included in the sample with certainty. The remaining districts were grouped into four size strata according to the number of classroom teachers in the district.

**Weighting the sample**

Because the sample design was drawn as a stratified sample with differing sampling rates (all districts with 2000 or more teachers were included), it is necessary to construct weights for each of the responding districts in each stratum. After all responses from the sample were received, expansion weights were computed for each responding district. Since the sample was selected using the numbers of teachers and geographic area to establish strata, the expansion weights for all responding districts within the $j^{th}$ (number of teachers by geographic area) cell were computed as the inverse of the response/population fraction; that is, $(R_j/P_j)^{-1}$ where $R_j$ = the number of districts responding in the $j^{th}$ cell and $P_j$ = the total number of districts in the $j^{th}$ cell (Hickman, 1990).

An example of the procedure can be seen in the stratum with 20 through 119 teachers and geographic area one (1). There were 428 districts in the population. A total of eight (8) districts out of the sample of 15 districts responded. The expansion weight for districts in this cell was $(R_j/P_j)^{-1} = (8/428)^{-1} = 53.50$. Table 2 lists the weights for each of the sample cells. Therefore, every cell has a unique weight.
Table 2. Sample districts' expansion weights listed by size and United States geographic area

<table>
<thead>
<tr>
<th>Geographic area</th>
<th>Number of teachers in the district</th>
<th>Expansion weights</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20-119</td>
<td>120-249</td>
</tr>
<tr>
<td>New England</td>
<td>53.50</td>
<td>16.82</td>
</tr>
<tr>
<td>Mideast</td>
<td>54.18</td>
<td>19.28</td>
</tr>
<tr>
<td>Southeast</td>
<td>65.33</td>
<td>18.55</td>
</tr>
<tr>
<td>Great Lakes</td>
<td>49.36</td>
<td>13.66</td>
</tr>
<tr>
<td>Great Plains</td>
<td>37.78</td>
<td>15.09</td>
</tr>
<tr>
<td>Southwest</td>
<td>45.60</td>
<td>17.25</td>
</tr>
<tr>
<td>Rocky Mountain</td>
<td>72.50</td>
<td>10.40</td>
</tr>
<tr>
<td>Far West</td>
<td>45.53</td>
<td>16.43</td>
</tr>
</tbody>
</table>

Collection of the Data

The survey materials were initially directed to the superintendent of schools in each of the 700 school districts in the sample. Mailing labels were produced from the Common Core of Data Public School Universe (1988) data tape and were not personalized; just the title, "Superintendent of Schools," name and address of the school district were inked on the label. Each envelope included a transmittal letter, questionnaire, a routing slip, computer data sheet, and a prepaid business reply envelope (see Appendix A). The transmittal letter was countersigned by Professor Richard P. Manatt and was printed on School Improvement Model stationery. The transmittal letter explained the purpose of the study, described the sample and provided
the potential respondents with specific directions as well as providing a mailing deadline. The routing slip was provided in order to insure that the person responsible for teacher performance evaluation would receive the questionnaire, particularly in very large districts which have several "layers" of central office administrators. The superintendent or designated official was given until January 19, 1990, to return the response forms.

A total of 382 responses were received as a result of the first mailing. An additional ten days were provided to the school districts sampled before a second mailing was distributed on January 29, 1990. The same materials were provided in the second mailing, with the exception of the inclusion of a yellow insert entitled "Your District's Response is Needed." This insert was attached to the original transmittal letter, urged assistance and requested that the questionnaire be returned by February 9, 1990. As a result of this, a follow-up mailing to the 318 nonresponding districts, another 91 sample districts completed the response forms and returned the questionnaire. This follow-up effort brought the number of districts responding to a total of 473, or a 68 percent response rate. A detailed explanation and table depiction of actual returns is provided in Chapter IV.

In an effort to determine whether or not the results were biased, a random sample of the nonresponding districts was drawn to determine if these districts are in some measurable way different from those who did respond. A telephone interview was conducted with either
the superintendent of schools or another school administrator responsible for teacher performance evaluation. Questions were posed to the designated school official from the original survey instrument.

Treatment of the Data

Each school district in the sample was given a code which corresponded to the order in which they were drawn from the sample. These pre-coded computer response sheets were placed in business reply envelopes which had been assigned the same code number. This was a precaution taken to assure that no district respondent would tamper with the pre-coded response sheet to protect their district's identity. This somewhat time-consuming precaution proved highly rewarding. As each computer sheet was received, the district code was checked against the sample roster. Also, new computer sheets were recopied for those respondent forms which were torn, had staple holes or were otherwise damaged. This extra step was performed because electronic scoring was impossible if any of the above flaws were present. In the event such a "damaged" card was reproduced, it was checked by two researchers to assure accuracy. The use of the envelope-size data forms was employed for ease in handling and aided in the data analyses.

Data analysis

The data gathered came from the 473 questionnaires returned as a result of the initial questionnaire transmittal on December 29, 1989, and the follow-up mailing one month later. Also, the Common Core of Data
Public School Universe (1988) was accessed to provide district size information. Initially, this information was to be gathered from the responses on questions number three and four. However, due to confusion on the part of several districts with the language used in question four, it was decided to access the information directly from the data tape, thus eliminating any possible question with regard to district size. This decision necessitated using the size information from the 1987/88 school year instead of the 1989/90 school year.

Inferential statistical procedures Iowa State University's mainframe computer system utilizing the SPSSx (Norusis, 1983) and the PC CARP (1986) IBM compatible statistical packages was utilized to perform the analytical procedures on the data. The mainframe program was used to complete sample weighting and descriptive tables with the microcomputer system utilized to test the two hypotheses.

Tests for independence of two characteristics such as size of district and training provisions were made using PC CARP (Personal Computer Cluster Analysis Regression Program), a software program designed to analyze survey data where the sample is drawn using stratification with unequal rates of selection. (All districts with 2000 or more teachers were selected for the sample.) Using the weighted data, PC CARP computes a statistic for a test of hypotheses when the entries in the population table of proportions are equal to the product of the marginal proportions. This statistic is called the test for proportionality and, in a table of counts using simple random sampling with complete response, is equivalent to the chi-square test of
independence. The distribution of the test statistics is approximately that of an F-statistic in large samples (Hickman, 1990).

As explained above, the test of proportionality was used to determine independence rather than a chi-square procedures because a weighted sample was employed to estimate to the population. Calculated Fs from this statistical procedure determined whether or not the various teacher evaluation training types were independent of the McGreal taxonomy of evaluation models and whether or not training types are independent of district size.

Specifically, the first test was completed using survey information gathered from the respondents on question 11—training types by question number five—taxonomy of evaluation models. The second test of proportionality was completed using district size stratum constructed in the original sampling procedure crossed by the training types reported in question eleven.

The .05 level of significance (p ≤ .05) was selected and the appropriate degrees of freedom were determined for the specific comparisons to be made. Therefore, if the calculated test value exceeds the .05 table value for the established degrees of freedom, the null hypothesis is rejected. This is to say, one would accept the alternative hypothesis which concludes that (1) type of training for teacher performance evaluation is dependent on the type of evaluation model identified by responding districts, and (2) the training provided to teachers is dependent on the size of the district determined by the number of teachers reported by responding districts.
Descriptive statistical proceduresActual frequencies or raw counts, weighted counts and weighted percentages were produced using the CROSSTABS command, found on SPSSx computer package at the Iowa State University Computer Center. These raw counts and weighted percentages were useful in completing tables to depict and categorize the data.

Treatment of Subjects from Responding Districts

The Committee on the Use of Human Subjects in Research of Iowa State University reviewed and approved this project. It was concluded that the rights and welfare of the human subjects were adequately protected, risks were outweighed by the potential benefits and expected value of the knowledge sought, confidentiality of data was assured, and informed consent was obtained by appropriate procedures.
CHAPTER IV. FINDINGS

Introduction

This investigation was conducted to determine the nationwide status of training for teacher performance evaluation. Data were gathered through the use of a 12-item questionnaire which was developed utilizing a research-based training model, the most widely cited and effective evaluation training content, and a detailed taxonomy of teacher performance evaluation models.

The initial mailing of this questionnaire took place on December 29, 1989, with a second distribution to nonresponding sample districts on January 29, 1990. A total of 700 districts were targeted in the national sample of 41 states and the District of Columbia. The nine excluded states were those with mandated teacher evaluation procedures and criteria.

Questionnaire Return Rate

This investigation was a nationwide effort, and the importance of the response rate deserves detailed analyses. Therefore, specific respondent categories listing return rate by state, geographic region, locale and size are provided in both narrative and table form.

State return rate

The overall return rate of the questionnaire was 68 percent or 473 districts responding out of a possible 700 sampled districts. United States locations returning all questionnaire/s mailed to them were:
Alaska, District of Columbia, Nevada and Wyoming. However, a review of the total number of requests to these areas reveals that less than five questionnaires were mailed to any one of these locations. Nonetheless, other more populated states, where more questionnaires were mailed, returned a high percentage of surveys (see Table 3). It is important to point out that only 41 states and the District of Columbia were surveyed; as noted in Chapter I, nine states mandating teacher evaluation procedures and criteria were not contacted.

The state of New York received 54 questionnaires and returned 30 completed instruments, which resulted in a 56 percent rate of return. Selected California schools were asked to complete 68 surveys and responded with 47 completed questionnaires for a 69 percent return rate. Illinois schools were mailed 40 requests for information and provided reports at a rate of 83 percent or mailed 33 completed surveys. Illinois had the best return rate from states receiving more than 35 questionnaires, and New York's return rate of 56 percent was the poorest from among the 6 states receiving the most mailings.

Overall, the state with the worst return rate was Maine, which returned only one of the three (33 percent) surveys mailed. However, it is important to note that the next poorest return rate was 50 percent from Connecticut.

Geographic region return rate

Table 4 reflects unweighted data and portrays the rate of response by United States regions. This table lists eight geographic areas and
Table 3. Questionnaire return rate categorized by state and district size

<table>
<thead>
<tr>
<th>State</th>
<th>Number of teachers in the district</th>
<th>% Return rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20-119</td>
<td>120-249</td>
</tr>
<tr>
<td>AK</td>
<td>MQ^a CQ^b</td>
<td>MQ^a CQ^b</td>
</tr>
<tr>
<td>AR</td>
<td>8 4 3 2 2 1 1 1</td>
<td>14 8</td>
</tr>
<tr>
<td>AZ</td>
<td>13 9 13 10 20 13 15 11</td>
<td>68 47</td>
</tr>
<tr>
<td>CA</td>
<td>3 1 1 1 3 3 2 1</td>
<td>10 6</td>
</tr>
<tr>
<td>CO</td>
<td>3 2 4 2 5 2 2 1</td>
<td>14 7</td>
</tr>
<tr>
<td>DC</td>
<td>-- -- -- -- --</td>
<td>1 1</td>
</tr>
<tr>
<td>FL</td>
<td>1 1 3 2 5 4 12 11</td>
<td>22 17</td>
</tr>
<tr>
<td>IA</td>
<td>12 11 2 2 2 2 2 2</td>
<td>18 17</td>
</tr>
<tr>
<td>ID</td>
<td>2 1 2 1 1 1 1 1</td>
<td>6 5</td>
</tr>
<tr>
<td>IL</td>
<td>20 15 9 9 7 6 3 2 1</td>
<td>40 33</td>
</tr>
<tr>
<td>IN</td>
<td>6 3 7 5 6 6 3 2 1</td>
<td>23 17</td>
</tr>
<tr>
<td>KS</td>
<td>8 6 2 1 1 1 1 1</td>
<td>13 10</td>
</tr>
<tr>
<td>KY</td>
<td>2 1 6 4 4 2 1 1</td>
<td>14 9</td>
</tr>
<tr>
<td>LA</td>
<td>-- --</td>
<td>4 3 5 1 4 14 8</td>
</tr>
<tr>
<td>MA</td>
<td>5 3 9 6 5 2 2</td>
<td>1 1</td>
</tr>
<tr>
<td>MD</td>
<td>-- --</td>
<td>1 1 3 2 5 5 9</td>
</tr>
<tr>
<td>ME</td>
<td>1 --</td>
<td>1 1</td>
</tr>
<tr>
<td>MI</td>
<td>10 5 13 11 9 8 6 4 1</td>
<td>1 39 29</td>
</tr>
<tr>
<td>MN</td>
<td>9 5 5 3 3 1 2 2 1</td>
<td>20 11</td>
</tr>
<tr>
<td>MO</td>
<td>11 10 4 3 4 2 2 2 2</td>
<td>1 23 18</td>
</tr>
<tr>
<td>MS</td>
<td>2 1 6 3 4 3 1 1</td>
<td>13 8</td>
</tr>
<tr>
<td>MT</td>
<td>3 2 --</td>
<td>1 1</td>
</tr>
<tr>
<td>ND</td>
<td>2 1 --</td>
<td>1 1</td>
</tr>
<tr>
<td>NE</td>
<td>6 5 1 1 1 1 1 1 1</td>
<td>1 10 9</td>
</tr>
<tr>
<td>NH</td>
<td>3 1 1 1</td>
<td>-- --</td>
</tr>
<tr>
<td>NJ</td>
<td>10 4 10 7 10 5 3 2 1</td>
<td>1 34 19</td>
</tr>
<tr>
<td>NM</td>
<td>1 1 1</td>
<td>2 1 1 1 1 1</td>
</tr>
<tr>
<td>NV</td>
<td>-- --</td>
<td>1 1</td>
</tr>
<tr>
<td>NY</td>
<td>13 9 17 7 16 9 5 3 2</td>
<td>54 30</td>
</tr>
<tr>
<td>OH</td>
<td>13 8 13 9 9 5 2 1 4</td>
<td>41 27</td>
</tr>
</tbody>
</table>

^aMQ = The number of questionnaires mailed to districts in individual size categories listed by state.

^bCQ = Returned completed questionnaires in individual size categories listed by state.
Table 3. (Continued)

<table>
<thead>
<tr>
<th>State</th>
<th>20-119</th>
<th>120-249</th>
<th>250-599</th>
<th>600-1999</th>
<th>≥2000</th>
<th>Total</th>
<th>Return rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MQ^a</td>
<td>MQ^b</td>
<td>MQ^a</td>
<td>MQ^b</td>
<td>MQ^a</td>
<td>MQ^b</td>
<td>MQ^a</td>
</tr>
<tr>
<td>OK</td>
<td>11</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>OR</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>PA</td>
<td>7</td>
<td>4</td>
<td>19</td>
<td>11</td>
<td>11</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>RI</td>
<td>--</td>
<td>--</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SD</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>UT</td>
<td>1</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>VT</td>
<td>3</td>
<td>2</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>WA</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>WI</td>
<td>10</td>
<td>5</td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>WV</td>
<td>1</td>
<td>--</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>WY</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Total</td>
<td>217</td>
<td>137</td>
<td>178</td>
<td>118</td>
<td>153</td>
<td>104</td>
<td>89</td>
</tr>
</tbody>
</table>
Table 4: Questionnaire return rate by United States geographic area

<table>
<thead>
<tr>
<th>United States geographic area</th>
<th>Number of districts in sample</th>
<th>Actual returns</th>
<th>Geographic percent return rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>New England</td>
<td>50</td>
<td>27</td>
<td>54.0</td>
</tr>
<tr>
<td>Mideast</td>
<td>139</td>
<td>83</td>
<td>60.0</td>
</tr>
<tr>
<td>Southeast</td>
<td>87</td>
<td>58</td>
<td>67.0</td>
</tr>
<tr>
<td>Great Lakes</td>
<td>165</td>
<td>120</td>
<td>73.0</td>
</tr>
<tr>
<td>Great Plains</td>
<td>92</td>
<td>71</td>
<td>77.0</td>
</tr>
<tr>
<td>Southwest</td>
<td>38</td>
<td>24</td>
<td>63.0</td>
</tr>
<tr>
<td>Rocky Mountains</td>
<td>30</td>
<td>23</td>
<td>77.0</td>
</tr>
<tr>
<td>Far West</td>
<td>99</td>
<td>67</td>
<td>68.0</td>
</tr>
<tr>
<td>Total</td>
<td>700</td>
<td>473</td>
<td></td>
</tr>
</tbody>
</table>

provides the number of returns in the individual regional categories depicted by number of districts sampled in each region, actual number of returns from each region, and total percentage return rate for each region.

The Great Lakes Region, the zone with the second largest representation from the sample, including the states of Illinois, Indiana, Michigan, Ohio and Wisconsin, had a 73.0 percent return rate. This response rate was second to 77.0 percent, which both the Great Plains and Rocky Mountain Region totaled after the data were analyzed.

District size return rate

Upon review of the overall response rates by district size (see Table 5), it can be seen that districts with 2000 or more teachers had
Table 5. District response rate by district size category

<table>
<thead>
<tr>
<th>Number of teachers in the district</th>
<th>Number of districts in population</th>
<th>Number of districts in sample</th>
<th>Number of districts responding</th>
<th>Percent response rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-119</td>
<td>6496</td>
<td>217</td>
<td>137</td>
<td>63.1</td>
</tr>
<tr>
<td>120-249</td>
<td>1891</td>
<td>178</td>
<td>118</td>
<td>66.3</td>
</tr>
<tr>
<td>250-599</td>
<td>986</td>
<td>153</td>
<td>104</td>
<td>68.0</td>
</tr>
<tr>
<td>600-1999</td>
<td>324</td>
<td>89</td>
<td>62</td>
<td>70.7</td>
</tr>
<tr>
<td>( \geq 2000 )</td>
<td>63</td>
<td>63</td>
<td>52</td>
<td>83.4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>9760</strong></td>
<td><strong>700</strong></td>
<td><strong>473</strong></td>
<td></td>
</tr>
</tbody>
</table>

the highest percentage return. It is important to note that this size stratum included all United States districts in the 41 sample states and District of Columbia which had 2000 or more teachers. Each of these, mostly urban, districts was included with certainty and returned questionnaires at an 83 percent level.

In descending order from the best return rate of 83 percent, the following progression is established: 2000 or more teacher districts had an 83 percent return rate; 600-1999 teacher districts accumulated a 70 percent return rate, while 250-599 teacher districts returned questionnaires at a 68 percent rate; 120-249 teacher districts accrued a 66 percent return rate; the smallest size stratum or districts with between 20-119 teachers garnered a return rate of 63 percent. The actual population counts and sample counts are listed above in Table 5.
**District locale return rate**

Table 6 represents unweighted data and lists the rate of response by the sample district's locale. This table lists seven locale codes identifying the district's level of ruralness, urban composition or otherwise metropolitan status based on census data. The number of returns in each of the locales is represented by districts sampled in each locale, actual number of returns from each locale, and total percentage return rate for each locale.

As can be seen after inspection of Table 6, the greatest percentage return rate is from the locale identified as large town. The description of a large town for the purpose of this study is a place not within a standard metropolitan statistical area, but with population greater than or equal to 25,000 and defined as urban by the United States Bureau of Census. The actual return rate from such a locale was 87.0 percent.

**Table 6. Return rate categorized by locale of districts**

<table>
<thead>
<tr>
<th>District locale</th>
<th>Number of districts in sample</th>
<th>Returns from districts in sample</th>
<th>Locale percent return rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large central city</td>
<td>53</td>
<td>33</td>
<td>62.0</td>
</tr>
<tr>
<td>Mid-sized city</td>
<td>114</td>
<td>80</td>
<td>70.0</td>
</tr>
<tr>
<td>Urban fringe of large city</td>
<td>115</td>
<td>83</td>
<td>72.2</td>
</tr>
<tr>
<td>Urban fringe of mid-size city</td>
<td>64</td>
<td>39</td>
<td>60.9</td>
</tr>
<tr>
<td>Large town</td>
<td>23</td>
<td>20</td>
<td>87.0</td>
</tr>
<tr>
<td>Small town</td>
<td>202</td>
<td>137</td>
<td>67.8</td>
</tr>
<tr>
<td>Rural</td>
<td>129</td>
<td>81</td>
<td>62.8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>700</strong></td>
<td><strong>473</strong></td>
<td></td>
</tr>
</tbody>
</table>
poorest representation of districts sampled was from the locale described as an urban fringe of mid-sized city, which is a place within a standard metropolitan statistical area of mid-size central city and defined as urban by the United States Bureau of Census. The calculated percent return rate for this locale was 60.9 percent.

Findings

This section is prefaced with a reminder that expansion weights were used to report frequencies and percents. Narratives and tables are used to explain the findings and comprise the majority of this chapter. All findings are presented to answer the six original research questions posed in Chapter I. Two of these questions were proposed as null hypotheses and are treated last.

Findings reported by weighted frequencies and percentages

As explained in Chapter III, sample weights have been calculated for each responding district. These weights are provided within cells constructed by using size and geographic area. In other words, responding districts from the sample categorized by geographic area and size represent a specified number of districts through the use of an expansion weight. Therefore, for the purpose of the findings reported herein, all frequencies are based on such expansion weights. Also, findings will refer to an estimated population as determined by these expansion weights. For example, in the smallest district size category (20-119 teachers) listed by the Rocky Mountain geographic region, one responding (sample) district represents the equivalent of 72.50 districts.
in the population.

Finally, for the purpose of reporting the data, weighted counts and weighted percentages are utilized to explain and graphically depict the findings. Actual counts are listed in order to provide the reader with the raw total of responding districts. Ultimately, these counts were used to calculate the expansion weight assigned to the responding district. However, all the findings provided in this section are explained as an estimate of the population using expansion weights to calculate counts and percentages; any actual count is listed as the "number of districts who completed the questionnaire."

Research questions one through four

The title of this study, for the most part, explains the major thrust of the research--Teacher Performance Evaluation: A Nationwide Status Report of Type, Content and Duration of Training for Public School Teachers. An investigation designed to provide a heretofore unavailable nationwide status report of training provisions for teacher performance evaluation has an obligation to describe actual national training practice. This description will serve to answer questions about type of training (characterized as either awareness, modeling, practice or feedback) content of training (represented by effective teaching, two-way communication, analyses and collection of data) and amount of training. Also, because a taxonomy of models was used to identify evaluation systems, this study will serve to provide information regarding the variety of evaluation processes that do exist.
The first research question

1. What are the types of teacher evaluation training among public school districts using specified evaluation models?

This section of the descriptive findings will primarily explain the results listed on the questionnaires with respect to the three most widely reported models of teacher performance evaluation—clinical supervision, common law and goal-setting. The product and artistic models of teacher performance evaluation were reported a combined total of eight times. These eight instances represent 200 districts in the population, characterizing their evaluation model as either product or artistic. More specifically, 51 or 0.6 percent of the districts in the sample represented themselves as product evaluation districts, while 149 or 1.60 percent of the responding districts indicated that they were artistic model districts (see Table 17).

This inappreciable number of districts reporting evaluation processes representative of the product and artistic model has precluded a detailed coverage of either of these evaluation models in the ensuing description of the questionnaire results. However, actual counts, weighted counts and percentages based on weighted counts for each of these scantily reported models have been listed on each of the following tables.

Awareness training by evaluation model When reviewed across evaluation models, awareness level training was reported by 55.7 percent of common law districts as a training strategy for their teacher evaluation systems (see Table 7). Goal-setting and clinical supervision
districts indicated that they used awareness level training 40.0 percent and 29.0 percent, respectively.

**Modeling training by evaluation model**

The modeling level training component, when viewed in the evaluation model context, was reported at a 20.7 percent rate by common law districts (districts who characterized their evaluation model as common law) as a training strategy for their teacher evaluation systems (see Table 7). Goal-

Table 7. Types of teacher evaluation training by evaluation models (first research question)

<table>
<thead>
<tr>
<th>Teacher evaluation models</th>
<th>Teacher evaluation training types</th>
<th>Awareness</th>
<th>Modeling</th>
<th>Practice</th>
<th>Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQ^a ED^b %^c</td>
<td>CQ^a ED^b %^c</td>
<td>CQ^a ED^b %^c</td>
<td>CQ^a ED^b %^c</td>
<td></td>
</tr>
<tr>
<td>Common law</td>
<td>66 1526 55.7</td>
<td>37 568 20.7</td>
<td>24 488 17.8</td>
<td>13 158 5.8</td>
<td></td>
</tr>
<tr>
<td>Goal setting</td>
<td>31 634 40.0</td>
<td>16 441 27.9</td>
<td>16 330 20.8</td>
<td>13 178 11.2</td>
<td></td>
</tr>
<tr>
<td>Product setting</td>
<td>-- -- --</td>
<td>1 5 9.9</td>
<td>-- --</td>
<td>1 46 90.0</td>
<td></td>
</tr>
<tr>
<td>Clinical supervision</td>
<td>40 802 29.0</td>
<td>42 729 26.4</td>
<td>37 734 26.6</td>
<td>26 497 18.0</td>
<td></td>
</tr>
<tr>
<td>Artistic</td>
<td>1 1 0.8</td>
<td>-- --</td>
<td>-- --</td>
<td>-- --</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>138 2962 40.8</td>
<td>96 1744 24.0</td>
<td>81 1681 23.1</td>
<td>49 878 12.1</td>
<td></td>
</tr>
</tbody>
</table>

^aCQ = Number of districts who completed the questionnaire.

^bED = Number of districts in the population estimated by the responding districts in the sample.

^c% = Percent usage of training types in evaluation model districts based on the expansion weights presented in Table 2.
setting and clinical supervision districts indicated that they used modeling level training 27.9 percent and 26.4 percent rate, respectively.

**Practice training by evaluation model**  Practice level training, when referenced to teacher evaluation models, was reported at a 17.8 percent rate by common law and 20.8 percent rate by goal-setting districts as a strategy for their teacher evaluation systems (see Table 7). Clinical supervision districts reported the highest frequency of this training strategy as they indicated use at a 26.6 percent rate.

**Feedback training by evaluation model**  The frequency rates of implementation for the evaluation model districts using feedback training in support of their teacher performance evaluation systems were quite different (see Table 7). This is noted as common law districts reported the use of feedback at a 5.8 percent rate; however, clinical supervision districts reported the use of feedback at an 18.0 percent rate. Goal-setting districts indicated use of feedback at a level between these two rates or 11.2 percent.

**The second research question**

2. What is the content and duration of teacher evaluation training utilized in public school districts when categorized by models of teacher evaluation?

This two-part question, when answered with the data obtained from this study, shall describe the content and duration of teacher evaluation training when districts are grouped by their model of evaluation. More specifically, the duration component of the questionnaire yielded data
which depict results of the question asked regarding "the time category that best represents the duration of teacher performance evaluation training provided to teachers." The responding districts were asked to indicate which time category most closely reflected the amount of training provided to teachers in their evaluation systems. The content portion was asked using three research-based content categories (effective teaching, interpersonal communications and observational data analysis) as choices for districts as they characterized their evaluation training content.

**Effective teaching training content by evaluation model**

Of the estimated districts reporting the use of training for teacher performance evaluation, 71.5 percent indicated that effective teaching was a component of their training program. Table 8 depicts this finding and is listed below. Another 28.5 percent of the estimated districts reported that effective teaching training content did not apply or was absent from their teacher evaluation training program. Of the districts utilizing this content, clinical supervision, goal-setting and common law models appeared to be most frequently represented. Clinical supervision districts reported 79.7 percent utilization of effective teaching content as a component of their evaluation training program. Goal-setting and common law districts reported use of this content at a 73.5 percent rate and 64.9 percent rate, respectively.

**Interpersonal communications training content by evaluation model**

Of the total estimated districts reporting interpersonal communication training for teacher performance evaluation, 72.8 percent identified this
### Table 8. Effective teaching training content by evaluation models (second research question)

<table>
<thead>
<tr>
<th>Teacher evaluation models</th>
<th>Effective teaching by models</th>
<th>Presence of effective teacher training for teacher evaluation</th>
<th>Absence of effective teacher training for teacher evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQ</td>
<td>ED</td>
<td>%</td>
</tr>
<tr>
<td>Common law</td>
<td>136</td>
<td>2476</td>
<td>64.9</td>
</tr>
<tr>
<td>Goal setting</td>
<td>58</td>
<td>1274</td>
<td>73.5</td>
</tr>
<tr>
<td>Product</td>
<td>1</td>
<td>5</td>
<td>9.9</td>
</tr>
<tr>
<td>Clinical supervision</td>
<td>141</td>
<td>2576</td>
<td>79.7</td>
</tr>
<tr>
<td>Artistic</td>
<td>3</td>
<td>88</td>
<td>59.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>339</td>
<td>6420</td>
<td>71.5</td>
</tr>
</tbody>
</table>

| aCQ = Number of districts who completed the questionnaire. |
| bED = Number of districts in the population estimated by the responding districts in the sample. |
| c% = Proportion of the estimated population responses based on the expansion weights presented in Table 2. |

Content as a component of their training program (see Table 9). These are somewhat similar results to those noted in the effective teaching content area. However, specific distributions among the models differed slightly from those reported in the effective teaching content category. The goal-setting, clinical supervision and common law evaluation models reported 73.7 percent, 88.1 percent and 58.8 percent, respectively, as they were categorized by interpersonal communications training content.

**Observational data analysis skills training content by evaluation model** When districts were asked whether or not they provide observation and analysis skills as a component of their evaluation.
Table 9. Interpersonal communications training content by evaluation models (second research question)

<table>
<thead>
<tr>
<th>Teacher evaluation models</th>
<th>Presence of interpersonal communications training for teacher evaluation</th>
<th>Absence of interpersonal communications training for teacher evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQ&lt;sup&gt;a&lt;/sup&gt;</td>
<td>ED&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Common law</td>
<td>103</td>
<td>2089</td>
</tr>
<tr>
<td>Goal setting</td>
<td>59</td>
<td>1278</td>
</tr>
<tr>
<td>Product</td>
<td>1</td>
<td>46</td>
</tr>
<tr>
<td>Clinical supervision</td>
<td>144</td>
<td>2828</td>
</tr>
<tr>
<td>Artistic</td>
<td>3</td>
<td>88</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>310</td>
<td>6328</td>
</tr>
</tbody>
</table>

<sup>a</sup>CQ = Number of districts who completed the questionnaire.

<sup>b</sup>ED = Number of districts in the population estimated by the responding districts in the sample.

<sup>c</sup>% = Proportion of the estimated population responses based on the expansion weights presented in Table 2.

training program, 48.2 percent replied "yes" (see Table 10). There was a somewhat equivalent frequency rate among the three most widely used models. Common law districts listed observation and analysis skills as a training component at 45.7 percent, with districts characterizing themselves as goal-setting reporting 38.8 percent. The model type having the highest percentage use rate of observation/data analysis training was
the clinical supervision model. Districts specifying this model reported
the use of such training content at a 56.8 percent rate.

**Overall evaluation training content by product and artistic models**

Those districts characterizing their evaluation model as either artistic
or product and who report the allowance for effective teaching training
represent an estimated 88 and five districts, respectively (see Table 8).
Interpersonal communications training responses yielded an estimated
frequency use rate of 46 product model districts and 88 artistic model
districts.

Table 10. Observational data analysis skills training content by
evaluation models (second research question)

<table>
<thead>
<tr>
<th>Teacher evaluation models</th>
<th>Presence of observational data analysis skills training for teacher evaluation</th>
<th>Absence of observational data analysis skills training for teacher evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQ^a ED^b</td>
<td>%c</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----------</td>
<td>----</td>
</tr>
<tr>
<td>Common law</td>
<td>91 1662</td>
<td>45.7</td>
</tr>
<tr>
<td>Goal setting</td>
<td>33 672</td>
<td>38.8</td>
</tr>
<tr>
<td>Product</td>
<td>1 5</td>
<td>9.9</td>
</tr>
<tr>
<td>Clinical supervision</td>
<td>104 1808</td>
<td>56.8</td>
</tr>
<tr>
<td>Artistic</td>
<td>4 74</td>
<td>49.4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>223 4221</td>
<td>48.2</td>
</tr>
</tbody>
</table>

^a CQ = Number of districts who completed the questionnaire.

^b ED = Number of districts in the population estimated by the responding districts in the sample.

^c % = Proportion of the estimated population responses based on the calculated weights presented in Table 2.
districts (see Table 9). Training in observational data analysis skills for product and artistic model districts is similarly low in frequency. An estimated five product model districts reported such training, while an estimated 74 artistic model districts employed observational data analysis skills training content in support of teacher evaluation (see Table 10).

**Common law model by duration of teacher evaluation training**

Overall, 74 percent of the districts that characterize their evaluation system as common law report that they offer less than 15 hours of training to teachers in support of teacher performance evaluation. The 15-25 hour time category was reported by common law districts at a 19.1 percent rate, while the 26-40 hour category yielded a 5.8 percent rate and 41-55 hour category showed only a fraction of 0.3 percent. There were no reports of common law districts offering 55 or more hours of training to their teachers.

**Goal-setting model by duration of teacher evaluation training**

The goal-setting evaluation model districts reported training amounts in each of the duration categories (Table 11). These districts reported the less than 15 hours of training category at a 50.6 percent frequency rate. Slightly more than half of the districts could be found in this, the highest selected category, whereas only 5.3 percent revealed that training totaled 55 hours or more. The training duration categories of 15-25, 26-40 and 41-55 hours were delineated by goal-setting districts at rates of 29.1 percent, 5.8 percent, and 2.1 percent, respectively.
Table 11. Duration of teacher evaluation training by evaluation models (second research question)

<table>
<thead>
<tr>
<th>Teacher evaluation models</th>
<th>Less than 15 hours</th>
<th>15-25 hours</th>
<th>26-40 hours</th>
<th>41-55 hours</th>
<th>55 hours or more</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQ</td>
<td>ED</td>
<td>%</td>
<td>CQ</td>
<td>ED</td>
</tr>
<tr>
<td>Common law</td>
<td>106</td>
<td>2201</td>
<td>74.7</td>
<td>28</td>
<td>564</td>
</tr>
<tr>
<td>Goal setting</td>
<td>39</td>
<td>727</td>
<td>50.6</td>
<td>17</td>
<td>417</td>
</tr>
<tr>
<td>Product</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1</td>
<td>46</td>
</tr>
<tr>
<td>Clinical supervision</td>
<td>70</td>
<td>1422</td>
<td>50.1</td>
<td>44</td>
<td>811</td>
</tr>
<tr>
<td>Artistic</td>
<td>3</td>
<td>77</td>
<td>58.5</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>218</td>
<td>4426</td>
<td>59.8</td>
<td>91</td>
<td>1842</td>
</tr>
</tbody>
</table>

\textsuperscript{a}CQ = Number of districts who completed the questionnaire.

\textsuperscript{b}ED = Number of districts in the population estimated by the responding districts in the sample.

\textsuperscript{c}\% = Percent usage in training duration categories specified by evaluation models based on the expanded weights presented in Table 2.
Clinical supervision model by duration of teacher evaluation training

In Table 11, data indicate that 50.1 percent of the clinical supervision districts describe the amount of evaluation training to be less than 15 hours. In 28.6 percent of these districts, training duration was disclosed to be 15-25 hours. Clinical supervision districts specified that 14.8 percent trained teachers between 26 and 40 hours. Only 1.1 percent of the districts reported training at 41-55 hours or more, whereas 5.5 percent of clinical model districts reported in the 55 hours or more category.

Artistic and product models by duration of teacher evaluation training

Of the estimated product model districts, 46 reported providing 15-25 hours of training while five indicated that 55 hours of training were provided to teachers in the product model districts. At least 77 artistic model districts indicated a training duration of less than 15 hours, with five districts in each of the 15-25 and 26-40 hour training categories. No artistic district reported training at a level at 55 hours or above (see Table 11).

The third research question

3. What type, content and duration of evaluation training is most frequently used by public school districts of various sizes?

Awareness type training by district size

The percentage of awareness level training is relatively consistent across district strata (see Table 12). For example, in order of size from small to large, 42.3 percent of the districts in the 20-119 teachers stratum employ the use of
Table 12. Types of teacher evaluation training by district size (third research question)

<table>
<thead>
<tr>
<th>Number of teachers in the district</th>
<th>Awareness</th>
<th>Modeling</th>
<th>Practice</th>
<th>Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQ&lt;sup&gt;a&lt;/sup&gt;</td>
<td>ED&lt;sup&gt;b&lt;/sup&gt;</td>
<td>%&lt;sup&gt;c&lt;/sup&gt;</td>
<td>CQ&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>20-119</td>
<td>42</td>
<td>2034</td>
<td>42.3</td>
<td>21</td>
</tr>
<tr>
<td>120-249</td>
<td>35</td>
<td>555</td>
<td>39.9</td>
<td>22</td>
</tr>
<tr>
<td>250-599</td>
<td>30</td>
<td>273</td>
<td>34.6</td>
<td>26</td>
</tr>
<tr>
<td>600-1999</td>
<td>13</td>
<td>80</td>
<td>29.7</td>
<td>16</td>
</tr>
<tr>
<td>≥ 2000</td>
<td>18</td>
<td>20</td>
<td>37.7</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>138</td>
<td>2962</td>
<td>40.5</td>
<td>97</td>
</tr>
</tbody>
</table>

<sup>a</sup>CQ = Number of districts who completed the questionnaire.

<sup>b</sup>ED = Number of districts in the population estimated by the responding districts in the sample.

<sup>c</sup>% = Percent usage of training types in district size categories based on the expansion weights presented in Table 2.
awareness training to support their evaluation process, 39.9 percent of
districts in the 120-249 stratum utilize awareness training, 34.6 percent
in the 250-599 stratum, 29.7 percent in the 600-1999 stratum and 37.7
percent of the responding sample districts in the largest or 2000 or more
teachers stratum applied the use of awareness training to support their
evaluation process.

**Modeling type training by district size**

With the exception of the size stratum 250-599 teachers, districts report consistent use of
modeling as a type of training strategy in support of teacher performance
evaluation. The districts in the 250-599 size stratum indicate modeling
usage at a 35 percent frequency level. Each of the remaining district
size stratifications range from 21 percent for the smallest district
stratum to 28 percent for the largest district stratum (see Table 12).

**Practice type training by district size**

The percentage rates
among the five district size strata in the implementation of practice as
a type of training strategy in support of teacher performance evaluation
differed only slightly. In fact, the range is from 24.5 percent for the
highest reported use among the districts in the 20-119 size stratum to
the lowest implementation rate of 19.7 percent for the 600-1999 size
stratum. Table 12 provides actual counts, weighted (expanded) counts and
percentages for the remaining strata.

**Feedback type training by district size**

Listed in Table 12 is
the feedback training component enumerated by size of district. It can
be seen that little differences exist among four of the five size strata
percentage rates when viewed in the feedback training strategy column.
The possible exception would be the 600-1999 teachers size stratum which reports 18.0 percent of the districts employ the use of feedback. The range begins at 10.0 percent for 120-249 size stratum and with no other size stratum reporting a frequency higher than 12.3 percent; the only exception is the 600-1999 teachers size stratum.

Effective teaching training content by district size

The data provided in Table 13 report information about effective teaching training content by district size reported as number of teachers in districts. Of

Table 13. Effective teaching training content by district size (third research question)

<table>
<thead>
<tr>
<th>Number of teachers in the district</th>
<th>Presence of effective teaching content for teacher evaluation training</th>
<th>Absence of effective teaching content for teacher evaluation training</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQ&lt;sup&gt;a&lt;/sup&gt;</td>
<td>ED&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>20-119</td>
<td>83</td>
<td>4032</td>
</tr>
<tr>
<td>120-249</td>
<td>80</td>
<td>1279</td>
</tr>
<tr>
<td>250-599</td>
<td>82</td>
<td>784</td>
</tr>
<tr>
<td>600-1999</td>
<td>52</td>
<td>276</td>
</tr>
<tr>
<td>≥ 2000</td>
<td>43</td>
<td>53</td>
</tr>
<tr>
<td>Total</td>
<td>340</td>
<td>6423</td>
</tr>
</tbody>
</table>

<sup>a</sup>CQ = Number of districts who completed the questionnaire.

<sup>b</sup>ED = Number of districts in the population estimated by the responding districts in the sample.

<sup>c</sup>Proportion of the estimated population responses based on the expansion weights presented in Table 2.
the total estimated districts in the 2000 or more teachers size stratum, 86.5 percent reported the incorporation of effective teaching content in their evaluation training, whereas districts in the smallest size category, 20-119 teachers, indicated that 68.0 percent of their training content included effective training content.

It is interesting to note that as the district size category increases, so too does the percentage use of effective teaching content increase. For example, districts with 120-249 teachers report the use of effective teaching strategy at 73.3 percent, districts with 250-599 teachers report 84.4 percent and districts with 600-1999 teachers report 86.5 percent level of use with regard to teaching effectiveness training in support of teacher evaluation processes.

**Interpersonal communications training content by district size**

The data provided in Table 14 provide information relative to interpersonal communications training content reported by district size. Of the estimated districts in the 120-249 teachers stratum, 74.2 percent report the use of interpersonal communications content in their evaluation training. Districts in the size category 20-119 exhibit similar results as they report 73.2 percent of their training provisions include interpersonal communications as a component of training. Districts in the remainder of the categories report similar interpersonal communications training content utilization rates.

Table 14 offers a sequential listing of these results. For example, districts with 250-599 teachers report 69.9 percent, districts with 600-1999 teachers 65.0 percent and districts with 2000 or more teachers
Table 14. Interpersonal communications training content by district size (third research question)

<table>
<thead>
<tr>
<th>District size</th>
<th>Presence of interpersonal communications training for teacher evaluation</th>
<th>Absence of interpersonal communications training for teacher evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQ</td>
<td>ED</td>
</tr>
<tr>
<td>20-119</td>
<td>87</td>
<td>4121</td>
</tr>
<tr>
<td>120-249</td>
<td>81</td>
<td>1308</td>
</tr>
<tr>
<td>250-599</td>
<td>68</td>
<td>649</td>
</tr>
<tr>
<td>600-1999</td>
<td>40</td>
<td>211</td>
</tr>
<tr>
<td>&gt; 2000</td>
<td>35</td>
<td>43</td>
</tr>
<tr>
<td>Total</td>
<td>311</td>
<td>6332</td>
</tr>
</tbody>
</table>

*CQ = Number of districts who completed the questionnaire.

*ED = Number of districts in the population estimated by the responding districts in the sample.

*Proportion of the estimated population responses based on the expansion weights presented in Table 2.

depict a level of use with regard to interpersonal communications training content at 72.9 percent.

Observational data analysis skills training content by district size

The data provided in Table 15 report information about observational and analytical skills training by district size. The districts with the largest percentage rate in this training content employ between 600 and 1999 teachers. These districts indicated provisions for observational data analysis skills content in teacher evaluation at a 67.5 percent
rate. Districts with 2000 or more teachers report the use of observational data analysis skills training at a slightly lower percentage rate (67.5 percent). It is noted that districts in the smallest size category, 20-119 teachers, have the smallest use rate (45.8 percent) for such training; however, it increases through the next two categories. For example, districts with 120-249 teachers detail observational data analysis training content at 48.4 percent while the next larger category is listed at 55.0 percent.

Table 15. Observational data analysis training content by district size (third research question)

<table>
<thead>
<tr>
<th>Number of teachers in the district</th>
<th>Presence of observational data analysis skills training for teacher evaluation</th>
<th>Absence of observational data analysis skills training for teacher evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQ^a</td>
<td>ED^b</td>
</tr>
<tr>
<td>20-119</td>
<td>53</td>
<td>2602</td>
</tr>
<tr>
<td>120-249</td>
<td>54</td>
<td>853</td>
</tr>
<tr>
<td>250-599</td>
<td>54</td>
<td>512</td>
</tr>
<tr>
<td>600-1999</td>
<td>41</td>
<td>216</td>
</tr>
<tr>
<td>≥ 2000</td>
<td>31</td>
<td>38</td>
</tr>
<tr>
<td>Total</td>
<td>233</td>
<td>4221</td>
</tr>
</tbody>
</table>

^aCQ = Number of districts who completed the questionnaire.

^bED = Number of districts in the population estimated by the responding districts in the sample.

^cProportion of the estimated population responses based on the expansion weights presented in Table 2.
Duration of teacher evaluation training by district size  

Table 16 represents the training duration by size of district determined by the number of teachers reported in each district. The greatest number of districts responded that they utilize less than 15 hours of teacher performance evaluation training. The district size strata beginning with the smallest (20-119 teachers) up to largest (2000 or more teachers) show little differences in percentages across all duration categories. The only possible exception might be districts with 600-1999 teachers, of which 11.7 percent provide 55 hours or more training whereas the other five district size strata report between 0 and 3.4 percent.

The fourth research question

4. What evaluation models are utilized most frequently by public school districts of various sizes?

Findings in this section will provide a description of data which represent the use of evaluation models as reported by districts in various size categories. However, as explained earlier, these reporting districts were assigned precise weights and ultimately served to estimate the entire population. In this section, types of evaluation models are categorized by district size and will be reported as weighted counts and percents. This descriptive information has been depicted in Table 17.

Common law evaluation model by district size  
The common law model is widely represented throughout the size categories listed in this study. The highest representation of the common law model appears to
Table 16. Duration of teacher evaluation training by district size (third research question)

<table>
<thead>
<tr>
<th>Number of teachers in the district</th>
<th>Less than 15 hours</th>
<th>15-25 hours</th>
<th>26-40 hours</th>
<th>41-55 hours</th>
<th>55 hours or more</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQ(^{a}) ED(^{b}) (^{c})</td>
<td>CQ(^{a}) ED(^{b}) (^{c})</td>
<td>CQ(^{a}) ED(^{b}) (^{c})</td>
<td>CQ(^{a}) ED(^{b}) (^{c})</td>
<td>CQ(^{a}) ED(^{b}) (^{c})</td>
</tr>
<tr>
<td>20-119</td>
<td>65 3024 62.3</td>
<td>24 1142 23.5</td>
<td>9 473 9.7</td>
<td>1 49 1.0</td>
<td>3 165 3.4</td>
</tr>
<tr>
<td>120-249</td>
<td>54 833 56.8</td>
<td>26 433 29.5</td>
<td>9 143 9.8</td>
<td>1 19 1.3</td>
<td>2 38 2.6</td>
</tr>
<tr>
<td>250-599</td>
<td>46 443 56.3</td>
<td>19 180 22.9</td>
<td>12 113 14.4</td>
<td>5 50 6.4</td>
<td>-- -- --</td>
</tr>
<tr>
<td>600-1999</td>
<td>26 133 47.3</td>
<td>12 75 26.7</td>
<td>8 40 14.4</td>
<td>-- -- --</td>
<td>7 33 11.7</td>
</tr>
<tr>
<td>≥ 2000</td>
<td>29 35 64.8</td>
<td>10 12 22.2</td>
<td>5 6 11.0</td>
<td>-- -- --</td>
<td>1 1 2.0</td>
</tr>
<tr>
<td>Total</td>
<td>220 4467 60.0</td>
<td>91 1842 24.8</td>
<td>43 776 10.4</td>
<td>7 119 1.6</td>
<td>13 237 3.2</td>
</tr>
</tbody>
</table>

\(^{a}\) CQ = Number of districts who completed the questionnaire.

\(^{b}\) ED = Number of districts in the population estimated by the responding districts in the sample.

\(^{c}\) \(\%\) = Percent usage in training duration categories specified by size based on the expansion weights presented in Table 2.
Table 17. Teacher evaluation models by district size (fourth research question)

<table>
<thead>
<tr>
<th>Number of teachers in the district</th>
<th>Common law</th>
<th>Goal setting</th>
<th>Product</th>
<th>Clinical supervision</th>
<th>Artistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$CQ^a$</td>
<td>$ED^b$</td>
<td>$%^c$</td>
<td>$CQ^a$</td>
<td>$ED^b$</td>
</tr>
<tr>
<td>20-119</td>
<td>54</td>
<td>2554</td>
<td>43.1</td>
<td>24</td>
<td>1151</td>
</tr>
<tr>
<td>120-249</td>
<td>52</td>
<td>829</td>
<td>45.8</td>
<td>22</td>
<td>356</td>
</tr>
<tr>
<td>250-599</td>
<td>36</td>
<td>355</td>
<td>37.0</td>
<td>18</td>
<td>160</td>
</tr>
<tr>
<td>600-1999</td>
<td>20</td>
<td>100</td>
<td>31.2</td>
<td>11</td>
<td>70</td>
</tr>
<tr>
<td>$\geq$ 2000</td>
<td>27</td>
<td>33</td>
<td>54.4</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>189</td>
<td>3871</td>
<td>42.7</td>
<td>81</td>
<td>1745</td>
</tr>
</tbody>
</table>

$^aCQ = \text{Number of districts who completed the questionnaire.}$

$^bED = \text{Number of districts in the population estimated by the responding districts in the sample.}$

$^c% = \text{Percent usage of teacher evaluation models in district size categories based on the expansion weights presented in Table 2.}$
exist in districts with 2000 or more teachers and is reported at a 54.4 percent rate, whereas the lowest incidences of districts reporting this model were found in the 600-1999 size category at a 31.2 percent rate. Overall, the intermediate percentages ranged from 45.8 percent (120-249 teachers) to 43.1 percent (20-119 teachers) to 37.0 percent (250-599 teachers) and, finally, to 31.2 percent (600-1999 teachers). This information is available in Table 17.

**Goal-setting evaluation model by district size**

There are differences among the reported percentages in Table 17 as district size is compared with the goal-setting model of teacher performance evaluation. The percentages range from 22.0 percent for districts with 600-1999 teachers to 12.1 percent for districts with 2000 or more teachers. However, the remaining size strata depict a tighter range as they are reported at 19.7 percent, 19.4 percent and 16.7 percent for 120-249, 20-119 and 250-599 size categories, respectively.

**Clinical supervision evaluation model by district size**

The clinical supervision model as identified by district size reflects a frequency level which places the highest use (46.3 percent) in the 250-599 size class and the lowest use (31.9 percent) in the size class with 2000 or more teachers. Districts falling in the 600-1999 size class indicate use of the clinical supervision model at a 43.5 percent rate, which is similar to the highest reported rate. The two remaining stratifications, 20-119 and 120-249 teachers, report 34.6 percent and 33.5 percent rates, respectively (see Table 17).
Artistic and product evaluation models by district size

As noted earlier, these models were reported so infrequently they are treated jointly. The product model was only reported by districts in either the 20-119 size class or the 600-1999 size stratum and these incidences of occurrence were 0.8 percent and 9.9 percent, respectively (see Table 17). However, the artistic model was cited as being utilized in all but the 120-249 size stratum. The actual weighted percentages did not vary more than 0.5 percent and ranged from 1.6 percent to 2.1 percent usage among the four district size strata reporting use of the artistic model.

Research questions five and six

Data were analyzed to determine whether or not the training components provided to teachers in their evaluation process were independent of district size and type of evaluation model used. The specific research questions were outlined in Chapter I as null hypotheses one and two. For the purposes of Chapters IV and V, they will be listed as research questions (null hypotheses) five and six.

5. Teacher performance evaluation training provided to teachers is independent of teacher evaluation models.

6. Teacher performance evaluation training is independent of district size.

These research questions were answered using the test of proportionality described in Chapter III. The analyses of the raw data are listed in table form; however, it is necessary to note that the test procedure produced cells that contained five or fewer responses.
Therefore, it was necessary to collapse those cells which did not have sufficient cases. Even after this collapse was initiated, cells remained too small for statistical analyses using the test of proportionality. Ultimately, the evaluation models defined as product and artistic were eliminated from the statistical analysis because of only an actual response of two and eight, respectively. These infrequently occurring cases produced virtually empty cells and the test of proportionality could not have been employed; therefore, these models were removed from consideration.

Teacher evaluation training is independent of teacher evaluation models

Research question five was written to test whether the teacher performance evaluation training provided to teachers is independent of models of teacher evaluation. A test of proportionality using the PC CARP statistical procedure (described in Chapter III) was applied to the 473 weighted district questionnaire responses. The analyses of the weighted data are represented in Table 18 as estimated district counts and percentages.

The test of proportionality was computed on the collapsed distribution or on three models rather than five. In comparing the districts responding as clinical supervision, goal-setting and common law evaluation districts with the four training types (awareness, modeling, practice and feedback), it cannot be said that these variables are independent. Therefore, the null hypothesis, teacher performance evaluation training types provided to teachers are independent of evaluation models, is rejected at the .05 level of significance (p <
Table 18. The fifth research question: Is teacher performance evaluation training provided to teachers independent of teacher evaluation models\(^a\)

<table>
<thead>
<tr>
<th>Teacher evaluation models</th>
<th>Teacher evaluation training types used in districts</th>
<th>Awareness</th>
<th>Modeling</th>
<th>Practice</th>
<th>Feedback</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ED(^b)</td>
<td>%(^c)</td>
<td>ED(^b)</td>
<td>%(^c)</td>
<td>ED(^b)</td>
</tr>
<tr>
<td>Common law</td>
<td></td>
<td>1525.70</td>
<td>55.7</td>
<td>568.44</td>
<td>20.7</td>
<td>487.82</td>
</tr>
<tr>
<td>Goal setting</td>
<td></td>
<td>633.92</td>
<td>40.1</td>
<td>441.22</td>
<td>27.9</td>
<td>329.91</td>
</tr>
<tr>
<td>Clinical supervision</td>
<td></td>
<td>801.63</td>
<td>29.0</td>
<td>728.86</td>
<td>26.4</td>
<td>733.85</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>2961.20</td>
<td>1738.50</td>
<td>1551.60</td>
<td>832.75</td>
<td>7084.10</td>
</tr>
</tbody>
</table>

\(^a\) Test of proportionality: Calculated F = 2.19; Degrees of freedom: F\(_6, 319\); Significance level: .05; F table value: 2.10.

\(^b\) ED = Estimated number of districts in the population.

\(^c\) % = Percentage based on expansion weights.
A further review of Table 18 reveals the calculated F value as 2.19 and, when compared to the table F (6, 319) degrees of freedom at the .05 level or 2.10, the null hypothesis is rejected and the alternative hypothesis is accepted. Therefore, it can be said that under the conditions of this investigation teacher evaluation training types appear to be associated with certain models of evaluation.

**Training for teacher evaluation is independent of district size**

Research question six was postulated in order to test whether or not types of evaluation training provided to teachers were independent of district size. The test of proportionality was used to assess whether or not there was independence among these variables (see Table 19).

An inspection of Table 19 reveals the calculated F to be 0.88 and, when compared to the table F (8, 328) degrees of freedom at the .05 level (p < .05) or 1.94, the null hypothesis is retained. To repeat for clarity, it can be stated that teacher performance evaluation training types are independent of district size.
Table 19. The sixth research question: Is teacher performance evaluation training independent of district size.

<table>
<thead>
<tr>
<th>Number of teachers in the district</th>
<th>Teacher evaluation training types used in districts</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Awareness</td>
<td>Modeling</td>
<td>Practice</td>
<td>Feedback</td>
<td>Totals</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$ED^b$</td>
<td>$%^c$</td>
<td>$ED^b$</td>
<td>$%^c$</td>
<td>$ED^b$</td>
<td>$%^c$</td>
<td>$ED^b$</td>
<td>$%^c$</td>
<td>$ED^b$</td>
</tr>
<tr>
<td>20-119</td>
<td>2034</td>
<td>42.3</td>
<td>1017</td>
<td>21.2</td>
<td>1175</td>
<td>24.5</td>
<td>577</td>
<td>12.0</td>
<td>4803</td>
</tr>
<tr>
<td>120-249</td>
<td>555</td>
<td>39.9</td>
<td>362</td>
<td>26.0</td>
<td>336</td>
<td>24.1</td>
<td>139</td>
<td>10.0</td>
<td>1391</td>
</tr>
<tr>
<td>250-599</td>
<td>273</td>
<td>34.6</td>
<td>276</td>
<td>34.9</td>
<td>143</td>
<td>18.1</td>
<td>97</td>
<td>12.3</td>
<td>788</td>
</tr>
<tr>
<td>600-1999</td>
<td>80</td>
<td>29.7</td>
<td>77</td>
<td>28.6</td>
<td>53</td>
<td>19.7</td>
<td>59</td>
<td>21.9</td>
<td>270</td>
</tr>
<tr>
<td>$\geq$ 2000</td>
<td>20</td>
<td>37.7</td>
<td>15</td>
<td>28.4</td>
<td>12</td>
<td>22.4</td>
<td>6</td>
<td>11.5</td>
<td>54</td>
</tr>
<tr>
<td>Totals</td>
<td>2962</td>
<td>1747</td>
<td>1719</td>
<td></td>
<td>878</td>
<td>7306</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Test of proportionality: Calculated $F = .88$; Degrees of freedom: $F_{8,328}$, Significance level: $.05$; $F$ table value: $1.94$.

$^b$ $ED = $ Estimated number of districts in the population.

$^c$ $% = $ Percentage based on expansion weights.
CHAPTER V. SUMMARY, CONCLUSIONS, LIMITATIONS,
DISCUSSION AND RECOMMENDATIONS

Summary—Overview of the Study

This nationwide investigation served to provide a heretofore absent status report on the type, content and duration of training provided to public school teachers in support of their own evaluation systems. There were 700 large, medium and small districts divided into five appropriate size categories, all of which were drawn from a total population of 9,760 in 41 states and the District of Columbia. (States mandating evaluation procedures and criteria were not included in this study.) A total of 473 districts responded or 68 percent of the sample provided information on the 12-item questionnaire which gathered data on training status.

The questionnaire served to obtain information on type, content and duration of training for teacher performance evaluation. The Joyce and Showers (1988) staff development action-research training types (awareness, modeling, practice and feedback) were defined on the questionnaire as possible choices from which the sample districts could select. Essential teacher evaluation training curricula (interpersonal communications, effective teaching and data analysis/observation skills) proposed by Stiggins and Duke (1988) were listed and served as choices for the respondents. Finally, participants providing training in support of teacher evaluation systems were asked to indicate the approximate amount of training by identifying into which time category (one of five) this training duration would fit.
Also, sample districts were asked to characterize their evaluation system as one of five evaluation models. The McGreal (1983) taxonomy of teacher evaluation systems was used to ascertain the actual evaluation model utilized in the 473 responding school districts. Clinical supervision, goal-setting and common law were those models most frequently identified as being used in sample districts. The remaining two models, artistic and product, were cited very infrequently.

Tests of proportionality of two characteristics were used to determine independence among the variables studied rather than a chi-square procedure because a weighted, nonproportional sample was used to estimate the population. The hypotheses tested whether teacher performance evaluation training was independent of the type of evaluation model and district size. These two research questions were treated statistically and descriptive data were depicted in a number of tables.

Conclusions—Overview of Results

Two research questions were tested statistically with the findings reported in Chapter IV. In this same chapter, descriptive data served to answer four additional research questions. No interpretations were provided; however, further analyses were drawn from these results and are presented in the following list of conclusions. The first four questions were answered by interpreting frequencies. The final two questions were answered with conclusions obtained from the statistical treatment.

These conclusions do not include the artistic or product evaluation model. They were found so infrequently in the population that
conclusions about public school district use of them relative to the variables identified in this study cannot be made with any reasonable accuracy.

The first research question

What are the types of teacher evaluation training among public school districts using specified evaluation models?

Awareness training described by evaluation models

Awareness or basic orientation type training was utilized most frequently (55.7 percent) in districts using the common law model. Perhaps this is due to the legalistic, procedural nature of common law models which promote teacher evaluation orientation sessions intended to provide notice as required by procedural due process.

Modeling training described by evaluation models

The most frequently reported evaluation models using this level training were goal-setting (27.9 percent) and clinical supervision (26.4 percent) districts. This result may indicate that as more complex evaluation systems are used, the need to strengthen the level of training also increases.

Practice training described by evaluation models

Districts characterizing themselves as using the clinical supervision model more frequently (26.6 percent) employed the use of practice level training techniques in support of their teacher performance evaluation program. Practice training is more complex than either awareness or modeling, and it would be expected that the more sophisticated evaluation model would
necessitate such advance level training.

**Feedback training described by evaluation model**

Clinical supervision model districts used feedback training more often than districts using other models (18.0 percent as compared to goal-setting at 11.2 percent), apparently to support the collegial and collaborative nature of this evaluation model. It could be said as a result of the data accumulated in this study, that two-way interaction between supervisor and supervisee is expected and feedback training enables the full and quality utilization of such growth-precipitating exchanges.

**The second research question**

What is the content and duration of teacher evaluation training utilized in public school districts when categorized by models of teacher evaluation?

**Effective teaching content described by evaluation model**

Districts characterizing themselves as using the clinical supervision model had the highest percentage (79.7 percent) utilization of effective teaching content in the training provisions for teacher evaluation. This could be attributable to the two-way communication and common language expectations associated with the traditional clinical supervision model of evaluation.

**Interpersonal communications content described by evaluation model**

The use of interpersonal communication skill building activities for the support of teacher performance evaluation training was at the highest
level for clinical supervision districts (88.1 percent). The importance of the conference which is central to clinical supervision may account for the high use rate of interpersonal communications training content.

**Observation and analysis skills content described by evaluation model**

It is widely accepted that clinical supervision requires detailed involvement for both supervisor and supervisee. Therefore, a common knowledge base with regard to classroom observation and data analysis would be extremely productive and may account for highest (56.8 percent) use of this training content among clinical supervision districts.

**Duration of training described by evaluation model**

Of all the evaluation models under study in this investigation, clinical supervision and goal-setting districts reported their percentage amount of training to be greater than any other models (29.1 and 28.6, respectively, in the 15-25 hour category). However, clinical supervision districts report a good deal of training in the 26-40 hour category (14.8 percent). This summary finding suggests the conclusion that both goal-setting and clinical supervision districts devote more time to training teachers in evaluation systems. Furthermore, it can be concluded that the complexity of these models surpass the common law model with regard to the need of additional training time.

**The third research question**

What types, content and duration of evaluation training are most frequently used by public school districts of various sizes?
Effective teaching content described by district size

In the population under study, it appears that the larger the school district, the greater the incorporation of effective teaching content in training programs established for teachers in support of teacher evaluation systems. (From smallest to largest district, the percentages are 68.0, 73.3, 84.4, 85.2, 86.5.) The results might be due to greater availability and utilization of staff development resources in larger districts.

Interpersonal communications content described by district size

District size appears to have very little, if any, direct influence on the use of interpersonal communications training for teacher performance evaluation.

Observation and analysis skills content described by district size

The results of this study may suggest that larger districts have greater availability of training opportunities which include classroom observation data capturing and analysis techniques (2,000 or more teacher districts had 64.4 percent, while 20-119 teacher districts had 45.8 percent).

Awareness training described by district size

When districts in the population are divided into the five size strata, the amount of awareness training in one district size category (600-1999 teachers) employs a noticeably smaller rate of awareness training than the other districts. Otherwise, similar awareness utilization rates are noted across all size categories.
Modeling training described by district size  The size of the district appears not to have any direct influence on whether or not modeling training is utilized in the population under study.

Practice training described by district size  All districts report about the same utilization of practice training when categorized by district size. Number of teachers in the district appears not to influence the provision for practice level training in support of evaluation training.

Feedback training described by district size  The size of the district does not appear to influence whether or not feedback training is provided to teachers in support of their performance evaluation system.

Duration of training described by district size  A consistency across districts is seen when training duration is categorized by district size. In both small and large districts, the spread of training amounts across size categories represents a reasonably even distribution.

The fourth research question

What evaluation models are utilized most frequently by public school districts of various sizes?

Evaluation models described by district size  No obvious utilization pattern of evaluation models is observed when categorized by district size. However, models of evaluation do not appear to spread evenly among the various sizes of districts in the population under study. The differing frequency rates did not permit general conclusions regarding the use of evaluation models in certain size school districts
but did point out some specific considerations.

One conclusions might suggest that the common law model is used most frequently in the largest size districts because this model meets expediency and legal requirements, both given priority status in large, bureaucratic public school districts. Another conclusion from these findings might indicate that the medium size school districts more frequently employ the use of the clinical supervision model because they are large enough to allocate necessary resources to facilitate this complex model, yet not too large so as to lose sight of the vital communications linkage between and training for supervisee and supervisor.

The fifth research question

Is teacher performance evaluation training provided to teachers independent of teacher evaluation models?

Training independent of models Null hypothesis number one was rejected at the .05 significance level, providing a conclusion which explains that teacher performance evaluation training varies by evaluation model utilized in the district. Therefore, there is a relationship between type of training (awareness, modeling, practice and feedback) and clinical supervision, goal-setting and common law models of evaluation.

Such models were selected by participating district officials as characterizations of their evaluation system and when compared with choices made concerning training practices yielded a dependent
relationship. It appears that the type of training for teacher performance evaluation is dependent on evaluation models established in 41 United States and one District of Columbia public school district. The test of proportionality utilized in this study allows this dependency inference; however, it does not permit one to propose the specific nature or the direction of the relationship.

The sixth research question

Is teacher performance evaluation training independent of district size?

Training independent of district size Null hypothesis number two was retained. No relationship exists between teacher evaluation training and the number of teachers in a district. District size is not a factor in type of training.

Limitations of the Study

A number of limitations resulted as by-products of this research design. The limitations are provided in the listing which follows.

1. Participation in this study was voluntary. The responding districts might have been motivated to return questionnaires because of what they perceived to be adequate or better than average training provisions, whereas those not responding might have been in some measurable way different from those that did respond.

An attempt was made to determine such response bias by using the original questionnaire through a telephone interview which served to gather information from nonresponding sample districts. However, time
constraints resulting from the attempted explanation of the five models, four training types and three contents prevented an accurate appraisal of bias among the nonresponding districts. An abbreviated questionnaire should be constructed and validated for the express purpose of telephone interviewing.

2. The duration (how many hours of teacher performance evaluation training) question provided very limited, general information and did not identify a specific time frame during which training was provided, i.e., year, month, evaluation cycle, probationary contract period, etc.

3. The questionnaire was mailed to superintendents of schools who in turn might have assigned the completion task to another school administrator more directly involved in teacher evaluation. However, it is important to note that teachers were not asked to complete a survey.

4. Responding districts who might be using a combination of evaluation models were not able to indicate this information. The survey allowed districts to select one best choice when depicting their evaluation model.

5. Participating districts using more than one class of training in support of their teacher evaluation process were asked to select only the level that best described that practice.

6. Other than questions provided to establish demographics, the questionnaire forced respondents to either indicate the presence or absence of type and content of training responses or to select one of five evaluation models and training duration categories. It is quite possible that responses to a Likert-type instrument could have
established the degree to which selected training type and content were being utilized in participating districts. A similar case could be made for the use of an ordinal scale when asking sample districts to select the type of evaluation model used.

7. An identical transmittal letter was mailed to nonresponding districts with bright yellow attachment explaining the urgency of the request with a plea for the return of a completed survey. Such a procedure might have confused some of the districts receiving this "second notice" because the date on the transmittal letter was the same as the date listed in the initial mailing. More returns might have been received if a completely "new" transmittal letter had been produced.

8. Question number four asked: "What is the enrollment of your school district?" This language appeared to be confusing and the data supplied from the responding districts were not consistent with the data from the preceding ("correctly" phrased) question. Therefore, it was necessary to access all the census data from the United States Office of Education common core tape which limited the size results to 1987/88 school year rather than current year information.

Discussion of the Study

The statistical findings in this investigation were consistent with intuitive expectations in place at the outset of the study. This is particularly true with regard to the relationship of training types on models of teacher performance evaluation. Previous researchers have established the need for mutual trust and a collegial, collaborative
relationship between evaluator and evaluatee. These enabling conditions do much to assure successful teacher evaluation that is ultimately intended to promote accountability and improvement goals (Darling-Hammond et al., 1983). The data in this study strongly suggest these prerequisite success factors as it demonstrates the existence of a dependent relationship between evaluation models and training. Specifically, when the frequency data are viewed, it is evident that clinical supervision and goal-setting districts provide higher levels of teacher evaluation training which contains pertinent content and inevitably promotes requisite enabling conditions.

Also, it was expected that size of district would have little influence on the training provisions offered to teachers in support of teacher evaluation. The statistical procedure applied to test this research question established this fact under the conditions and within the population. This is not alarming, in light of the fact that successful teacher evaluation is predicated on organizational, contextual issues and, most importantly, trust and joint ownership, all of which can and do exist in any size district.

The descriptive information revealed in this national study has placed into perspective the utilization of those models believed to be most widely implemented in United States public schools. It also points to the fact that the artistic and product evaluation models are infrequently cited by United States public school officials as models of evaluation. Furthermore, it was shown that awareness (orientation) training levels were widely noted as the class of training most often
associated with teacher performance evaluation. This alarming fact fortifies the practitioner's and researcher's call for enhanced training provisions for teachers in teacher evaluation (McGreal, 1983; McLaughlin & Pfeifer, 1986).

The information gathered with regard to content of training for teacher performance evaluation was quite revealing. It was not surprising to find that interpersonal communication and effective teaching skills were rated as the number one and two most common content (respectively) associated with teacher evaluation training content. This fact, when considered in the context of the clinical supervision model, parallels recent research finding by Roettger (1990) for evaluation in two-year colleges.

The data accumulated from an extensive literature review clearly established the need to increase the knowledge base of the evaluator. However, data presented as findings and conclusions strongly suggest that the same be done for the "consumers" of evaluation, the teachers. This status report not only points to the "low" level and amount of training provided to teachers in evaluation models, but it supports the appeal by the researchers that additional training comprised of essential content must be provided if evaluation is to serve accountability and improvement objectives simultaneously.

Continued investigation along these lines will, hopefully, prevent the false assumption that accountability and improvement are incompatible. For if this canard is allowed to persist and perpetuate, teacher performance evaluation will be cast in the role of laggard in the
emerging scheme to restructure and improve schools. This study has been one of many initial steps toward securing necessary research support to ultimately assure a reciprocal understanding between evaluator and evaluatee of effective teaching methods, quality two-way communications and the observation, collection and analyses of classroom/artifact data, all essential before accountability and improvement measures can unite to form an instrument capable of playing a starring role in efforts aimed at school improvement and restructuring.

With improved utilization of what is known about adult learning and professional development, such teacher evaluation content will be transferred to the "executive control" of teachers and administrators and no longer will it be characterized as weak and vulnerable. Adequate training in essential elements will produce teacher performance evaluation that is valid, reliable and serving both accountability and growth goals.

Recommendations for Practice

The research is replete with calls for improved training provisions for teachers in their own evaluation system. Also, the better than expected response rate to the survey used in this study indicates a good deal of interest in training for teacher performance evaluation among practitioners. The results of practitioner responses to the data gathering instrument employed in this investigation has prompted the following recommendations.

1. Teacher evaluation training, for the most part, is conducted at
the basic stages and for transfer to take place implementation should occur at higher levels such as modeling, practice and feedback training levels.

2. Training in effective teaching, communications and classroom observation and data analysis provides a common knowledge level among evaluator and evaluatee and should be included as a portion of all professional development activities conducted in school districts.

3. Districts should identify the type of evaluation model utilized and plan training activities to support the intended goals of their individual evaluation system.

4. Aggregate summative data should be organized by school district and building to supply valuable staff development needs assessment information. This information could be ultimately used to help facilitate and focus staff development efforts.

5. Equal levels of performance evaluation skills between evaluator and evaluatee should be encouraged. Such equality of professional development will increase the likelihood that teachers will become agents of their growth. Ultimately, this will decrease undesirable dependency on the administrator as evaluator and elevate their role to one of a professional development coach.

6. Effective teacher evaluation systems and supporting training programs should be encouraged for all evaluation models, at a minimum, to assure teachers that their property and due process rights are protected.

7. Evaluation models that involve teachers in the same training as provided administrators should be the rule rather than the exception.
Such provisions will promote positive trust and help foster a climate of growth.

8. School districts, regardless of size, should provide necessary resources, both human and financial, to support training activities for teacher evaluation.

Recommendations for Further Research

As a result of data obtained in this study and subsequent analyses, certain focalized and data-supported research speculations have emerged. Such teacher performance evaluation training research recommendations are suggested and listed below:

1. A study should be conducted to determine whether or not higher training levels in pertinent teacher performance evaluation content (as described in this study) produce increased acquisition of those skills listed as essential content.

2. A follow-up case study of responding districts should be conducted to learn actual evaluation practice (particularly in clinical supervision, goal-setting and common law model districts) and compare these case study findings with those reported on the initial questionnaire.

3. An investigation should be completed with the same sample districts to compare teacher perceptions of training provisions offered in support of their evaluation model with those reported by the responding district officials.

4. A comparison of nationwide training provisions for teachers and
administrators in the acquisition of skills associated with teacher performance evaluation should be conducted.

5. An investigation should be conducted to learn of teacher performance evaluation training commonalities and differences between the 41 states and District of Columbia public school districts surveyed in this study and those remaining districts found in the nine states which mandate teacher performance evaluation procedures, criteria and training.

Suggestions for Replication and Refinement

This nationwide investigation utilized a complex methodology and produced results which researchers may wish to replicate. Therefore, several tips are provided to assist future investigators in such endeavors.

1. A more detailed questionnaire using the population parameters employed in this study would produce a more complete understanding of training provisions for teachers in support of their own evaluation procedures.

The duration question should provide the respondents with a specific time frame, i.e., year, month, week, evaluation cycle, etc., in which training is expected to take place. Districts should be allowed to characterize whether or not they use a combination of evaluation models and/or training types. Also, a data entry form which has more ample response fields should be utilized. Additional data fields would allow the use of only one district size question rather than two; a split question and response requirement should not be utilized.
2. A questionnaire employing the use of ordinal or Likert-type ratings could be designed to determine the degree to which districts implemented various training content and type, thus allowing for correlation and, possibly, predictions. Also, additional inquiries regarding district demographics should be included and precise language determined for each item to assure meaningful and accurate response.

3. A post card should be mailed following receipt of the second returns with a note encouraging the return of the questionnaire. Also, a separate transmittal letter should be developed for the follow-up mailing to prevent confusion.
BIBLIOGRAPHY


Wiese, M. J. & Maxwell, G. L. (1939). As we are taught, so do we teach. *Journal of Adult Learning, 11*(2), 174-175.


ACKNOWLEDGMENTS

The realization of this exceptional career development experience came with much sacrifice, especially from my loving, supportive and extremely tolerant family. This study is humbly dedicated to Mary, Joey, Mariana and Gabe, from whom I gained strength and resolve to complete this undertaking. Without their unceasing encouragement and consistent support, completion of this titanic project would not have been realized.

I wish to gratefully acknowledge and sincerely thank my major professor, Dr. Richard P. Manatt; without his coaching and enduring belief in my ability, I would never have embarked on this prodigious adventure. My association with Dr. Manatt has not only contributed extensively to my professional growth, but also served to install me as a member of the School Improvement Model project team, a family of highly dedicated professionals, many of whom have become valued friends. Two of whom especially assisted me through this memorable journey and unselfishly offered time and expertise to lighten my load for easier passage; for this generosity, I offer sincere, special thanks.

I wish to thank the other members of my dissertation and program of study committee—Dr. Shirley Stow, Dr. Tony Netusil, Dr. Jackie Mitchell and Dr. Detroy Green—for their advice and scholarly counsel throughout this ambitious project. It is not often one has the extraordinary opportunity to capture the undivided attention and energetic guidance of such an expert group of education professionals.
December 29, 1989

Dear Superintendent of Schools or Designated School Official:

The School Improvement Model (SIM) Projects Team of Iowa State University has been involved in teacher performance evaluation research for the past decade. Among the many SIM research projects, nation-wide training provisions for teacher performance evaluation ranks as one of its priorities. We are currently conducting a national study of training status for teacher performance evaluation and need your assistance.

Your organization is among 700 school districts sampled from the more than 15,550 public districts nation-wide. Your district is a representative from a specific size stratification and completion of this brief, 12 item questionnaire is vital for the success of the study. Please, if you do not intend to personally complete the questionnaire, use the enclosed routing form to forward it to an individual responsible for teacher evaluation in your district. With your help, the results of our research will prove to be extremely valuable in describing current national training provisions for teacher performance evaluation.

You can be assured of the anonymity of your responses to the enclosed questionnaire. No individual responses will be reported in the findings. The enclosed Data Sheet III (blue "bubble") response form has been coded in order to assure geographic representation and to permit follow-up mailings of conclusions, should this be your wish. You are asked to indicate whether or not you would like a report of findings on item 10 of the questionnaire.

Please return the completed Data Sheet III as soon as possible but no later than January 19, 1990. We have enclosed a pre-addressed, postage-paid business reply envelope for the Data Sheet III (blue "bubble") response form. You may keep the actual questionnaire. Our primary concern is receiving your responses listed on the completed Data Sheet III.

Thank you for your time and assistance. If you have questions, please telephone during school hours at (office telephones) [515] 227-3217/674-4111 or call after hours at (home telephone) [515] 233-6722. Your assistance is greatly appreciated.

Sincerely,

Richard P. Manatt, Professor and Director

Joseph M. Petrone, Doctoral Candidate
TRAINING FOR TEACHER PERFORMANCE EVALUATION
QUESTIONNAIRE

A Note to the Respondents

Your school organization has been selected as a participant in a nation-wide research project about training provisions offered to public school teachers in their performance evaluation processes.

For the purpose of this questionnaire, teacher performance evaluation training is defined as: learning activities provided to teachers for increased understanding of principles, acquisition of skills and application of behaviors identified as teacher performance criteria in a district's teacher evaluation process. The ultimate objective of such teacher performance evaluation training is the attainment of the highest quality teacher performance possible.

MARK ON THE ENCLOSED DATA SHEET III (BLUE "BUBBLE" RESPONSE FORM) USING A NO. 2 PENCIL. NO IDENTIFYING INFORMATION IS REQUIRED. THE RESPONSE SHEET HAS BEEN CODED FOR YOUR CONVENIENCE.

An abstract of the complete report will be forwarded if requested on item 10 of the questionnaire.

1. Does your school district/organization currently utilize a formal (district-wide) teacher performance evaluation process?
   A. Yes
   B. No [If No, complete only question 10 and follow the mailing instructions provided on the last page of the questionnaire].

2. What title best describes your position?
   A. Superintendent
   B. Assistant Superintendent
   C. Principal
   D. Director
   E. Coordinator/Facilitator

   PLEASE ANSWER BOTH ITEMS THREE (3) AND FOUR (4)

3. What size category most accurately represents the number of full-time classroom teachers employed in your school district during the 1989/90 school year?
   A. Under 75 teachers
   B. 76 - 120 teachers
   C. 121 - 180 teachers
   D. 181 - 250 teachers
   E. More than 250 teachers [See question number four (4).]

4. What is the enrollment of your school district?
   A. Less than 250 teachers
   B. 251 - 425 teachers
   C. 426 - 600 teachers
   D. 601 - 1,200 teachers
   E. More than 1,201 teachers
In his book *Successful Teacher Evaluation*, Thomas McGreal provides a taxonomy for the classification of evaluation models and procedures. The five statements that follow are representative explanations of the models and procedures. Select the one best statement that describes your evaluation model/process and mark the corresponding letter on the response sheet.

5. A. ___ The model is characterized by high supervisor - low teacher involvement; evaluation is synonymous with observation; major emphasis is on summative evaluation; standardized criteria; and comparative judgments. This process usually relies on definitions, procedures, and processes that are traditional.

B. ___ This model is characterized by an emphasis on an individualized approach to evaluation. Instructors and evaluators meet and confer to set and monitor goals. Generally, no checklist of criteria is used. Self-evaluation may be a component of this model.

C. ___ The model is characterized by evaluation that is based on the results or outcomes of student achievement tests or on competency-based evaluations, but not on methods, styles, or processes. Generally, the instruments for assessing student growth are norm-referenced and criterion-referenced tests.

D. ___ The model is characterized by a close relationship between the instructor and the supervisor with emphasis on collegial rather than authoritarian orientation. It takes its principal data from the classroom and is designed to improve instructor’s performance.

E. ___ The model is characterized by a belief that teaching is an art, that the quality of the performance the instructor exhibits is likened to an aesthetic experience. The evaluation is more subjective and, perhaps, less precise.

In their book *Case for Commitment to Teacher Growth Research on Teacher Evaluation*, Richard J. Stiggins and Daniel Duke identify elements of training for teacher evaluation. The three statements that follow explain the content identified. Should more than one of the descriptions which follow reflect the teacher performance evaluation training currently conducted in your district, mark the item’s corresponding letter which represents a description of your district’s practice.

6. A. ___ The training content in our district contains effective teaching methodology based on research or derived from a coherent set of values and aspirations.

B. Does not apply.

7. A. ___ The training content in our district contains effective, two-way communication founded on a clear set of descriptive terms, good speaking and listening, and interpersonal relations.

B. Does not apply.

8. A. ___ The training content in our district contains technical skills in the collection and analyses of descriptive data on teaching.

B. Does not apply.

9. Please mark the corresponding letter on the response sheet of the statement most representative of your district.

A. ___ At least one (1) of the above teacher performance evaluation content areas is currently utilized to support our teacher performance evaluation process.

B. ___ None of the above content descriptions represent training provided to teachers in support of their teacher performance evaluation process.

C. ___ No training is provided to teachers in support of their teacher performance evaluation process. [If C, SEE INSTRUCTIONS FOLLOWING ITEM 10.]
10. Do you wish results of this national study?
A. Yes
B. No

11. In their book *Student Achievement Through Staff Development*, Bruce Joyce and Beverly Showers identify several training components. The four statements that follow are performance evaluation-specific representations of the components identified. Select the one statement that best describes the type of teacher performance evaluation training process consistently used in your district and mark the corresponding letter on the response sheet.

   A. An exploration of the district's evaluation theory and practice through discussions, readings, lectures, etc. provides an understanding of the rationale behind teacher performance criteria and the principles/procedures that govern its use.

   B. The demonstration or modeling of skills associated with the district's teacher performance criteria are demonstrated in settings that simulate the workplace, mediated through film or videotape, or conducted live in the training setting.

   C. The practice of skills associated with the district's teacher performance criteria which enables the trainee to profit from one another's ideas and skills. Opportunities are provided for small group interaction following peer teaching and practice conducted in classroom settings.

   D. Feedback about each other's achievement in skill areas associated with the district's teacher performance criteria is shared among teacher trainees and, utilizing audio, video recording or direct classroom observation, critiques are offered once they have a clear idea of the skills and how to use them.

12. Please select the time category that best represents the duration of teacher performance evaluation training provided to teachers in your district and mark the corresponding letter on the response sheet.
   A. Less than 15 hours
   B. 15 - 25 hours
   C. 26 - 40 hours
   D. 41 - 55 hours
   E. More than 55 hours

PLEASE RETURN ONLY YOUR COMPLETED DATA SHEET III (BLUE “BUBBLE” RESPONSE FORM). THERE IS NO NEED TO RETURN THE ACTUAL QUESTIONNAIRE. A SELF-ADDRESSED, POSTAGE PAID BUSINESS REPLY ENVELOPE HAS BEEN PROVIDED FOR YOUR CONVENIENCE.

SHOULD YOU HAVE QUESTIONS, PLEASE CALL:

JOE PETRONE, RESEARCH ASSOCIATE
SCHOOL IMPROVEMENT MODEL PROJECTS
IOWA STATE UNIVERSITY
OFFICE TELEPHONES (SCHOOL HOURS)
   515 227-3217
   515 674-4111
HOME TELEPHONE (EVENING)
   515 233-6722

THANK YOU. YOUR TIME AND ASSISTANCE IS GREATLY APPRECIATED.
APPENDIX B. HUMAN SUBJECTS APPROVAL FORM
INFORMATION ON THE USE OF HUMAN SUBJECTS IN RESEARCH
IOWA STATE UNIVERSITY
(Please follow the accompanying instructions for completing this form.)

1. Title of project (please type): Nation-wide survey of teacher training in support of teacher evaluation processes

2. I agree to provide the proper surveillance of this project to insure that the rights and welfare of the human subjects are properly protected. Additions to or changes in procedures affecting the subjects after the project has been approved will be submitted to the committee for review.

   Joseph M. Petrone 11/30/89
   Typed Name of Principal Investigator

214 16 Street, Ames, Iowa 50010
   Campus Address

   233-6722
   Campus Telephone

3. Signature of others (if any) Date Relationship to Principal Investigator

   Richard Manatt 11/30/89 Major Professor

5. ATTACH an additional page(s) (A) describing your proposed research and (B) the subjects to be used, (C) indicating any risks or discomforts to the subjects, and (D) covering any topics checked below. CHECK all boxes applicable.

   □ Medical clearance necessary before subjects can participate
   □ Samples (blood, tissue, etc.) from subjects
   □ Administration of substances (food, drugs, etc.) to subjects
   □ Physical exercise or conditioning for subjects
   □ Deception of subjects
   □ Subjects under 14 years of age and (or) □ Subjects 14-17 years of age
   □ Subjects in institutions
   □ Research must be approved by another institution or agency

6. ATTACH an example of the material to be used to obtain informed consent and CHECK which type will be used.

   □ Signed informed consent will be obtained
   □ Modified informed consent will be obtained

   Anticipated date on which subjects will be first contacted: 01 05 1990
   Anticipated date for last contact with subjects: 03 02 1990

7. If Applicable: Anticipated date on which audio or visual tapes will be erased and (or) identifiers will be removed from completed survey instruments: 03 02 1990

8. Signature of Head or Chairperson Date Department or Administrative Unit

   11/30/89 Prof. Studies

9. Decision of the University Committee on the Use of Human Subjects in Research:

   □ Project approved □ Project not approved □ No action required

   George G. Karas 12/7/89
   Name of Committee Chairperson Date

   Signature of Committee Chairperson

   Principal investigator will be using identification code for matching purposes. - Per Joseph M. Petrone 12/11/89