

1988

The effects of performance appraisal training on the skills and confidence levels of teacher evaluators and trainers of teacher evaluators

Leonard Anthony McIntyre
Iowa State University

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

 Part of the [Educational Administration and Supervision Commons](#), and the [Teacher Education and Professional Development Commons](#)

Recommended Citation

McIntyre, Leonard Anthony, "The effects of performance appraisal training on the skills and confidence levels of teacher evaluators and trainers of teacher evaluators " (1988). *Retrospective Theses and Dissertations*. 8871.
<https://lib.dr.iastate.edu/rtd/8871>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

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. These are also available as one exposure on a standard 35mm slide or as a 17" x 23" black and white photographic print for an additional charge.

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.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600



Order Number 8909173

The effects of performance appraisal training on the skills and confidence levels of teacher evaluators and trainers of teacher evaluators

McIntyre, Leonard Anthony, Ph.D.

Iowa State University, 1988

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106



The effects of performance appraisal training on the
skills and confidence levels of teacher evaluators and
trainers of teacher evaluators

by

Leonard Anthony McIntyre

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

1988

TABLE OF CONTENTS

	PAGE
CHAPTER I. INTRODUCTION.....	1
Statement of the Problem.....	5
Purpose of the Study.....	6
Research Hypotheses.....	7
Basic Assumptions.....	7
Delimitations of the Study.....	8
Definitions of Terms.....	9
Human Subjects Release.....	10
CHAPTER II. REVIEW OF LITERATURE.....	11
Introduction.....	11
Teacher Evaluation.....	11
Evaluation Training.....	13
Confidence Level of Administrators.....	16
Lesson Analysis/Observation and Data Gathering... ..	18
Conferencing and Feedback Skills.....	20
Personnel Decisions.....	24
Personnel Decisions Connected to Teacher Eval.	24
Pay for Performance.....	27
Tenure/Dismissal.....	29
Supervisor-Teacher Relationships.....	32
Summary of the Chapter.....	34
CHAPTER III. METHODS AND PROCEDURES.....	35
Collection of Study Data.....	36

Research Design.....	36
The Sample.....	37
Materials Development.....	39
Video-Tapes.....	40
Instrumentation.....	41
Procedures.....	43
Analysis of Data.....	45
CHAPTER IV. FINDINGS.....	47
Descriptive Data.....	48
Sample.....	48
Interrater Reliability.....	48
Confidence Level.....	55
Inferential Statistics.....	62
Hypotheses.....	62
Hypotheses Testing.....	62
CHAPTER V. DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS.....	68
Discussion and Conclusions.....	68
Findings.....	69
Other Findings.....	72
Limitations.....	75
Recommendations for Further Research.....	75
SELECTED BIBLIOGRAPHY.....	77
ACKNOWLEDGMENTS.....	84
APPENDIX A - INSTRUCTIONAL PLANS.....	86
APPENDIX B - HUMAN SUBJECTS MODIFIED INFORMED CONSENT STATEMENT.....	93
APPENDIX C - TPE RATING SCALE.....	95
APPENDIX D - THE SUPERVISORY ATTITUDE SURVEY.....	97

LIST OF TABLES

	PAGE
TABLE 1. Distribution of sample by job classification.....	38
TABLE 2. Reliability distribution of composites of the posttest ratings of confidence levels as teacher evaluators in selected components.....	44
TABLE 3. Reliability distribution of composites of posttest ratings of confidence levels as trainer of teacher evaluators in selected components.....	44
TABLE 4. Distribution of the standard deviations of pretest posttest ratings (1=low to 3=high) of selected skills as teacher evaluators.....	52
TABLE 5. Frequency of responses of evaluators' ratings for each teaching skill.....	53
TABLE 6. Pretest/posttest ratings of confidence level as teacher evaluators for selected components.....	57
TABLE 7. Pretest/posttest ratings of confidence levels as trainers of teacher evaluators for selected components.....	60
TABLE 8. Analyses of pretest/posttest ratings of confidence level as teacher evaluators in selected components.....	65
TABLE 9. Analyses of pretest/posttest ratings of confidence level as trainers of teacher evaluators in selected components.....	66

CHAPTER I. INTRODUCTION

The emerging knowledge base from the recent research on effective teaching linked to student achievement will challenge administrators to develop their skills in teacher evaluation in order to achieve optimal student learning. Stow and Sweeney (1981) maintain that to be successful an evaluation system must include the development of evaluators' skills for assessing teacher performance. They reported that school districts across the nation have discovered that teacher performance evaluations are "the essential building blocks of accountability" (p.539). Researchers overall agree that student achievement is greatly influenced by the techniques and strategies used by teachers (Anderson et al., 1979; Good & Grows, 1979). Among educators there is a general consensus as to the merits and justification of teacher evaluation, (McGreal, 1983); moreover, educators are basically in accord with Bolton (1973) regarding its general purpose: "to safeguard and improve the quality of instruction received by students" (p. 27). Consequently, teacher evaluation is most important for those administrators whose objective is to improve student achievement; however, there is some concern relative to administrators "possessing the professional skills necessary to participate in the evaluation

process" (Kowalski, 1978). Manatt (1982) wrote, "Teacher Performance Evaluation (TPE) is a skill (or series of skills) and like skiing, tennis...TPE can be enhanced by training" (p. 2).

Teacher evaluation can be subjected to a task analysis and divided into various skills. The skills to be acquired include observation and analysis skills, data gathering skills, conferencing and feedback skills as well as those skills associated with teaching and coaching.

Research indicates that observation skills and data gathering comprise one of the major components of teacher evaluation training. Researchers support the idea that classroom observation skills can be enhanced by training (McGreal, 1983; Medley et al., 1984; and Wise et al., 1984). Bolton included, in his five steps for evaluation, a plan for data acquisition through evaluator training. According to McGreal (1983), the major factors in the success and effectiveness of the teacher evaluation system depend on the quality of observations and the way supervisors collect and share data with teachers. Certainly, the collection and analysis of specific data gathered during classroom observation is the foundation for improving teacher performance; providing feedback derived from classroom observation

allows teachers and supervisors to analyze teaching strategies and improve performance (Borg & Gall, 1983; Latham et al., 1975).

Another major component of teacher evaluation training is conferencing and providing feedback. Supervisors view the post-observation conference as an important and effective means for improving teacher performance in spite of little evidence in research to support this. The failure of the post-observation conference to improve teacher performance is attributed to the supervisor's lack of ability to effectively conduct the conference (Blumberg, 1970). Brandt (1982) and Olivero (1982) found that supervisors rated one of the highest needs for professional development, the improvement of their conference skills.

Some of the knowledge and skills involved in conducting effective conferences was supported by the literature. Maier (1976) found that successful supervisory conferences shared a common structure, i.e., a good opening, a body, and a good closing. Blumberg (1970) and Sullivan & Walker (1981) noted that the climate of the conference was important to conference success. Lefton et al. (1981) deemed effective probing and questioning important. Hunter (1982) suggested teachers must be involved and encouraged, and that

feedback and reinforcement are necessary in any learning endeavor, such as the supervisory conference. She (1980) further stated, "The principles of learning apply to teachers as apply to students." The originator and driving force behind the performance objectives approach, Redfern (1980), indicated that the most useful personnel evaluation program will be based on the concept that evaluatee and evaluator jointly establish work objectives, agree on well established action plans, and measure accomplishment in terms of outcomes and results. Meyer et al. (1965) found in the private sector that subordinate participation in setting improvement goals was more likely to improve performance than the setting of improvement goals by the supervisor without consulting with the subordinate.

Many situations when first encountered are likely to create apprehension and hesitancy particularly on the part of those who lack the necessary confidence to perform. For teacher supervisors, conducting teacher performance evaluations (TPEs) typically falls into this category. The relationship between confidence and performance is an important one. According to Saunders (1984) self-confidence is basic to success. It is understandable then that the untrained evaluator approaches the evaluation process with apprehension and

hesitancy. Bandura (1978, p. 141) stated that efficacy expectations "determine how much effort people will expend and how long they will persist in the face of obstacles." The stronger the perceived efficacy, the more active the efforts to persist. Feltz and Mugno (1983) found that a change in self-efficacy, as a result of training, improved performance levels of trainees. It follows, then, if administrators are provided proper training in teacher performance evaluation, the confidence level of the administrator will increase and they will be more willing to become involved and persist in this most important supervisory activity. Confidence, then, seems worthy of inclusion in training and research. As Rice (1986) suggested, "it (confidence) appears to influence the amount of energy we allocate to an activity, the extent which we persist in the activity, and our performance in the activity."

Statement of the Problem

It is apparent that more work needs to be done to explicate the relationships between the skills and purposes that make up the major components of teacher evaluation and the effects of teacher evaluation training on each. The skills comprise observation and analysis skills, data gathering skills, conferencing and feedback

skills, as well as those skills associated with teaching and coaching. The purposes include making personnel decisions, developing and maintaining supervisor-teacher relationships, and improving instruction. While it is understandable that the untrained supervisor approaches teacher evaluation with apprehension and anxiety, we don't know whether training will result in the increase level of confidence that administrators need to improve performance as teacher evaluators and trainers of teacher evaluators.

In truth, we know little about the effects of training on these components or on the other elements which comprise supervisor training. We know little about how training affects those being trained or the factors which interact with training. It is this problem of assessing the effects of training that is addressed by this study.

Purpose of the Study

The expressed purpose of the study was to assess the effects of training on the skills and confidence level of teacher evaluators and trainers of teacher evaluators.

Below are the questions which strike at the heart of the investigation:

- a. Are the skills of teacher evaluators enhanced by training?

- b. Is the confidence level of teacher evaluators enhanced by training?
- c. Will there be a difference in the confidence level needed to become a trainer of teacher evaluators based on training?

Research Hypotheses

This study was designed to gather data to test the following hypotheses:

1. There will be significantly higher posttest ratings of confidence level of teacher evaluators for selected components after training.
2. There will be significantly higher posttest ratings of confidence level needed to become a trainer of teacher evaluators for selected components after training.

Basic Assumptions

The study was predicated on the following basic assumptions:

1. The instruments, survey procedures, and data collection method used in the study are reliable and valid.
2. Respondents to the assessment instruments will reply honestly.

3. The participants are knowledgeable and skilled and extreme differences exist.
4. Improved skills and confidence level of teacher evaluators and trainers of teacher evaluators should lead to improved teacher evaluator performance.
5. Improved teacher performance evaluation in the classroom should lead to improved instruction.
6. Improved instruction in the classroom has a positive effect on student learning.
7. A task analysis of the training content and method will result in better training.
8. The opportunity to practice and demonstrate the newly acquired training skills and confidence level will improve evaluation skills.

Delimitations of the Study

The following factors limited the scope of the investigation.

1. The study was conducted with a limited number of administrators and supervisors primarily from the state of Iowa and may have had similar goals and expectations.
2. Subjects analyzed a taped lesson segment of only one grade level, in one subject area, using one particular teaching approach.

Definition of Terms

The following definitions of terms give clarity to their use and meaning in this study:

1. Teacher Skills - those 22 behaviors that evaluators observed and rated during lesson observation and analysis.

2. Evaluator Skills - data gathering, lesson analysis, lesson analysis and observation, and use data for feedback.

3. Selected Components - data gathering, lesson analysis, lesson analysis and observation, and use data for feedback, improving instruction, compensation pay decisions, tenure decisions, and maintaining a positive supervisory relationship.

4. Interrater Reliability - the degree to which evaluators agree on their ratings for each of the 22 teacher skills.

5. Participants - 64 administrators/supervisors (trainees) receiving the I-LEAD Train-the-Trainers' evaluation training.

6. Evaluator - participants whose role is to appraise teacher performance.

7. Trainers of evaluators - participants who will later train teacher evaluators.

8. Confidence - the perception of competence the participants have in their ability in each of the eight components.

Human Subjects Release

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

CHAPTER II. REVIEW OF LITERATURE

Introduction

This study investigated the effects of training on the skills, and confidence levels of teacher evaluators and trainers of teacher evaluators. This review is limited to literature essential to the study, addressing subtopics related to performance evaluation training, confidence, lesson analysis and observation skills, conferencing and feedback, personnel decisions, and teacher supervisor-teacher relationships.

Teacher Evaluation

Over a dozen reasonably distinct purposes for teacher evaluation have been suggested, such as improving teacher performance, aiding administrative decisions, guiding students in course selections, meeting state and institutional mandates, promoting research on teaching, and the like (Millman, 1981). Among educators there is a general consensus as to the merits and justification of teacher evaluation (McGreal, 1983). Basically, researchers agree that student achievement is greatly influenced by the techniques and strategies used by teachers (Anderson, Evertson, & Brophy, 1979; Good & Grows, 1979). In McGreal (1980, p. vii), Bolton

suggested that educators overall are in accord regarding the general purpose of teacher evaluation: "to safeguard and improve the quality of instruction received by students." He listed the following specific functions of teacher evaluation as the means for fulfilling this major purpose:

1. To improve teaching through the identification of ways to change teaching systems, teaching environments, or teaching behaviors;

2. To supply information that will lead to the modification of assignments, such as placements in other positions, promotions, and terminations;

3. To protect students from incompetence, and teachers from unprofessional administrators;

4. To reward superior performance;

5. To validate the school system's teacher selection process;

6. To provide a basis for teachers' career planning and professional development.

Consequently, teacher evaluation is most important for those administrators whose objective is to improve student achievement; however, there is concern as to whether or not administrators "possess the professional skills necessary to participate in the evaluation process" (Kowalski, 1978). This is supported by McGreal

(1983) who stated that in most instances, the difficulties arise not with the concept or the general purposes, but from the way evaluation is conducted. Weber (1987, p. 11) found, "Nearly everyone agrees that the ultimate aim of teacher evaluations is to create competent, effective teachers who will improve student performance. But the road toward this goal is strewn with controversies." Stow and Sweeney (1981) maintain that to be successful an evaluation system must include the development of evaluators' skills for assessing teacher performance. McGreal (1980) noted that one of the major difficulties associated with developing effective teacher evaluation systems lies with the general lack of training of teachers and supervisors in the evaluation process.

Evaluation Training

Beach (1980, p. 358) defined training as "the organized procedure by which people learn knowledge and/or skills for a definite purpose, such as to aid in the achievement or organizational goals." McGreal (1983) recommended that the school district provide all the members of the school with appropriate training and guided practice in the skills and knowledge necessary to implement and effectively maintain or increase the

effectiveness of its teacher evaluation system. Faast and Stow (1984) reported that many evaluators feel poorly prepared to do a suitable job of teacher evaluation. Therefore, it is necessary for school districts to offer evaluator training as a part of an on going in-service program. According to Barth (1980), supervisors often leave the university setting, the sight of their preservice preparation, with insufficient skills. He further stated that the ineffectiveness of university preparation was due, in part, to the fact that fledgling supervisors were often unsure of what they would face in the field until the actual situations or problems presented themselves. Consequently, it seems that much of the actual training of supervisors must occur in the field. Therefore, teacher evaluation programs will not work without effective methods for training supervisors in the field.

Experts in the field identify a number of components in which administrators must be trained and provide some direction as to how the training should be conducted. The need for administrative training, according to Streifer (1987), comes in three areas: knowledge of the teacher effectiveness literature, data collection/evaluation techniques, and conference techniques. Faast and Stow (1984, p. 130) stated that

"each evaluator should be able to identify effective teaching behaviors, be proficient in lesson plan analysis, gather descriptive data in classroom observation, conduct an evaluation conference with coaching and feedback, and complete the summative evaluation report which adequately rates the teacher on the defined criteria." They contend that training must address itself to the immediate application of skills if it is to accomplish these tasks. Nance's (1986) study showed that trainees analyzed lesson plans more effectively, captured data more accurately, and conducted better conferences after evaluation training.

Wickert (1987) concluded that in order to be fair to the teachers that administrators evaluate, they should be consistent in their practices by establishing a required level of administrator competency within a single district.

Training also is needed to reduce rating errors. The lack of reliability in the observation of behavior can be largely attributed to well-known rating errors. "Rating errors are errors in judgement that can occur when one individual observes another." It would seem logical that to solve the problem of rating errors, the observer must be trained (Latham et al., 1975). Bayroff et al. (1954) three decades ago stated that raters are

seldom skilled in making systematic work-related behavior observations. They need to become adept at observing and recording relevant job behaviors so they may be better equipped with the information necessary for making accurate evaluations of teacher performance. Criteria on the evaluation instrument should be those that can be clearly described so that all raters will have the same kind of behavior in mind. Borman and Dunnette (1975) suggested that if raters can first be trained to observe work-related behaviors more competently, and second to use scales more accurately, it is possible that more error-free portrayals of performance can be made. Bolton (1973) cited the lack of training of administrators as one source of low reliability in evaluating teachers. He stated that training can increase reliability.

Most of this section addressed the need for teacher evaluation training for administrators. The following section focuses on an indirect outcome of evaluation training; namely, the impact of the training on confidence level of administrators.

Confidence Level of Administrators

There appears to be a relationship between training/confidence level and the performance level of administrators in evaluation. According to Saunders

(1984) self-confidence is basic to success. How one views oneself is the result of others' interpretation of our behavior and affects our morale and the degree to which we are enthusiastic, courageous, and ambitious. Bandura (1978) reported that expectations of personal efficacy are derived from several sources, personal accomplishments being the most important. He further stated that efficacy expectations "determine how much effort people will expend and how long they will persist in the face of obstacles" (p. 141). The stronger the perceived self-efficacy, the more active will be the efforts to persist. It seems likely that if we train administrators well in the skills associated with teacher performance evaluation, the confidence levels of the administrators will increase and they will be more willing to participate and persist in this most important supervisory activity.

The relationship between self-confidence and performance is highly important. Feltz and Mugno (1983) found that change in self-efficacy, as a result of training, improved performance levels of trainees. They found a reciprocal effect between self-efficacy and performance. This effect was found to be greatest in the initial stages of training and was characterized by improvements in self-confidence followed by increased

performance which, in turn, produced additional positive changes in self-confidence. The process seemed to produce a cycle much like the commonly known self-fulfilling prophesy. Improved performance influences self-concept which, in turn, influences performance.

Confidence seems to be a variable worthy of consideration in training and research. In summary, then, it appears that self-confidence is critical to improving the performance of administrators in conducting teacher performance evaluations. Eight components of teacher evaluation were selected to determine if training would, in fact, influence the administrators' level of confidence. The literature supporting their inclusion is discussed in the following sections.

Lesson Analysis/Observation and Data Gathering

According to McGreal (1983) classroom observation is the most practical procedure for collecting formal data about teacher performance. He further stated that the quality of observations and the ways supervisors collect and share data with teachers are major factors in the success and effectiveness of teacher evaluation systems.

He also noted that training to improve observation skills is most effective when supervisors have already adopted an appropriate attitude about observation.

Contemporary views of observation, based on research and experience, strongly suggests that the appropriate role for the supervisor in visiting classrooms is to be a collector of descriptive data on a predetermined aspect of the teacher's performance (p. 96).

Evaluation experts find that many principals were unable to conduct effective teacher evaluations because they lacked the skills needed to analyze classroom teaching behaviors (Gudridge, 1980; Krajewski, 1976; Robinson, 1978; Wise et al., 1984). They further concluded that since observation skills play an important role in the success and effectiveness of teacher evaluation, principals, as observers, must develop these skills. Edward's (1985) study supported this finding. She found the need for training in lesson observation to be widespread and common, for almost 80 percent wished for a better way to record what they see in the classroom. Her findings are consistent with Acheson (1982) and Hawley (1982) who reported that a high percentage of administrators felt a need to improve their classroom observation skills.

The literature revealed that observation skill has the potential to be enhanced by training. Manatt (1982) wrote, "Teacher Performance Evaluation (TPE) is a skill (or series of skills) and like skiing, tennis...can be enhanced by training" (p. 2). A study by Faast (1982)

which confirmed Manatt's statement found that the training of evaluators led to greater success in classroom observation. For those principals who lack the observational skills needed for effective classroom observation, training can aid them in acquiring these skills and is essential for effective leadership (Acheson, 1985; Ishler, 1984).

After a review of the literature, Edwards (1985) pointed out: "...classroom observation for purposes of supervision and evaluation is an accepted practice in today's classrooms; problems were found that need to be addressed. These problems center around confusion about the purposes of observation, the brevity and frequency of observations, the lack of validity and reliability, the lack of observational skills, and teachers' concerns about observations" (p. 21).

Rice (1986) stated there is an obvious need to collect accurate data regarding a person's performance so that feedback can be accurate. Feedback, discussed in the next section, is especially important activity for effective conferencing.

Conferencing and Feedback Skills

In successful teacher evaluation systems, formal feedback from the supervisor to the teacher occurs in two

separate, but related types of conferences - the post observation conference and the final conference at the conclusion of the evaluation period. In the post-observation conference, the data are based on the single observation that just occurred. The final conference is built on all the data collected over the full evaluation period. Another difference is that the post-observation conference has a formative evaluation emphasis and can be conducted with a collegial orientation. On the other hand, the final conference is summative, and a judgement must be rendered by a supervisor (McGreal, 1983). Sweeney (1983) noted that there is a sharp distinction between the end-of-the-year conference and the conference following lesson observation. While the former, he explains, is designed to provide teachers with a valid, reliable evaluation of their performance in classroom and non-teaching duties, the post-observation conference is formative, not evaluative: its goal is to help teachers become more effective in the classroom. The threatening nature of summative judgements often makes it difficult for a final conference to be as productive as it might. In spite of the differences between the two types of conferences, there is a set of concepts, principles, and techniques applicable to either type of feedback

situation (McGreal, 1983). Feedback, the information provided to a practitioner about the strengths and weaknesses of teaching or administrative performance of a new methodology or skill, is often unclear, subjective, inaccurate, or irrelevant (McGeoch & Lindsey, 1967; and Acheson & Gall, 1980).

Spencer (1985) reported that not all authorities in the field believe that conferences result in the improvement of teaching practices. Some maintain that conferences are often used to discuss unimportant aspects of teaching and the conferences have little impact on the improvement of teaching. Blumberg (1970), for example, is one who opined that the supervisory conference is not likely to produce teacher growth. Critics attributed this to inadequate training of supervisors. Most authorities in the field, however, viewed the conference in a positive manner with potential for improving teaching behaviors (Spencer, 1985). Sweeney (1982) stated that a lack of proper planning contributes significantly to the problem. He suggested that a post-observation conference requires careful, strategic planning. According to Sweeney, the primary goal of the post-observation conference is to encourage teachers to examine their own effectiveness. This requires skill in data gathering and an ability to provide feedback on the

teaching behaviors observed. A precondition for success is to make a careful analysis of the data gathered during the lesson observation. Sweeney (1983) stated that the post-observation conference is conducted to assist competent professionals to improve their performance. He added, "When administrators are able to use their know-how to deepen teachers understanding of how they function in the classroom and to help them set goals for maintaining and improving performance, they are truly supervisors" (p. 35).

Joyce and Showers (1980) reported that, in order for inservice to be successfully implemented, the participants must study the theoretical basis or rationale for the methodology to be learned, observe the methodology being implemented by an expert, practice the methodology and receive feedback from their colleagues on the strengths and weaknesses of their performance. In a study involving professional engineers, Ivancevich (1982) found when supervisors were trained to give specific feedback to subordinates on their performance and then collaboratively set goals for performance improvement, the subordinates felt that the accuracy, fairness, and clarity of the appraisal interview improved. Moreover, Meyer, Kay and French (1965) observed a 65% improvement rate in employee performance when performance feedback

was translated into improvement goals. However, when performance feedback was not translated into goals, the improvement rate was 27%. Madeline Hunter (1980) suggested that the importance of conferences seems obvious and that conferences do receive mixed reviews with respect to how effectively supervisors conduct them. She contends that many conferences do not result in teacher change, often because the supervisor has not had sufficient training in conducting conferences.

The components discussed thus far relate to evaluation skills that administrators need to successfully evaluate teachers. The following components respond to the purposes for which these evaluation skills serve.

Personnel Decisions

Personnel Decisions Connected to Teacher Evaluation

The 1983 Rand study of teacher evaluation noted that the new concern for the quality of education and of teachers is being translated into decisions about teacher status, merit-pay, and career-ladders that presuppose the existence of effective teacher evaluation systems. When the National Commission on Excellence in Education published *A Nation at Risk: The Imperative for Educational Reform* several of the commission's

recommendations concerned with teaching would require teacher evaluation:

Persons preparing to teach should be required to meet high standards, to demonstrate an aptitude for teaching, and to demonstrate competence in an academic discipline Salaries for the teaching profession should be competitive, market sensitive, and performance-based. Salaries, promotion, tenure, and retention decisions should be tied to an effective evaluation system that includes peer review so that superior teachers can be rewarded, average ones encouraged, and poor ones either improved or terminated (p. 30).

Action for Excellence, a report of the Task Force on Education for Economic Growth, Education Commission of the States, echoed some of the same recommendations:

We recommend that boards of education and higher education in each state - in cooperation with teachers and school administrators - put in place, as soon as possible, systems for fairly and objectively measuring the effectiveness of teachers and rewarding outstanding performance (p. 30).

Teacher evaluation may serve four basic purposes: individual staff development, school improvement, individual personnel decisions, and school status decisions. The second two purposes involve accountability. For the purposes of accountability, teacher evaluation processes must be capable of yielding fairly objective, standardized, and externally defensible information about teacher performance.

A preliminary survey identified 32 districts as having highly developed teacher evaluation systems. In most of the 32 districts, the teacher evaluation system has led to personnel actions. Although few districts used evaluation outcomes to terminate tenured staff, non-tenured staff were dismissed on the basis of evaluation in most sample districts.

Personnel decisions demand the highest reliability of evaluation results. Evaluation criteria must be standardized and evaluators must apply these criteria with consistency when the results are to be used for personnel decisions regarding tenure, dismissal, pay, and promotion. Despite the differences in level of development and diversity of local implementation choices, the major problems associated with teacher evaluation practices were similar in the 32 districts surveyed. Respondents ranked inadequate training for evaluators as one of them.

The validity of a teacher evaluation process depends upon its accuracy and comprehensiveness in assessing teaching quality as defined by the agreed-on criteria. Although school districts may seek to finesse the issue of validity by striving for measurement reliability in their evaluation process, they cannot ignore the validity process when they use the results as a basis for

personnel decisions. The purpose of the evaluation, the inference to be drawn, the help to be given, the decision to be made, determines the validity of the evaluation process. One district trains evaluators in the same teaching principles that guide teacher staff development. This training enhances the correlation between the evaluators's judgements and the standard of practice adopted by the district.

Two recommendations that respond to one of the conclusions of the study address training. Namely, regular assessments as to quality of evaluation should provide input into the continuing evaluator training process, and evaluators should be trained in observation and evaluation techniques, including reporting, diagnosis, and clinical supervision skills, when it adopts a new teacher evaluation process.

By respondent report, a substantial amount of teacher discomfort results from a third problem area; the lack of uniformity and consistency within a school district. While inconsistency in evaluation judgements stems in part from instrumentation, it also reflects inadequate training for evaluators (Wise et al., 1984).

Pay for Performance

If present trends continue, more and more states will mandate career ladder plans. Different conceptions

and interpretations of the career ladder lead to different consequences for teachers as professionals, for the status and functioning of supervisory roles and responsibilities, and for teaching and learning.

Ideally, the career ladder is a system for sharing school responsibilities with teachers and for enhancing their roles as professional partners and school leaders. Merit salary increases are allocated to teachers to reflect this additional responsibility. Within the career ladder, appraisal determines who is meritorious for advancement. Advancement leads to enhanced leadership roles and more responsibility. Leadership roles and responsibility result in salary increases.

At present many states are implementing large scale teacher incentive programs (i.e., career ladder, merit pay, pay for performance, etc.) to develop programs for teachers and administrators. Teacher performance lies at the center of all the programs being implemented or already in operation. The current trend in measurement procedures emphasize peer evaluation, classroom observation, student achievement outcomes, and data from principals, teachers and students (Southern Regional Education Board, 1987).

"The more difficult task, and one that is essential for successful performance-based career plans, is the

development of an evaluation system that is fair and is perceived as fair" (Southern Regional Education Board, 1987, p. 3). There is no question that teacher evaluation is changing, peer review, classroom observations, and use of student achievement are replacing ratings by principals. While early career ladder and incentive programs were to a large degree centralized at the state level, states establishing more recent programs are using more local involvement and control. More structure is now evident. Districts in many of those states are asking for more technical help in developing evaluation procedures. No ideal teacher evaluation model has evolved, nor is a single one likely, but practice is providing directions for change. Career ladder and incentive programs provide ways to restructure schools, reward superior teaching, and focus on student learning (Southern Regional Education Board, 1987).

Tenure/Dismissal

Successful dismissal of a tenured teacher for incompetence hinges upon the administrator's ability to persuade an impartial third-party that she or he has provided such proof (Bridges, 1986). According to Rosenberger and Plimpton (1975), incompetence is a concept without meaning. They further concluded that "There seems to have been no legal need to define

competence" (p. 470), and that "conventional wisdom and common sense, rather than precise standards, have been used in judging incompetence claims" (p. 486).

Bridges (1986) stated that the tenure decision is the single most important personnel decision that administrators make. He believes that the tenure decision is the last opportunity to enforce high performance standards on the teaching staff. In his judgement, the evaluation of probationary teachers and the decision to grant or deny tenure should receive top priority. Bridges suggested making it as hard for teachers to obtain as it is for them to lose tenure. Strike and Bull (1981) stated that the major role of tenure is to secure for the teacher the right of continued employment. Though the tenure system may have been designed to protect teachers from administrative whims, to assure some organizational stability, and, perhaps, to offer security to compensate for relatively low salaries, it has also, according to Blumberg (1980), had some unexpected results for schools in general and supervision in particular. He states, "A teacher who has been granted tenure has at his disposal a formidable device with which to insulate himself from the pressures of other teachers or his supervisor" (p. 47). Strike and Bull (1981) declared that once teachers attain tenure,

their legal status changes dramatically. The full protection of the due process clause of the 14th amendment thereby becomes available.

The tenured teacher, they explained, can only be dismissed for cause and has a wide range of due process rights. Incompetence or its legal synonym, inefficiency, is a legitimate cause for termination in most state tenure statutes. It is the charge to which teacher evaluation is most obviously and immediately applicable. Strike & Bull (1981) suggests that three conclusions may be drawn relative to dismissal: First, courts are likely to rely on the professional judgement of administrators in the substantive aspects of evaluation. Secondly, judicial review of dismissal decisions is likely to be more restrictive when dealing with the procedural aspects of dismissal. Third, despite the lack of an authoritative legal definition as well as jurisdictional variations in interpretation, a general and widely accepted core of meaning for teaching incompetence can be discerned in case law.

Strike and Bull asserted that the personnel file will be appealed to in any action taken with respect to a teacher. They suggested that the administrator should thus focus on generating an evaluation system that produces a record about the teacher that is characterized

by informative descriptions and clear assessments appealing to known and relevant criteria and collected according to a known and orderly process. The need to make personnel decisions in a legal and morally responsible manner is one of the summative uses and justifications for a systematic program of teacher evaluation (Strike and Bull, 1981).

Supervisor-Teacher Relationships

"The most successful evaluation systems are designed to increase the quality of supervisor-teacher time, not the quantity" (McGreal, 1983, p. 104). When dealing with clinical supervision as supervisory model (Blumberg and Amidon, 1965; Boyan and Copeland, 1974; and Shinn, 1976), heavy emphasis was placed on collegial relationships, non-directive technique, and reliance on assumptions about teachers being willing and able to assume major responsibility. Hyman (1975) illustrated the importance of the pre-conference when he talks about the value of teachers and supervisors conferring together to develop goals. Iwanicki (1981, p. 226) provided a good summary of the major strengths and weaknesses of goal setting. One of the strengths he cited is that it "fosters a positive working relationship between teacher and evaluator."

McGreal (1983) noted that it is increasingly apparent that all participants must not only receive adequate training, but also must be provided with a system that supports and enhances supervisory-teacher relationships. The concern for the system and for its procedures is not intended to deny the importance of the individual relationship between a supervisor and a teacher. Moreover, the system that works best, imposes the fewest possible infringements upon the supervisor-teacher relationship. Experience shows that a positive, supportive relationship between a knowledgeable supervisor and a committed teacher is still a very effective way to improve instruction.

One of the assumptions of clinical supervision models that Sergiovanni (1982) noted was that supervision is a "partnership in inquiry" with the supervisor as a more experienced practitioner instead of an aloof expert. This type of relationship can, in many cases, supersede an inadequate evaluation system (McGreal, 1983).

Sweeney (1982) stated that if principals are to improve teacher performance, it must be in the context of a helping rather than authoritative relationship. While some individuals possess that innate ability to communicate empathy, understanding, and a desire to help, most principals need to work on their behavior in this

regard. Sweeney (1983) suggested that supervisors must be able to exhibit behaviors consistent with sound human relations and management principles. He recommended that supervisors be sensitive to teachers' professional pride, as well as to their attitudes and feelings. To be successful, supervisors must develop a climate of engendering confidence and trust, and exhibit excellent interpersonal skills (Sweeney, 1983).

Summary of the Chapter

This chapter has focused on effective teacher performance evaluation and training. The impact of training on confidence level, and the following evaluation skills and purposes were discussed: lesson analysis and observation and data-gathering skills, conferencing and feedback skills, personnel decisions, supervisor-teacher relationships, and improving instruction. The literature divulges that much is known about evaluator training and the components of effective teacher evaluator training.

The present study examined the delivery to administrators of content related to teacher performance evaluation, and the effects of that training on the confidence of the administrators as evaluators and trainers of evaluators.

CHAPTER III. METHODS AND PROCEDURES

The purpose of this chapter is to discuss the methods and procedures used to assess the effects of performance evaluator training on the skills and confidence levels of teacher evaluators and trainers of teacher evaluators. The data for this study were collected as part of the train-the-trainer workshops under the federally funded Iowa-Leadership in Educational Administration (I-LEAD) project. The objective of the train-the-trainers project was twofold: to train a group of Iowa Administrators in teacher evaluation and to train them how to train others to become teacher evaluators.

This chapter, which describes the methods and procedures used to gather and analyze the data required for the study, has been divided into two major sections. The first section, "Collection of Study Data" describes the research design, the sample, materials development, video-tapes, the instrumentation used to collect data for the study, and the procedures. The second section, "Analysis of Data," reviews the statistical methods used in the treatment of the data.

Collection of Study Data

Research Design

This study includes two phases: (a) development of the materials and training method, and (b) training. Phase one included the development of a training process which utilized a task analysis derived from a series of strategy meetings. These meetings resulted in instructional plans (Appendix A) that detailed the content, presenters, resources, and time schedules for the training sessions. A major undertaking in phase one was the development of a training manual for the training sessions. Included in the training manual are the goals and objectives tied to the standards for evaluator approval in the state of Iowa. A companion manual containing instructional graphics and the assessment instruments also was developed.

Phase two of this study was conducted using a pretest/posttest design. Six days of training were provided in two-day sessions to sixty-four Iowa School administrators. Of the six days training, five were devoted to teacher evaluation and one day to administrator evaluation. Professors of the Educational Administration faculty at Iowa State University provided the training. Some time during the six days was used to

assist participants in how to deliver the training since the participants were committed to conduct training later through their Area Education Agency (AEA). Permission to conduct the study was secured from the ISU Committee on the Use of Human Subjects in Research in November, 1987 (Appendix B).

The Sample

The sixty-four participants including superintendents, principals, and other supervisory personnel in Iowa were selected at random from a pool nominated by fifteen Area Education Agency educational service directors. Letters were sent to the fifteen AEA educational services directors on October 1, 1987 asking that they nominate a representative number of superintendents, principals, and central office personnel for the train-the-trainer project. Letters were sent on November 2 to sixty-four prospective participants inviting them to participate; three declined, others replaced them. Of the sixty-four participants, make-up sessions were held for ten of them as a result of scheduling conflicts. A cross section of Iowa's administrators and supervisory

TABLE 1. Distribution of sample by job classification

	Administrators	Public School	Non-Public School
1. Superintendents.....	12		
2. Central Office.....	7		
3. Secondary Principals.....	8		1
4. Assist. Sec. Principals....	1		
5. Middle/Jr. Principals.....	3		
6. Elementary Principals.....	13		4
7. Assist. Elem. Principals...	1		
8. AEA Personnel.....	11		
9. College Professors.....	3		
TOTAL	59		5
GENDER	NUMBER		PERCENT
MALE	53		83
FEMALE	11		17
TOTAL	64		100

personnel were represented (Table 1); both in terms of levels of experience and geographical location in the state. The group comprised eleven females and fifty-three males. Five of the twenty-six principals were non-public school principals. The majority of the participants were public school male principals.

Materials Development

The I-LEAD training manuals were the primary resources used in training evaluators. Below is an outline of the organization of the manuals.

ORGANIZATION

Manual

- I. Introduction

- II. Sessions
 - A. Instructional Plan
 - B. Objectives
 - C. Teaching Materials
 - D. Training Tips
 - E. Literature/References

- III. Training Materials

Companion Manual

- I. Transparencies/Graphic

The brief statements that follow indicate the purposes and uses of the various sections that are included in the manuals.

The instructional plan provides a breakdown of the topics, presenters, visuals, handouts, and time schedule for each session.

The objectives provide the focus for each module.

The teaching materials provide important information which can be used by both the trainer and trainees.

The training tips provide information about the transparencies and their sequence.

The literature/references are either summaries of important materials which can be used by the trainers or trainees, or are citations of research or articles to read for further information.

The training materials are those materials that could be used primarily in a training activity.

Transparencies/Graphics are the transparencies used for training (companion manual).

Video-Tapes

Video tapes were used during the training. The tapes were selected because they depicted average teaching performance, revealed teachers who exhibited reasonably explicit strengths as well as weaknesses, were of appropriate length (approximately 30 minutes), and

were at the middle school level and could be used with K-12 administrators.

One of the video tapes, Gerry Page I, was used in conjunction with the Teacher Performance Rating Scale instrument for pre/posttesting of the participants lesson observation and analysis skills to determine if training positively affected interrater reliability, i.e., whether the participants agreed more on their ratings of twenty-two teacher skills after training. The Larry Mann and Mary Curtin video tapes were used to give the participants an opportunity to practice lesson observation and analysis skills during training.

Note: For further information about the video tapes used in this study please contact Dr. Jim Sweeney of Iowa State University.

Instrumentation

Two instruments were utilized in this study; namely, TPE Rating Scale (Appendix C) and the Supervisory Attitude Survey (Appendix D). The TPE Rating Scale instrument received human subjects approval and had been used in research previously conducted at Iowa State University. The Supervisory Attitude Survey was designed specifically for this study. This instrument was constructed in consultation with the researcher's major professor and was modified several times for clarity and to make the items more conducive to the purposes of this

study. Both instruments were administered to participants during the I-LEAD Train-the-Trainers workshop.

A description of, and information for, each instrument follows:

TPE Rating Scale- A three-point Likert rating scale containing twenty-two teacher skills was used in conjunction with the Gerry Page I video tape for participants to record their ratings during Lesson Analysis and Observation pre- and posttesting.

Supervisory Attitude Survey- This twenty-four-item instrument was designed to gather data related to the administrators' level of confidence as teacher evaluators and trainers of teacher evaluators in important lesson analysis/observation skills, data-gathering, and conferencing and feedback skills as well as in making personnel decisions, maintaining positive supervisor-teacher relationships, and improving instruction. A nine-point Likert scale was used for this instrument. The instrument measured supervisor responses to twenty-four statements on a scale from 1, "strongly disagree," to 9, "strongly agree." Participants completed this survey as a pre- and posttest. The data in Tables 2 and 3 were derived from the reliability test of the composites of the posttest ratings of confidence

levels in selected components. These composites were based on theoretical concepts derived from the literature. The Tables show the degree to which the measure will yield similar results for the same subjects at different times, i.e., the consistency of the instrument. The reliability coefficient alphas of .71, .80, and .81 in Table 2 and .88, .86, and .73 in Table 3 suggest that the composites on this instrument are highly free of error variance and are a measure of true differences among persons in the dimensions assessed. The reliability coefficient alphas were obtained using the Cronbach's Coefficient Alpha test. The average item correlations indicate the correlation of the dimensions within each composite.

Procedures

The sixty-four Iowa School Administrators received six days of training, in three two-day sessions. The study took place in Ames, Iowa on December 7, 8, 1987, January 25, 26, 1988, and February 8, 9, 1988. Of the six days training, five were devoted to teacher evaluation. Some time during the six days was used to help participants to be trainers.

The following procedures were implemented during the six days of workshops:

1. A statement regarding the purpose of the workshop and the related research was read to the

TABLE 2. Reliability distribution of composites of the posttest ratings of confidence levels as teacher evaluators in selected components

SELECTED COMPONENTS	AVERAGE ITEM CORRELATION	ALPHA
Data Gathering What to Record Observe and Record	.55	.71
Lesson Analysis Identify Specific Areas Lesson Design Effectiveness of Strategy	.66	.80
Lesson Analysis/Observation Observation/Lesson Analysis Interpret Data	.69	.81

TABLE 3. Reliability distribution of composites of posttest ratings of confidence levels as trainers of teacher evaluators in selected components

SELECTED COMPONENTS	AVERAGE ITEM CORRELATION	ALPHA
Data Gathering What to Record Observe and Record	.79	.88
Lesson Analysis Identify Specific Areas Lesson Design Effectiveness of Strategy	.75	.86
Lesson Anal/Observ Observation/Lesson Anal Interpret Data	.58	.73

participants. The voluntary nature of the research was emphasized in the statement and read to the participants to ensure adherence to the Human Subject's Committee's guidelines.

2. Participants viewed a videotaped lesson and completed the teacher performance evaluation rating scale.

3. Participants completed the Supervisory Attitude Survey.

4. Participants received training in Lesson Observation and Analysis.

5. Participants received training in Conferencing.

6. Participants received training in writing Professional Improvement Commitments.

7. Participants received training in Formative and Summative Evaluation.

8. Participants viewed a videotaped lesson and completed the Teacher Performance Evaluation Rating Scale.

9. Participants completed the Supervisory Attitude Survey.

Analysis of Data

After the two completed instruments were received, the data were delivered to the test and evaluation

center. The data were then transferred to the computation center. Statistical treatment of the data was completed using the Statistical Package for the Social Science (Norvsiis, 1983) computer program. Descriptive statistics (frequencies, means, and standard deviations) were computed to study the relative value of study variables. Frequencies and paired t-tests were used to assess the mean differences between pretest and posttest confidence ratings.

CHAPTER IV. FINDINGS

The basic problem for this study was to assess the effects of training on the skills, and confidence level of teacher evaluators and trainers of teacher evaluators. Two instruments were used to collect data for this study. The data were collected from sixty-four Iowa educators including superintendents, principals, assistant principals, central office personnel, Area Education Agency personnel, and college professors. These Iowa educators received six days of performance evaluator training designed to help them to become better evaluators as well as enable them to train other administrators and supervisors in performance evaluation.

The study was conducted in two phases. During the first phase of the study, materials and training methods were developed. Several planning sessions of those who designed and delivered the training were held from mid October to late November. The development of a training process resulted in a training manual which contained the workshop content and design, resources, and time schedules for the training sessions. A companion manual containing instructional graphics and assessment instruments was also developed.

In the second phase of the study professors of the Educational Administration faculty at Iowa State University provided training. The cluster of skills addressed were those needed to improve school administrators' ability to evaluate, analyze, and improve teacher performance. Two instruments and a pretest/posttest design were used to assess the workshop participants' growth in skills and confidence as a result of training. These instruments may be seen in Appendices C-D.

The purpose of this chapter is to report the results of the investigation. Each of the research hypotheses presented in Chapter I is presented and the results of the statistical tests are discussed and displayed in table form.

Descriptive Data

Sample

Sixty-four evaluators participated in the study which took place in Ames, Iowa on December 7, 8, 1987, January 25, 26, 1988, and February 8, 9, 1988. Subjects for the study were superintendents, principals, assistant principals, Area Education Agency personnel, college professors and other central office

supervisory/administrator personnel involved in teacher performance evaluation. The sixty-four participants were selected at random from a pool provided by 15 AEA educational service directors. The criteria for selection to the pool included "expertise and interest in teacher evaluation" and "ability to provide training." Letters were mailed on November 2 to the sixty-four prospective participants inviting them to participate. Three of the sixty-four were not able to participate and were replaced from the pool. A cross section of Iowa's administrative and supervisory personnel was represented; both in terms of levels of experience and geographical location. Of the sixty-four participants, make up sessions were held for ten of them as a result of scheduling conflicts.

Interrater Reliability

One of the objectives of the study was to determine if evaluators were more likely to agree on ratings of teacher performance after training. A reduction in the standard deviations from the pretest to the posttest implies that the raters are in more agreement after training. Tables 4 and 5 offer summaries of the evaluators' ratings. Table 4 displays the data for the pretest and posttest of the participants' rating of

teacher performance in twenty-two teaching skills. After viewing a videotaped lesson, the participants were asked to complete a formative evaluation report using a 3 point likert ratings scale (1=low, 3=high). It should be noted that the same videotape was used for the pretest and posttest. Table 5 displays the standard deviations of the pretest/posttest ratings of teacher behaviors. The standard deviations were computed using the 1, 2, and 3 responses with non-response representing 0 values, which accounts for the difference in the number of responses for each skill item. There is a general decrease in the standard deviations of the ratings, for nineteen of twenty-two teaching strategies. The mean decline in the standard deviations for the twenty-two items on the pretest (.62) to the posttest (.34) suggests that the raters were in greater agreement in their ratings after training. The largest difference was for "guided practice skill" (-.93) where all the participants agree on their ratings on the posttest, while the smallest decreases occur for "communication skills" and "student participation" (-.07 and -.09 respectively). The most noteworthy exception was "state objectives," where there was a substantial increase in the standard deviation, indicating that the evaluators were in greater disagreement on the teacher's performance in stating the

objectives after training. The highest standard deviation was in the "enrichment/ remediation" skill (.99) for the pretest but was reduced substantially to a standard deviation of .58 on the posttest.

Table 5 displays the frequency counts for each point on the scale and for each skill (the 0 point indicates non-response). The frequency distribution among the rating points for each skill differs markedly from pretest to posttest. The large percentage of non-responses in the pretest data (249) is substantially reduced in the posttest data (65). This is accompanied by a rise in the percentage of "high" responses which dominate the posttest numbers. Both of these observations are consistent with what is seen in Table 4, and provide an explanation of why the standard deviations behave the way they do: The decrease in non-responses and the large percentage of "high" responses decrease the standard deviations, conversely for the "personal organizational skills," "learning structure," and "state objectives" items.

TABLE 4. Distribution of the standard deviations of pretest/posttest ratings (1=low to 3=high) of selected skills as teacher evaluators

Selected Skills	PRETEST		POSTTEST		Diff.
	N	SD	N	SD	
1. Personal Organ'1 Skills	58	.00	63	.13	+.13
2. Organizes Stud for Eff Instruc	55	.59	60	.33	-.26
3. Lrng Structure	62	.34	61	.53	+.19
4. State Objectives	60	.35	57	.92	+.57
5. Input	61	.58	62	.41	-.17
6. Modeling	37	.63	62	.13	-.50
7. Stu Participation	63	.47	63	.38	-.09
8. Clear Directions	63	.33	63	.13	-.20
9. Eff Questioning	56	.74	62	.47	-.27
10. Guided Practice	37	.93	63	.00	-.93
11. Checks for Understanding	58	.90	63	.53	-.37
12. Paces Lesson	59	.69	61	.43	-.26
13. Feedback	52	.75	62	.25	-.50
14. Enrich/Remediat'n	25	.99	55	.58	-.41
15. Communicat'n Skls	56	.37	63	.30	-.07
16. Eval Activities	23	.82	44	.62	-.20
17. Knowledgeable	60	.28	63	.00	-.28
18. Motivates Stu	51	.80	56	.61	-.19
19. Time On Task	48	.88	62	.22	-.66

Table 4 Continued

Selected Skills	PRETEST		POSTTEST		Diff.
	N	SD	N	SD	
20. High Stndrds for Stud Behavior	51	.51	57	.00	-.51
21. Sensitivity	44	.54	61	.13	-.41
22. Equal Treatmnt	54	.67	59	.39	-.28
Overall Rating		.62		.34	-.28

TABLE 5. Frequency of responses of evaluators' ratings for each teaching skill

Skills	0	Pretest			3	Posttest			
		1	2			0	1	2	3
1. Personal Organ'l Skills	5	-	-	58	-	-	1	62	
2. Organizes Stud for Eff Instruc	4	4	9	42	3	1	3	56	
3. Lrng Structure	1	-	8	54	2	4	3	54	
4. State Objectives	3	1	4	55	6	25	10	22	
5. Input	2	5	3	53	1	2	3	57	
6. Modeling	26	3	6	28	1	-	1	61	
7. Stu Participation	-	1	13	49	-	1	6	56	
8. Clear Directions	-	1	3	59	-	-	1	62	
9. Eff Questioning	7	8	15	33	1	2	8	52	
10. Guided Practice	26	15	6	16	-	-	-	63	

Table 5 Continued

Skills	0	Pretest			3	Posttest			
		1	2			0	1	2	3
11. Checks for Understanding	5	23	12	23	-	3	9	51	
12. Paces Lesson	4	7	6	46	2	2	4	55	
13. Feedback	11	8	8	36	1	1	-	61	
14. Enrich/Remediat'n	38	14	1	10	9	4	5	45	
15. Communicat'n Skls	7	-	9	47	-	-	6	57	
16. Eval Activities	40	12	6	5	19	4	4	36	
17. Knowledgeable	3	-	5	55	-	-	-	63	
18. Motivates Stu	12	10	12	29	7	4	11	41	
19. Time on Task	15	13	6	29	1	-	3	59	
20. High Standrds for Stud Behavior	12	3	2	46	6	-	-	57	
21. Sensitivity	19	3	2	39	2	-	1	60	
22. Equal Treatment	9	6	5	43	4	2	1	56	
Total	249	137	141	855	65	64	79	1187	
Cum %	18	10	10	62	5	4	6	84	

Scale: No Mark = Absent OK
 1 = Absent, Not OK
 2 = Present, Unacceptable
 3 = Present, Acceptable

Confidence Level

Table 6 displays summary statistics for pre- and post-ratings of the participants' levels of confidence as teacher evaluators in each of the selected skill components. The participants were asked to complete the 24 item "Supervisor Attitude Survey", in which they were asked to indicate their level of agreement with statements concerning their confidence levels. The 9 point Likert scale (1=strongly disagree, 9=strongly agree) is used to measure their level of confidence as teacher evaluators. The participants had the most confidence in "interpreting data recorded" (6.69), their "ability to record data", (6.67), and "what data to record", (6.66). The mean ratings suggest that prior to training, the participants were least confident in their ability to make "compensation pay decisions" for teachers; they indicated they were "not sure" on the rating scale (5.02). With the exception of "compensation pay decisions", all components fall roughly into the middle of the "agree" range (6 or 7), indicating that the participants' level of confidence was fairly high for most components prior to training. All of the participants were even more confident after training. The largest rise in confidence was perceived in their

ability to "improve instruction" (1.64). The smallest increase in the participants confidence level was in their data gathering abilities (.79 and .65 respectively). However, it should be noted that the participants were highly confident in their ability to gather data prior to training. While the participants levels of confidence for "what data to record" and "interpreting data recorded" were similar, the participants levels of confidence for the two differed substantially after training with mean differences of .79 and 1.07. Overall, the participants' ratings of their confidence after training increased markedly except for "compensation pay decisions", where they rated themselves least confident. According to the rating scale used, the overall pretest mean of 6.32 (agree), and the mean of 7.46 (strongly agree) for the group after training indicate a 1.14 overall increase in their level of confidence as teacher evaluators.

Table 7 displays pre- and postscores reflecting the participants' levels of confidence as trainers of teacher evaluators for each of the selected components. The

TABLE 6. Pretest/posttest ratings of confidence level as teacher evaluators for selected components

SELECTED COMPONENTS									
Data Gathering									
What Data to Record								
Ability to Record Data								
Lesson Analysis									
Identify Specific Areas for Growth								
Lesson Design Decisions								
Teaching Strategy Decisions								
Lesson Analysis/Observation									
Interpret Data Recorded								
Observation and Lesson Analysis								
Use Data for Feedback								
Improving Instruction								
Compensation Pay Decisions								
Tenure Decisions								
Maintaining Positive Supervisory Relationship									
Relationship								
Overall								
Scale =	1	2	3	4	5	6	7	8	9
	Strongly disagree		Disagree		Not sure		Agree		Strongly agree

	N	<u>PRETEST</u> MEAN	SD	N	<u>POSTTEST</u> MEAN	SD	Mean Dif's
.....	58	6.66	1.37	58	7.45	1.60	.79
.....	58	6.67	1.22	58	7.32	1.58	.65
.....	58	6.43	1.46	58	7.41	1.38	.98
.....	58	6.14	1.53	58	7.29	1.19	1.15
.....	58	6.24	1.20	58	7.40	1.06	1.16
.....	58	6.69	1.41	58	7.76	.84	1.07
.....	58	6.26	1.21	58	7.72	.93	1.46
.....	58	6.59	1.44	58	7.74	.77	1.15
.....	58	6.21	1.51	58	7.85	1.04	1.64
.....	58	5.02	1.99	58	5.98	1.67	.96
.....	58	6.36	1.54	58	7.69	1.13	1.33
.....	58	6.57	1.29	58	7.95	.80	1.38
.....		6.32			7.46		1.14

participants were asked to complete the "Supervisor Attitude Survey", a 24 item, 9 point likert scale, that was used to measure their level of confidence as trainers of teacher evaluators. The participants showed the most confidence as trainers in their "ability to record data," "observation and lesson analysis," and "use data for feedback," (6.14 each). The data suggest that prior to training, the participants were least confident in their ability to train evaluators to make "compensation pay decisions (5.36). The data show that the participants' confidence ratings increased after training, with the largest increase in confidence occurring for the ability to train evaluators to "improve instruction" (1.72). The data gathering components show the smallest increases (1.12 and 1.19) after training, with the exception of "compensation pay decisions (0.62).

Overall, the participants' confidence increased markedly after training, except for "compensation pay decisions" (.62) where they were least confident after training (5.98). According to the rating scale used, the overall pretest mean of 6.02 (agree), and the mean of 7.33 (strongly agree) for the group after training indicate a 1.31 increase in their level of confidence to train teacher evaluators.

TABLE 7. Pretest/posttest ratings of confidence levels as trainers of teacher evaluators for selected components

SELECTED COMPONENTS

Data Gathering

What Data to Record
 Ability to Record Data

Lesson Analysis

Identify Specific Areas for Growth
 Lesson Design Decisions
 Teaching Strategy Decisions

Lesson Analysis/Observation

Interpret Data Recorded.....
 Observation and Lesson Analysis.....

Use Data for Feedback

Improving Instruction

Compensation Pay Decisions.....

Tenure Decisions

Maintaining Positive Supervisory Relationship

Overall

Scale = 1 2 3 4 5 6 7 8 9
 Strongly Disagree Not Agree Strongly
 disagree sure Agree agree

	N	<u>Pretest</u> MEAN	SD	N	<u>Posttest</u> MEAN	SD	Mean Diff.
.....	58	6.05	1.34	58	7.17	.99	1.12
.....	58	6.14	1.54	58	7.33	.98	1.19
.....	58	6.12	1.45	58	7.43	.96	1.31
.....	58	6.12	1.46	58	7.33	1.11	1.21
.....	58	6.03	1.49	58	7.33	.93	1.30
.....	58	6.05	1.36	58	7.45	.94	1.40
.....	58	6.14	1.50	58	7.67	.85	1.53
.....	58	6.14	1.41	58	7.62	.91	1.48
.....	58	5.85	1.53	58	7.57	.92	1.72
.....	58	5.36	1.71	58	5.98	1.74	0.62
.....	58	6.09	1.45	58	7.33	1.18	1.24
.....	58	6.09	1.55	58	7.69	.80	1.60
.....		6.02			7.33		1.31

Inferential Statistics

This section reports findings on the hypotheses tested in this study.

Hypotheses

Below are the hypotheses which provided the foci for the study:

1. There will be significantly higher posttest ratings of confidence level of teacher evaluators for selected components after training.
2. There will be a significantly higher posttest ratings of confidence level needed to become a trainer of teacher evaluators for selected components after training.

Hypotheses Testing

Below are the null hypotheses and the results of hypotheses testing. In order to test the hypotheses, it was predetermined that six of the eight components must show significant differences between pre- and posttests to reject the null hypothesis. The eight components

utilized in this study are as follows: Data Gathering, Lesson Analysis, Lesson Analysis and Observation, Use Data for Feedback, Improving Instruction, Compensation Pay Decisions, Tenure Decisions, and Maintaining a Positive Supervisory Relationship. The paired t-test was used to test the differences between pre- and posttests for significance at the .05 level for each component. All probabilities less than .05 were reported. A discussion of each hypothesis follows.

Ho 1 There is no significant difference between pre- and posttest ratings of confidence level of teacher evaluators for selected components after training.

Research hypothesis 1 was designed to determine if the participants were more confident in their ability in each of the eight selected components after training. The lowest t-value of 3.20 for data gathering with a posttest mean of 7.39 was .73 points higher than the pretest mean of 6.66, ($t=3.20$, $p < .01$). The highest t-value of 9.54 for improving instruction with a posttest mean of 7.84 was 1.63 points higher than the pretest mean

of 6.21, ($t=9.54$, $p < .01$). All eight components were significant at the .01 level. The null hypothesis was rejected, indicating that the group became significantly more confident as teacher evaluators in the eight components after training.

Table 8 displays results from paired t-tests of the null hypothesis of no difference between pre- and posttest confidence ratings of teacher evaluators for these components.

Ho 2 There is no significant difference between pre- and posttest ratings of confidence level needed to become a trainer of teacher evaluators for selected components after training.

Research hypothesis 2 was designed to determine if the participants were more confident in their ability in each of the selected components after training. The lowest t-value of 2.35 for "compensation pay decisions" with a posttest mean of 5.98 was .62 points higher than the pretest mean of 5.36, ($t=2.35$, $p < .05$). The highest t-value of 8.33 for "improving instruction" with a posttest mean of 7.57 was 1.73 points higher than the pretest mean of 5.84, ($t=8.33$, $p < .01$). Seven

TABLE 8. Analyses of pretest/posttest ratings of confidence level as teacher evaluators in selected components

SELECTED COMPONENTS	N	MEAN	SD	T	TWO-TAIL PROBABILITY
Data Gathering					
Pretest		6.66	1.40		
Posttest	58	7.39	1.14	3.20	0.00 **
Lesson Analysis					
Pretest		6.27	1.02		
Posttest	58	7.37	1.11	7.50	0.00 **
Lesson Analysis/Observation					
Pretest		6.47	1.15		
Posttest	58	7.74	0.81	9.21	0.00 **
Conferencing					
Pretest		6.59	1.45		
Posttest	58	7.74	0.77	6.35	0.00 **
Improve Instruction					
Pretest		6.21	1.51		
Posttest	58	7.84	1.04	9.54	0.00 **
Compensation Pay Decisions					
Pretest		5.02	1.99		
Posttest	58	5.98	1.67	3.73	0.00 **
Recommend Tenure					
Pretest		6.36	1.54		
Posttest	58	7.69	1.13	6.02	0.00 **
Positive Supervisory Relationship					
Pretest		6.57	1.29		
Posttest	58	7.95	0.80	7.87	0.00 **

** .01.

TABLE 9. Analyses of pretest/posttest ratings of confidence level as trainers of teacher evaluators in selected components

SELECTED COMPONENTS	N	MEAN	SD	T	TWO-TAIL PROBABILITY
Data Gathering					
Pretest		6.09	1.40		
Posttest	58	7.25	0.93	6.07	0.00 **
Lesson Analysis					
Pretest		6.09	1.39		
Posttest	58	7.36	0.89	6.90	0.00 **
Lesson Analysis/Observation					
Pretest		6.09	1.34		
Posttest	58	7.56	0.79	8.29	0.00 **
Conferencing					
Pretest		6.14	1.40		
Posttest	58	7.62	0.91	7.70	0.00 **
Improve Instruction					
Pretest		5.84	1.53		
Posttest	58	7.57	0.92	8.33	0.00 **
Compensation Pay Decisions					
Pretest		5.36	1.71		
Posttest	58	5.98	1.74	2.35	0.02 *
Recommend Tenure					
Pretest		6.09	1.45		
Posttest	58	7.33	1.18	5.60	0.00 **
Positive Supervisory Relationship					
Pretest		6.09	1.55		
Posttest	58	7.69	0.79	7.53	0.00 **

* .05.

** .01.

components were significant at the .01 level and one component was significant at the .05 level. The null hypothesis was rejected, indicating that the group became significantly more confident as trainers of teacher evaluators in the eight components after training.

Table 9 displays results from the paired t-tests of the null hypothesis of no difference between pre- and posttest confidence ratings of trainers of teacher evaluators for these components.

CHAPTER V. DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

The primary purpose of the study was to assess the effects of training on the skills and confidence level of teacher evaluators, and trainers of teacher evaluators. In this chapter, conclusions based on the findings are reported and discussed and recommendations for further research are made. The chapter has been organized into the following sections: (a) discussion and conclusions, (b) limitations, and (c) recommendations for further research.

A discussion of the findings, based on data gathered in the fall of 1987 and the spring of 1988, from those participating in teacher evaluation training and a discussion of those findings follow.

Discussion and Conclusions

This study has very important implications for teacher evaluators, and those who train teacher evaluators and trainers of teacher evaluators. If the results found in this study are supported by further research, one might conclude that the confidence level of teacher evaluators and trainers of teacher evaluators is strongly influenced by teacher evaluation training and that this training may result in more effective teacher

evaluation. One might also consider that interrater reliability can be increased by training. Both could have a significant effect on improving instruction.

Sixty-four educators involved in teacher evaluation training in Ames, Iowa provided data for the study. During the fall 1987 and the spring of 1988, they participated in an activity designed to assess the effects of training on the skills and confidence level of teacher evaluators and trainers of teacher evaluators. The findings follow:

1. The evaluation training significantly influenced the confidence level of teacher evaluators for selected components.
2. The evaluation training significantly influenced their confidence level as trainers of teacher evaluators in selected components.
3. The evaluation training did affect the degree to which the evaluators agreed on ratings of teacher performance. There was more interrater reliability following training.

The first two findings which resulted from hypothesis testing are presented and discussed first, followed by other findings and discussion related to the descriptive data for interrater reliability.

Findings

Two highly significant findings resulted from this study: 1) evaluation training influences the confidence level of teacher evaluators in selected components, and

2) evaluation training influences the confidence level needed to become a trainer of teacher evaluators in selected components.

The findings showed a significant difference in the participants' level of confidence as teacher evaluators in the selected components after training. The participants' overall level of confidence as evaluators increased after training and they became significantly more confident in their knowledge and skills in each of these eight components: Data Gathering, Lesson Analysis, Lesson Analysis and Observation, Use of Data for Feedback, Improving Instruction, Compensation Pay Decisions, Tenure Decisions, and Maintaining Positive Supervisory Relationships.

The increase in their level of confidence while significant for each component, was greater in some than in others. The greatest increase in their confidence level was in their ability to improve instruction. The least increase in confidence was in data gathering. However, it should be noted that they were highly confident in the data gathering skill prior to training. They were least confident (before and after training) in their ability to make valid decisions about teacher compensation.

It is interesting to note that their confidence as teacher evaluators parallels their confidence level as trainers of teacher evaluators. There was a significant difference after training in their level of confidence as trainers in each of the selected components and in their overall level of confidence. Both increased significantly after training. The confidence level of the participants' ability to train evaluators to "improve instruction" showed the greatest increase. The least gain in confidence was in their ability to train evaluators to make valid decisions about teacher compensation. Finally, they were least confident in their ability to train evaluators to make valid decisions about teacher compensation.

The teacher evaluation training had a significant effect on the confidence level of the participants as evaluators and as trainers of teacher evaluators. Since Bandura et al. (1977) reported that expectations of personal efficacy are derived from several sources, personal accomplishments being the most important, this rise in level of confidence has implications for teacher evaluation training. It follows, then, if we provide administrators the knowledge and skills they need for teacher evaluation, their confidence will increase and perhaps their personal efficacy as well. The

participants' level of confidence as evaluators and trainers of evaluators was highest in their ability to improve instruction. This is not surprising since they improved in all the areas that are the essential elements for improving instruction. The participants were least confident in their ability to make "compensation pay decisions." This also was not surprising since this component is regarded as one of the most difficult tasks associated with teacher evaluation and is associated more with politics than skill. What we do not know for certain is whether the low confidence level in making compensation decisions is ascribed to the political nature of the task or their concern with their individual skills.

This study indicates that given proper training and sufficient practice, teacher evaluators and trainers of teacher evaluators can become more confident in their skill and perhaps more proficient.

Other Findings

One of the objectives of the study, not subjected to hypothesis testing, was to determine if, after training, the evaluators were more likely to agree on ratings of teacher performance. The data suggest that the training made a difference because the raters were in greater agreement in their ratings of teacher performance after

training. The variance between raters was reduced in 19 of the 22 selected teacher skills after training. Some of these reductions were substantial. For example, the greatest reduction in variance between raters was for "guided practice". The ".93" standard deviation in this skill was reduced to ".00," indicating that the participants unanimously agreed after training. The greatest disagreement for "enrichment/remediation" was reduced by ".41." It was ".99" before training and ".58" after training. There was more variance in the participants ratings for "personal organization skills," "learning structure," and "state objectives," indicating that they disagreed more after training.

Training positively influenced the interrater reliability (consistency among raters) of teacher performance. While inconsistency in evaluation judgements stems in part from instrumentation, it also reflects inadequate training of evaluators (Wise et al., 1984). This is consistent with Borman & Dunnette (1975) who observed that if raters can first be trained to observe work-related behaviors more competently, and second to use scales more accurately, it is possible that more-error free portrayals of performance can be made. It also supports research by Bolton (1973) who noted that untrained administrators are a source of low reliability

in evaluating teachers and that training can increase interrater reliability. The increases in the degree of agreement on teacher ratings for 19 of the 22 teacher skills indicate that training positively influenced interrater reliability. It should be pointed out that this does not imply that the ratings are valid as there is no standard from which to judge. These findings should be viewed with caution because the instrument's scale may have created confusion for the participants. The participants rated the teacher's performance in each skill using 1 (low) to 3 (high) or could not respond. However, a non-response indicated "not observed, ok," implying that it was not present but acceptable if absent. For example, 26 participants rated "guided practice" "not observed, ok" and it was not needed in the lesson. This is acceptable. The participants disagreed more after training for "state the objectives," becoming even more diversified. It should be pointed out that this is in a sense a positive aspect, since the rating should not have been a "3" according to workshop trainers. The increase in variance for "organizational skills" and "learning structure" are difficult to assess since these increases are quite small. What is disconcerting is that there are selected teacher skills which were not marked and which by their very nature must

be present, such as, sensitivity, equal treatment, and motivating students. For example, 19 participants did not mark "sensitivity" indicating that it was "not observed, ok." One is concerned as to why they were not rating "sensitivity" which should be observed and rated in every lesson. Thus, the interpretation of these findings are somewhat clouded with uncertainty.

Limitations

The following factors limited the scope of the investigation:

1. No specific information as to the amount and type of prior training if any in teacher performance evaluation was obtained.
2. No information was obtained as to years of experience or level of education.
3. There was no opportunity for participants to conduct a conference after pretest data were collected.

Recommendations for Further Research

Below are the suggestions and recommendations for further research:

1. An assessment of the effects of training on the participants' knowledge should be made in order to determine the effects of training on the knowledge level

of participants. A pretest/posttest design can be used for this purpose.

2. Teacher evaluation training should place more emphasis on the task of "compensation pay decisions," since the participants were least confident in this task both before and after training.

3. Participants should be given the opportunity to practice their conferencing and feedback skills during the workshop in order to assess the effects of evaluation training on the participants ability to conduct conferences and provide feedback. This information could be obtained using a rating scale developed by the researcher.

4. Provide an opportunity for participants to practice and demonstrate their skill at training. This added activity could be used to determine if an even greater impact on their confidence levels would be realized as a result of this practical experience.

5. Demographic and vital statistics should be ascertained in order to assess the effects of evaluation training by gender, experience, and education level.

6. To increase the validity of the findings relative to interrater reliability, the scale on the instrument must be explained clearly so as to eliminate any difficulty in interpreting the data.

SELECTED BIBLIOGRAPHY

- Acheson, K. (1982). The principal's role in instructional leadership. Eugene, OR: University of Oregon, Oregon School Study Council, Center for Advanced Technology in Education. (ERIC ED 154 940)
- Acheson, K., & Gall, M. (1980). Techniques in the clinical supervision of teachers. New York: Longman.
- Anderson, L. M., Evertson, C. M., & Brophy, J. E. (1979). An experimental study of effective teaching in first-grade reading groups. Elementary School Journal, 79, 193-222.
- Bandura, A., Adams, N. E., & Beyer, S. (1977). Cognitive processes mediating behavioral change. Journal of Personality and Social Psychology, 35, 125-139.
- Bandura, A. (1978). Self-efficacy: 2. Toward a unifying theory of behavioral change. Advances in Behavior Research and Therapy, 1, 139-161.
- Barth, Roland S. (1980, March). How to ensure an effective principalship. The National Elementary Principal, 59, 10-20.
- Bayroff, A. C., Haggerty, H. R. & Rundquist, E. A. (1954). Validity of ratings as related to rating techniques and conditions. Personnel Psychology, 7, 93-113.
- Beach, Dale S. (1980). Personnel: The management of people at work. New York: Macmillan.
- Blumberg, A., & Amidon, E. (1965, September). Teacher perceptions of supervisor-teacher interaction. Administrator's Notebook, 14, 1.
- Blumberg, Arthur. (1970). "Supervisor-teacher relationships: A look at the supervisory conference." Administrator's Notebook, 19, 1-4.
- Blumberg, Arthur. (1980). Supervisors and teachers: A private cold war. Berkeley, CA: Mc Cutchan.
- Bolton, D. L. (1973). Selection and evaluation of teachers. Berkeley, CA: McCutchan.

- Borg, W. R., & Gall, M. D. (1983). Educational research: An introduction. New York: Longman.
- Borich, G., & Madden, S. (1977). Evaluating classroom instruction: A sourcebook of instruments. Reading, MA: Addison-Wesley.
- Borman, W. C., & Dunnette, M. D. (1975). Behavior-based vs. trait-oriented performance ratings: An empirical study. Journal of Applied Psychology, 60, 561-563.
- Boyan, N. J., & Copeland, W. D. (1974). A training program for supervisors: Anatomy of an educational development. Journal of Educational Research, 68, 100-116.
- Brandt, Ron. (1982). On improving teacher effectiveness: 2. A conversation with David Berliner. Educational Leadership, 40, 12-4.
- Bridges, Edwin M. (1986). The incompetent teacher. Philadelphia, PA: Falmer.
- Cawelti, G. (1982). Training for effective school administrators. Educational Leadership, 39, 324-329.
- Edwards, Suzanne. (1985). Development and analysis of a structured data-capturing technique (SDCT) for classroom observation. Dissertation, Iowa State University.
- Faast, D. A. (1982). An analysis of the effectiveness of training teacher evaluators in specific steps in the process of clinical supervision and teacher performance evaluation. Unpublished Doctoral Dissertation, Iowa State University.
- Faast, Dorothy A. & Stow, S. B. (1984). Appraisal training. Clearing House, 58, 128-130.
- Feltz, Deborah L., & Mugno, Denise A. (1983). Replication of the path analysis of the causal elements in Bandura's theory of self-efficacy and the influence of autonomic perceptions. Journal of Sports Psychology, 5, 263-277.
- Goldhammer, R., Anderson, R., & Krajewski, R. (1980). Clinical supervision. 2nd ed. New York: Holt, Rinehart, and Winston.

- Good, T. L., & Grows, D. A. (1979). The Missouri mathematics effective project. Journal of Educational Psychology, 71, 355-362.
- Gordon, Thomas. (1977). Leadership effectiveness training. New York: Wyden Books.
- Griffith, F. (1973). A handbook for the observation of teaching and learning. Midland, MI: Pendell.
- Gudridge, B. M. (1980). Teacher competency: Problems and solutions. AASA Critical Issues Report. Sacramento, CA: American Association of School Administrators. (ERIC ED 182 868)
- Hawley, R. C. (1982). Assessing teacher performance. Amherst, MA: Education Research Associates.
- Hunter, M. (1980). Six types of supervisory conferences. Educational Leadership, 37, 404-412.
- Hunter, Madeline. (1982). Mastery teaching. El Segundo, CA: TIP.
- Hyman, R. (1975). School administrator's handbook of teacher supervision and evaluation methods. Englewood Cliffs, NJ: Prentice-Hall.
- Ishler, P. (1984). Upgrading education means upgrading the teacher evaluation system: Merging evaluation information and effective teaching research--An inservice approach. Paper presented at the Annual Meeting of the Association of Teacher Education. New Orleans, LA. (ERIC ED 241 486)
- Ivancevich, John M. (1982). Subordinates' reaction to performance appraisal interviews: A test of feedback and goal setting techniques. Journal of Applied Psychology, 67, 581-587.
- Iwanicki, E. (1981). Contract plans: A professional growth-oriented approach to evaluating teacher performance. In J. Millman (Ed.), Handbook of Teacher Evaluation. Beverly Hills, CA: Sage.
- Joyce, Bruce, & Showers, Beverly. (1980). Improving inservice training: The messages of research. Educational Leadership, 37, 379-85.

- Kindsvatter, R., & Wilen, W. (1981). A systematic approach to improving conference skills. Educational Leadership, 38, 525.
- Kowalski, Joan P. (Comp.). (1978). ERS Report: Evaluating teacher performance. Arlington, VA: Educational Research.
- Krajewski, R. J. (1976). Clinical supervision to facilitate teacher self-improvement. Journal of Research and Development in Education, 9, 58-66.
- Latham, G. P., Wexley, K. N., & Pursell, E. D. (1975). Training managers to minimize rating errors in the observation of behavior. Journal of Applied Psychology, 60, 550-555.
- Lefton, R. E., Buzzota, V. R., & Sherberg, Manuel. (1981). Improving productivity through people skills. Cambridge: Ballinger.
- Maier, Norman R. F. (1976). The appraisal interview: Three basic approaches. LaJolla, CA: University Associates.
- Manatt, R. (1982). Teacher performance evaluation--Practical application of research. Ames: Iowa State University, Research Institute for Studies in Education. (ERIC ED 225 275).
- McGeoch, Dorothy M.; & Lindsey, Margaret. (1967). Supervisory conferences and the analysis of teaching. The Study of Teaching. Washington, D. C.: Association for Student Teaching.
- McGreal, T. (1980). Helping teachers set goals. Educational Leadership, 37, 414-420.
- McGreal, T. L. (1983), Successful teacher evaluation. Alexandria, VA: Association for Supervision and Curriculum Development.
- Medley, D., & Mitzel, H. (1963). Measuring classroom behavior by systematic observation. In N. L. Gage (Ed.), Handbook of Research on Teaching. Skokie, IL: Rand-McNally.
- Medley, D. M., Coker, H., & Soar, R. S. (1984). Measurement-based evaluation of teacher performance. New York: Longman.

- Meyer, Herbert H., Kay, Emmanuel; & French, John R. P., Jr. (1965). Split roles in performance appraisal. Harvard Business Review, 43, 396-9.
- Millman, Jason. (1981). Introduction. In Jason Millman (Ed.), Handbook of Teacher Evaluation. CA: Sage.
- Nance, Steven R. (1986). Training school administrators in the writing of quality professional improvement commitments. Dissertation. Iowa State University.
- Norvsis, Marija J. (1983). User's guide SPSSX. New York: McGraw Hill.
- Olivero, James L. (1982). Principals and their inservice needs. Educational Leadership, 39, 340-4.
- Redfern, G. (1963). How to appraise teaching performance. Columbus, OH: School Management Institute.
- Redfern, G. (1980). Evaluating teachers and administrators: A Performance Objectives Approach. Boulder, CO: Westview.
- Rice, Ronald. (1986). Effectiveness of a systematic, media facilitated approach to training school administrators to conduct postobservation conferences. Ph.D. dissertation, Iowa State University.
- Robinson, J. (1978). The observational report--A help or a nuisance? NASSP Bulletin, 62 (416), 22-26.
- Rosenberger, D. S., & Plimpton, R. A. (1975). Teacher incompetence and the courts. Journal of Law and Education, 4, 468-486.
- Rosenshine, B., & Furst, N. (1973). The use of direct observation to study teaching. In R. M. Travers (Ed.), Second Handbook on Research on Teaching. Skokie, IL: Rand-McNally.
- Saunders, Mauderie. (1984). The interrelationship of self concept and morale building. The Delta Kappa Gamma Bulletin 28.
- Sergiovanni, T. J. (1982). Clinical supervision and teacher evaluation. Chicago, IL: University of Illinois.

- Shinn, J. L. (1976). Teacher perceptions of ideal and actual supervisory procedures used by California elementary principals. Doctoral Dissertation, University of Oregon.
- Southern Regional Education Board. (1987). More pay for teachers and administrators who do more: Incentive pay programs. Career Ladder Clearing House. Atlanta, GA: SREB.
- Spencer, Karen M. (1985). Post-observation conference: Factors related to success. Dissertation, Iowa State University.
- Stallings, J. (1977). Learning to look. Belmont, CA: Wadsworth.
- Stow, S. B., & Sweeney, J. (1981). Developing a teacher performance evaluation system. Educational Leadership, 38, 538-541.
- Streifer, Philip A. (1987). Teacher evaluation systems: A review of critical issues and the current state of the art. MA: The Regional Laboratory.
- Strike, Kenneth, & Bull, Barry. (1981). Fairness and the legal context of evaluation. In Jason Millman (Ed.), Handbook of Teacher Evaluation. Beverly Hills, CA: Sage.
- Sullivan, John J., & Walker, Joseph J. (1981). Irritating habits - How would your teachers rank you? NASSP Bulletin, 65, 35-7.
- Sweeney, J. (1982). Planning makes a difference: Improving the post-observation conference. NASSP Bulletin, 66, 28-30.
- Sweeney, J. (1983). The post-observation conference: Key to teacher improvement. The High School Journal, 66, 31-37.
- Weber, James R. (1987). Teacher evaluation as a strategy for improving instruction. Eugene, OR: University of Oregon.
- Wickert, Donald M. (1987). Using teacher evaluation for improving instruction. Clearing House, 61, 23-24.

Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1984). Teacher evaluation: A study of effective practices. Santa Monica, California: The Rand Corporation.

ACKNOWLEDGEMENTS

It is with sincere appreciation that I thank all those who helped to make this research a meaningful and productive experience.

An expression of gratitude is extended to my doctoral committee, Dr. R. Manatt, Dr. H. McNab, Dr. D. Robinson, Dr. J. Sweeney, and Dr. R. Warren, whose expert advice and suggestions contributed to the successful completion of this study.

I especially thank Dr. James Sweeney who served most competently as chairman. The challenges he presented through high expectations and a demand for excellence exemplifies the kind of collegial relationship that ideally exists between a major professor and student. His encouragement and confidence in my ability resulted in a rewarding growth experience.

I am deeply grateful to my mother and father for instilling in me the value of education. This entire effort could not have been realized without their love and the love and support from my entire family.

This dissertation is dedicated in memory of my grandmother, Beulah McIntyre, who was a continuous source of strength, understanding, and inspiration.

All praise and glory to the Divine Teacher for the many blessings and capabilities given me to achieve this milestone.

**APPENDIX A -
INSTRUCTIONAL PLANS**

INSTRUCTIONAL PLAN

Title TRAINING TEACHER EVALUATORS Page # 1
 Group or School PROJECT LEAD of 2
 Date(s) MON/TUES, DEC 7/8, 1987
 Attending TRAINERS

Presenting Consultant(s)
SWEENEY
MANATT
STOW

TIME	TOPIC	PRESENTER	MODE	VISUALS	HANDOUTS	REMARKS
8:30	Welcome/Introductions	Sweeney	LGI	Names	Notebooks	--
8:45	Purposes and Overview	Manatt	LGI	O/H	--	Whole Six Days
9:15	Pretest	Stow	IS	Booklet	Scanforms	(1-1)
10:15	Break	OYO	--	--	--	(1-2)
10:30	Pretest (continued)	Stow	IS	Videotape	ASCD Packet	Benchmark "Judy O'Connell er"
11:00	Professional Teacher Performance Evaluation and Training Tips	Manatt	LGI	O/H	Notebook	Improvement (1-3)
12:00	Lunch	LEAD Hosted	--	--	--	--
1:00	Teaching Paradigm	Manatt	LGI	O/H	--	Models to Consider (1-4)
2:00	Break	OYO	--	--	--	--
2:15	Lesson Observation and Analysis (Wide Lens)	Sweeney	LGI	O/H Videotape	Notebook	(1-5)
3:30	Training Questions/ Discussion	Manatt	LGD	O/H	All Materials	--
4:00	Dismissal					

INSTRUCTIONAL PLAN

Title TRAINING TEACHER EVALUATORS Page # 2
 Group or School TRAINERS of 2
 Date(s) TUESDAY, DECEMBER 8, 1987
 Attending TRAINERS

Presenting Consultant(s)
SWEENEY
MANATT
STOW

TIME	TOPIC	PRESENTER	MODE	VISUALS	HANDOUTS	REMARKS
8:30	Review and Preview	Sweeney	LGI	O/H	Notebook	--
8:45	Lesson Observation and Analysis (Wide Lens) Guided Practice I	Sweeney	LGI	O/H	Notebook	Mann
10:00	Break	OYO	--	--	--	
*10:15	Mann Video Debriefing (cont.)	Sweeney	LGI	O/H	Feedback Forms	
10:45	Lesson Observation and Analysis Guided Practice II	Manatt	LGI	O/H Videotape	Notebook	Page I
12:00	Lunch	LEAD Hosted	--	--	--	88
1:00	Lesson Observation and Analysis (Narrow Lens) - Techniques of Data Recording: 1. Classroom Interaction 2. Verbal Flow	Stow	LGI	O/H	Notebook	Curtin
2:00	Break	OYO	--	--	--	
2:15	Data Recording (Wide and Narrow Lens) Guided Practice III	Stow	LGI	O/H Videotape	Notebook	
3:30	Discussion	Manatt/ Sweeney/ Stow	LGD	O/H	All Materials	
4:00	Dismissal					
	*Only 10 minutes when trainers are training.					

INSTRUCTIONAL PLAN

Title Training Teacher Evaluators
 Group or School I-LEAD Project
 Date(s) Mon./Tues., January 25/26, 1988
 Attending Trainers

Page # 1
 of 2

Presenting Consultant(s)
Dr. James Sweeney
Dr. Shirley Stow

TIME	TOPIC	PRESENTER	MODE	VISUALS	HANDOUTS	REMARKS
8:30	Review/Preview	Sweeney	LGI	--	--	Questions
8:45	Preobservation Conference	Stow	LGI/SGD	Overhead	Notebook	--
9:45	Guided Practice I	Sweeney	LGI/SGI	Videotape	Notebook	Have Plans
10:15	Break	OYO	--	--	--	--
10:30	Guided Practice I (continued)	Sweeney	LGI/SGI	Overhead	--	--
11:15	Supervisory Conference	Sweeney	LGI	Overhead	Notebook	Planning Elements
12:00	Lunch	LEAD Hosted	--	--	--	--
1:00	Supervisory Conference	Sweeney	LGI	Overhead	Notebook	Communication Strategies
3:30	Small Group/Discussion/ Questions	Stow/ Sweeney	SGI	--	--	Questions/ Comments
4:00	Dismissal	--	--	--	--	--

INSTRUCTIONAL PLAN

Title Training Teacher Evaluators
 Group or School I-LEAD Project
 Date(s) Mon./Tues., January 25/26, 1988
 Attending Trainers

Page # 2
 of 2

Presenting Consultant(s)
Dr. James Sweeney
Dr. Shirley Stow

TIME	TOPIC	PRESENTER	MODE	VISUALS	HANDOUTS	REMARKS
8:30	Supervisory Conference (continued)	Sweeney	LGI	Overhead	Notebook	Difficult Person
10:00	Break	OYO	--	--	--	--
10:15	Writing Professional Improvement Commitments	Stow	LGI/SGD	Overhead	Notebook	--
12:00	Lunch	LEAD Hosted	--	--	--	-- 06
1:00	From Formative Data to Summative Evaluation	Stow	LGI/SGD	Overhead Videotape	Notebook	--
3:00	Small Group/Discussion/ Questions	Stow/ Sweeney	SGI	--	Notebook	Questions/ Comments/ Concerns
4:00	Dismissal	--	--	--	--	--

INSTRUCTIONAL PLAN

Title Train-the-Trainers
 Group or School I-LEAD Project
 Date(s) February 8 and 9, 1988
 Attending Trainers

Page # 1
 of 2

Presenting Consultant(s)
Dr. Jim Sweeney
Dr. Jerry Herman

TIME	TOPIC	PRESENTER	MODE	VISUALS	HANDOUTS	REMARKS
8:30	Review/Preview	Sweeney	LGD	Overhead	--	
8:45	Legal Aspects	Sweeney	LGD	Overhead	Notebook	
9:30	Documentation	Sweeney	LGD	Overhead	Notebook	
10:00	Break	OYO	--	--	--	
10:15	Documentation (continued)	Sweeney	LGD	Overhead	Notebook	
10:45	Supervising the Marginal Employee	Sweeney	LGD	Overhead	Notebook	
12:00	Lunch	LEAD Hosted	--	--	--	
1:00	Administrator Performance Evaluation - Purpose and Process	Herman	LGD	Overhead	Notebook	
1:45	- District, Unit, Personal, and Growth Goals	Herman	LGD	Overhead	Notebook	
2:15	Break	OYO	--	--	--	
2:30	- The Essential Elements	Herman	LGD	Overhead	Notebook	
3:15	- Data Gathering	Herman	LGD	Overhead	Notebook	
3:45	Questions	Sweeney/ Herman	LGD	--	--	

INSTRUCTIONAL PLAN

Title Train-the-Trainers Page # 2
 Group or School I-LEAD Project of 2
 Date(s) February 8, and 9, 1988
 Attending Trainers

Presenting Consultant(s)
Dr. Jim Sweeney
Dr. Jerry Herman

TIME	TOPIC	PRESENTER	MODE	VISUALS	HANDOUTS	REMARKS
8:30	Review/Preview	Sweeney	LGD	--	--	
8:45	Observation	Herman	LGD	--	--	Julie Brown Videotape
10:00	Break	OYO	--	--	--	
10:15	Written Work Samples/ Other Data/Logs	Herman				
11:00	Summative Evaluation	Herman				92
11:30	Wrap-up	Sweeney				
12:00	Lunch	LEAD Hosted	--	--	--	
1:00	Post-Test	Sweeney				

**APPENDIX B -
HUMAN SUBJECTS MODIFIED INFORMED CONSENT STATEMENT**

**MODIFIED INFORMED CONSENT STATEMENT TO BE USED
WITH LEONARD MCINTYRE'S
THE EFFECTS OF I-LEAD PERFORMANCE APPRAISAL TRAINING ON THE
KNOWLEDGE, SKILLS, AND CONFIDENCE OF EVALUATORS**

**Ph.D Dissertation
Jim Sweeney
Advisor**

The following statement will be read to each participant:

"Today you will participate in training to improve your evaluation knowledge, skills, and confidence. Because the data collected during training will compare your previous knowledge, skills, and confidence with same after having participated in training, you have a right to refuse to participate in the pre- and post- testing. Your decision to participate in this training is greatly appreciated as most educators want to improve their abilities in the area of conducting evaluations. If you are willing to take part in this undertaking please turn in your materials at the close of the exercises.

Submitting the materials will be construed as a modified consent to participate.

**APPENDIX C -
TPE RATING SCALE**

**APPENDIX D -
THE SUPERVISORY ATTITUDE SURVEY**

SS# _____-_____-_____

SUPERVISORY ATTITUDE SURVEY

DIRECTIONS: This survey is designed to help us examine your confidence level in evaluating teachers and in training teacher evaluators. It is important information for your workshop facilitators. Please be sure to fill out your social security numbers as it is important for matching purposes only. Please read each statement carefully and circle the number which reflects your level of confidence. Please use the following response scales:

1	2	3	4	5	6	7	8	9
Strongly		Disagree		Not	Agree		Strongly	
disagree				sure			agree	

Please Circle One Response

S D D N A S A

As a teacher evaluator:

- | | | | | | | | | | |
|--|---|---|---|---|---|---|---|---|---|
| 1. I find it difficult to make decisions about <u>what to record</u> when observing a lesson. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 2. I find it difficult to <u>observe and record important teacher and student behaviors and other occurrences</u> observed during the lesson. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 3. <u>Identifying specific important areas</u> which would help the teacher to improve his or her classroom performance is difficult for me. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 4. I feel confident I can <u>interpret the data</u> I have recorded when observing the lesson. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 5. I feel confident I can <u>use the data</u> I have recorded in providing specific examples for feedback to help the teacher improve. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

E D D N A E A

As a teacher evaluator:

- 6. I feel confident I am able to make good decisions about the design used in the lesson. 1 2 3 4 5 6 7 8 9
- 7. I feel confident I understand how to make decisions about the effectiveness of the teaching strategies or methods employed during a lesson. 1 2 3 4 5 6 7 8 9
- 8. I feel confident that, given the evaluation skills I now have and the process I have been trained in, I can use them to significantly improve instruction. 1 2 3 4 5 6 7 8 9
- 9. I feel confident, given sufficient observations, I can make valid decisions about teachers' compensation based on my evaluation of their performance. 1 2 3 4 5 6 7 8 9
- 10. I feel confident, given sufficient observations, I can make a valid decision about whether or not to recommend a teacher for tenure and provide adequate documentation. 1 2 3 4 5 6 7 8 9
- 11. I feel confident, given the evaluation skills I now have using the evaluation process I have been trained in, I can maintain a positive supervisory relationship and improve instruction.. 1 2 3 4 5 6 7 8 9
- 12. I feel confident in my observation and lesson analysis skills. 1 2 3 4 5 6 7 8 9

S D D N A S A

I feel confident that I can train teacher evaluators:

- | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 13. to make good decisions about <u>what to record</u> when observing a lesson. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 14. to be proficient at <u>observing and recording important teacher/student behaviors and other occurrences</u> observed during the lesson. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 15. to be proficient at <u>identifying the specific important areas</u> which would help the teacher to improve his or her classroom performance. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 16. to <u>interpret the data</u> they have recorded when observing the lesson. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 17. to <u>use the data</u> they recorded in providing specific examples for feedback to help the teacher improve.. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 18. to make good decisions about the <u>design used in the lesson</u> | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 19. to understand how to make good decisions about the <u>effectiveness of the teaching strategies or methods</u> employed during a lesson. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 20. given the training skills I now have, to have the necessary <u>knowledge and skills</u> to improve instruction. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 21. given sufficient observations, to <u>make valid decisions about teachers' compensation</u> based on their evaluation of teachers' performance. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 22. given sufficient observations, to make valid decisions about <u>whether or not to recommend a teacher for tenure</u> and provide <u>adequate documentation</u> | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 23. given the training skills I have now, to <u>maintain a positive supervisory relationship and improve instruction</u> .. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 24. to significantly improve their <u>observation and lesson analysis skills</u> in order to improve classroom instruction. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |