

1-1-1984

What makes a difference in inservice teacher education : a meta-analysis of the research.

Ruth Konhaus Wade

University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_1

Recommended Citation

Wade, Ruth Konhaus, "What makes a difference in inservice teacher education : a meta-analysis of the research." (1984). *Doctoral Dissertations 1896 - February 2014*. 4148.

https://scholarworks.umass.edu/dissertations_1/4148

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations 1896 - February 2014 by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

FIVE COLLEGE
DEPOSITORY

What Makes a Difference in Inservice Teacher Education:
A Meta-Analysis of the Research

A Dissertation Presented

By

RUTH KONHAUS WADE

Submitted to the Graduate School of the
University of Massachusetts in partial fulfillment
of the requirements for the degree of

DOCTOR OF EDUCATION

February 1984

Education

© Ruth Konhaus Wade
All Rights Reserved

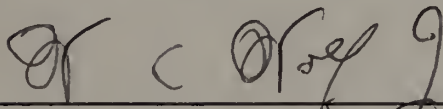
What Makes a Difference in Inservice Teacher Education:
A Meta-Analysis of the Research

A Dissertation Presented

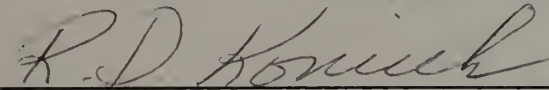
By

RUTH KONHAUS WADE

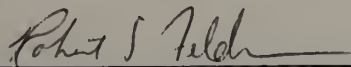
Approved as to style and content by:




W. C. Wolf, Jr., Chairperson of Committee



Richard Konicek, Member



Robert S. Feldman, Member



Mario Fantini, Dean
School of Education

ACKNOWLEDGEMENTS

This dissertation is in one sense an individual effort and in another sense a collective one. To those who contributed to the collective effort, I wish to express my sincerest appreciation.

Bill Wolfe for his support, encouragement, and advice throughout my doctoral program.

Dick Konicek and Bob Feldman for serving on my dissertation committee and for their helpful comments and suggestions.

Eva Goldwater for providing crucial support in the processing and analysis of the data.

The University of Massachusetts Interlibrary Loan staff for assistance in obtaining dissertations.

Mary Carlson for the expert editing and typing of this manuscript. Her attention to detail and deadlines ensured an early completion of the final draft of this manuscript.

Genie Hersh, who has run over 8,000 miles with me during the completion of my doctoral work, has provided daily 6:00 AM briefings which have kept me abreast of current events.

Jennifer, Courtney, and Alison for living with a particularly "driven" mother during the past year. They helped me to retain a balanced perspective regarding the

importance of work and family.

George for most everything.

ABSTRACT

What Makes a Difference in Inservice Teacher Education:
A Meta-Analysis of the Research

February, 1984

Ruth Konhaus Wade, B.S., Pennsylvania State University
Ed.D., University of Massachusetts

Directed by: Professor W. S. Wolf, Jr.

The purpose of this study was to gather existing quantitative data on inservice teacher education in order to analyze and synthesize the findings. Data were gathered on ninety-one research studies presented between 1968 and 1983 that were available through the ERIC system, dissertations, or journals. Meta-analysis was used to draw generalizations regarding the efficacy of various inservice practices. Effectiveness was measured at four different effect levels: participants' reactions to training, participants' learning, behavior change of participants, and results in terms of participants or their students.

Findings indicate that inservice teacher education programs reported in the literature are moderately effective. When the data are grouped by effect level, it becomes apparent that attempts to increase participants'

learning through inservice teacher training are highly effective, attempts to change participants' behavior and to elicit positive reactions are moderately effective; while attempts to demonstrate results in the school environment are only mildly effective.

Specific findings are: (1) the number of participants in an inservice training program, the number of treatment hours, and the length of the treatment period do not significantly influence effect size results; (2) outside originated programs are generally more effective than in-school originated ones; (3) inservice training programs which include both elementary and secondary educators are more effective than for either group individually; (4) enhanced status and college credit are the incentives most likely to increase effect size results; and (5) training programs which use observation, micro teaching, video/audio feedback, or practice show greater effects than those programs not using these methods. Programs which included discussion, lecture, games/simulations, and guided field trips were significantly less effective than those using other instructional methods. Of course these findings do not preclude the possibility that these training methods could be more effective under certain circumstances. Coaching, modeling, mutual assistance, printed material, production of instructional material, and film as used were

not associated with significant effects.

The results of this study were used to suggest implications for staff developers as well as to make suggestions for future study.

TABLE OF CONTENTS

ACKNOWLEDGEMENT	iv
Chapter	
I. INTRODUCTION	1
The Problem	3
Purpose of the Study	5
Reaction	6
Learning	6
Behavior	6
Results	7
Specific Objectives	9
Significance	10
Limitations	11
Methodology	14
Definition of Terms	16
II. REVIEW OF RELATED LITERATURE	19
Introduction	19
Historical Perspective	20
Preservice Education	20
Inservice Education	21
Summary	27
Theoretical Framework	30
Levels of Evaluation	36
Reaction	37
Learning	38
Behavior	40
Results	42
The Evaluation Matrix	44
Research Evidence	45
Context	47
Design	51
Duration	51
Location and Scheduling	52
Focus and Nature	54
Methods of Instruction	55
Process	56
Summary	63
III. METHODOLOGY	64
Identifying and Collecting the Studies	64
Identifying, Operationally Defining and Coding	

The Variables	66
Meta-Analysis	67
Calculating Effect Size	68
Computer Processing	72
Analyzing the Data	73
IV. PRESENTATION OF THE DATA	75
Description of the Data Base	75
Grand Mean Effect Size Results	76
Grand Effect Size by Dependent Variables	77
Analysis by Independent Variables	78
Instructional Focus	79
Source	79
Number of Participants	79
Location	86
Initiator	86
Grade Level Taught	89
Participation	89
Participant Incentives	92
Funding	95
Instructor	95
Type of Structure	100
Length of Treatment	103
Emphasis	103
Schedule	106
Duration	109
Activities	109
Participant Role	109
Goals	113
Group Focus	113
Status	113
Follow-Up	117
Assignment to Groups	117
Internal Validity	120
Reactivity Level	120
Source of Measurement	124
Type of Measurement	126
Effect Formula	128
Types of Instruction	130
Summary	132
V. DISCUSSION AND IMPLICATIONS	135
Analysis by Dependent Variables	135
Reaction	136
Learning	137
Behavior	137
Results	138
Conclusions	139

Staff Development Research Findings	140
Dependent Variables	140
Duration	142
Location and Scheduling	143
Process	143
Focus and Nature	146
Instruction	147
Implications for Staff Developers	150
Recommendations for Future Study	152
.	
REFERENCES	156
APPENDIX A	170
APPENDIX B	186
APPENDIX C	190
APPENDIX D	205
APPENDIX E	211

LIST OF TABLES

1.	Relative Effect Size	68
2.	Source of Research Studies	76
3.	Grand Mean Effect Size	77
4.	Grand Mean Effect Size by Dependent Variable Measures	78
5.	Mean Effect Size by Instructional Focus by Effect Level	80
6.	Analysis of Variance Results for Instructional Focus and for Effect Level	81
7.	Mean Effect Size by Source by Effect Level	82
8.	Analysis of Variance Results for Source and for Effect Level	83
9.	Mean Effect Size by Number of Participants by Effect Level	84
10.	One-Way Analysis of Variance Results for Number of Participants	85
11.	Mean Effect Size by Location by Effect Level	87
12.	Analysis of Variance Results for Location and for Effect Level	87
13.	Mean Effect Size by Initiator by Effect Level	88
14.	One-Way Analysis of Variance Results for Initiator	89
15.	Mean Effect Size by Grade Level Taught by Effect Level	90
16.	Analysis of Variance Results for Grade Level Taught by Effect Level	91
17.	Mean Effect Size by Participation by Effect Level	91
18.	Analysis of Variance Results for Participation by Effect Level	92
19.	Mean Effect Size by Participant Incentive by Effect Level	93
20.	One-Way Analysis of Variance Results for Participant Incentive	94
21.	Mean Effect Size by Funding by Effect Level	96
22.	One-Way Analysis of Variance Results for Funding	97
23.	Mean Effect Size by Instructor by Effect Level	98
24.	One-Way Analysis of Variance Results for Instructor	99
25.	Mean Effect Size by Type of Structure by Effect Level	101
26.	One-Way Analysis of Variance Results for Structure	102
27.	Mean Effect Size by Length of Treatment by Effect Level	104
28.	Analysis of Variance Results for Length of	

	Treatment and for Effect Level	105
29.	Mean Effect Size by Emphasis by Effect Level . . .	105
30.	Analysis of Variance Results for Emphasis and for Effect Level	106
31.	Mean Effect Size by Schedule by Effect Level . . .	107
32.	One-Way Analysis of Variance Results for Schedule	108
33.	Mean Effect Size by Duration by Effect Level . . .	110
34.	Analysis of Variance Results for Duration and for Effect Level	110
35.	Mean Effect Size by Activities by Effect Level . .	111
36.	Analysis of Variance Results for Activities and for Effect Level	111
37.	Mean Effect Size by Participant Role by Effect Level	112
38.	Analysis of Variance Results for Participant Role and for Effect Level	112
39.	Mean Effect Size by Goals by Effect Level	114
40.	Analysis of Variance Results for Goals and for Effect Level	114
41.	Mean Effect Size by Group Focus by Effect Level	115
42.	Analysis of Variance Results for Group Focus and for Effect Level	115
43.	Mean Effect Size by Status by Effect Level	116
44.	Analysis of Variance Results for Status and for Effect Level	116
45.	Mean Effect Size by Follow-Up by Effect Level . .	118
46.	Analysis of Variance Results for Follow-Up and for Effect Level	118
47.	Mean Effect Size by Assignment to Groups by Effect Level	119
48.	One-Way Analysis of Variance Results for Assignment to Groups	120
49.	Mean Effect Size by Internal Validity by Effect Level	121
50.	Analysis of Variance Results for Internal Validity and for Effect Level	122
51.	Mean Effect Size by Reactivity Level by Effect Level	123
52.	Analysis of Variance Results for Reactivity Level and for Effect Level	124
53.	Mean Effect Size by Source of Measurement by Effect Level	125
54.	One-Way Analysis of Variance Results for Source of Measurement	126
55.	Mean Effect Size by Type of Measurement by Effect Level	127
56.	One-Way Analysis of Variance Results for Type of Measurement	128

57.	Mean Effect Size by Effect Formula by Effect Level	129
58.	Analysis of Variance Results for Effect Formula and for Effect Level	130
59.	Mean Effect Size by Type of Instruction	131
60.	A Comparison of Mean Effect Size by Effect Level as Determined by Wade, Joslin, and Lawrence and Harrison	141

LIST OF FIGURES

1.	Characteristics of the Four Evaluation Criteria	8
2.	Brethower and Rummeler's Evaluation System	33
3.	Brethower and Rummeler's Systems Model	36
4.	Operational Evaluation Model	46

C H A P T E R I

INTRODUCTION

The professional development of teachers through locally-developed inservice education activities has begun to emerge as a potentially valuable means for improving schools, upgrading the skills of educators, and providing opportunities for personal growth. Issues related to inservice education have not only received a great deal of interest in the past five to seven years, but federal and state dollars have also been appropriated to expand and improve professional development programs and practices.

There appear to be many reasons for the increased attention given to inservice education. Declining student enrollments and resulting budget reductions coupled with reduced teacher turnover have greatly limited the hiring of new teachers. Previously, new ideas and innovations were introduced into a school district when it annually hired a substantial number of new teachers. A related source of stimulation has been increased interest from undergraduate and graduate teacher training institutions which, due to declining enrollments, have had to turn their attention from preservice training to inservice training in order to

economically survive. Third, even though many dollars were expended to improve education from 1965 to 1975, little change resulted. Subsequently, a great deal of attention was focused on understanding the implementation process with the resulting conclusion that change is more a function of people, process, and organizations rather than of technology. This, in turn, focused greater attention on the school and local inservice programs as an important element in the change process. Finally, the popular press along with just about everyone else is decrying the state of public elementary and secondary education throughout the United States. Test scores have declined, standards have been lowered, and the teaching profession has failed to attract enough qualified people to its ranks. An April, 1983 status report on American education by the National Commission on Excellence in Education suggests through its title "A Nation at Risk: The Imperative for Educational Reform" that our education system is lacking. The consensus throughout the country seems to be that public education needs to be upgraded, because our students cannot compete with those from other industrialized nations.

In the 1980s, there is no indication that inservice education has experienced the loss of momentum that has characterized most innovations in education. Rather, there is continued attention given to and acceptance of inservice

education as the means of improving education. Evidence of this trend includes the formation of two national organizations in the past decade, the National Staff Development Council and the National Council of States on Inservice Education, a proliferation of books, doctoral dissertations, articles on the topic, and a growth of entries in the ERIC Clearinghouse on Teacher Education from 938 entries for "inservice education" in 1975 to 11,127 in 1982.

The Problem

Trying to keep abreast of the rapidly growing research on inservice education is a nearly impossible task. Even if one can keep up with it, the results reported are often contradictory and confusing. The research techniques used, the measurements taken, and the groups studied may vary greatly from one study to the next. Drawing conclusions from such a heterogeneous conglomeration can lead to frustration and further confusion. Yet most research reviews and integrative works continue to be largely a pattern of reviewers' personal judgements, individual creativity, and preferred styles (Jackson, 1978).

Edelfelt (1981) reports that even though many more people are writing about staff development, telling the whole story is a rare occurrence. Loucks et al. (1982) report that most accounts of staff development are simply

statements of participant satisfaction which are then used to determine the success of a program. Most professional development programs have narrow and short-term objectives that are unrelated to a larger purpose or rationale. Edelfelt (1975) characterized the existing approach to staff development as piecemeal and the result as patchwork. While participant satisfaction and local support are invaluable to inservice programs, there is a need to systematically determine the efficacy of various inservice practices. Effectiveness needs to be measured not only at the level of the teacher-participant, but also at the level of the students with whom the teacher-participants interact.

What is needed is a systematic method for integrating findings across independent studies by converting them to a common base. Integrative analysis, or what has been termed meta-analysis by Glass (1976), provides the necessary perspective for this systematic integration.

Meta-analysis refers to the analysis of analyses. I use it to refer to the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings. It connotes a rigorous alternative to the causal, narrative discussions of research studies which typify our attempts to make sense of the rapidly expanding research literature.
(p. 3)

Glass and Smith (1980, p. 2) view meta-analysis "not

as a technique" but rather as "a perspective that uses many techniques of measurement and statistical analysis." Meta-analysis will be used in this study to provide the necessary "perspective" to statistically review the experimental literature on staff development in order to determine the efficacy of various staff development practices.

Purpose of the Study

The proposed project will be an effort to gather existing quantitative data on inservice education published or presented over the past fifteen years in order to analyze and synthesize the findings. Meta-analysis will be used to draw generalizations that indicate the types of policies, formats, materials, leadership roles and styles, and procedures that are most effective in promoting the professional development of educational personnel.

In attempting such an integration of research findings, it is important to start with a clear framework from which staff development can be examined and synthesized. The "Kirkpatrick approach to training evaluation" has served as the classic model for trainers in the field of business and industry over the past two decades and provides the framework needed here. This model, introduced by Kirkpatrick (1959) conceptualizes different levels for

evaluating training including reaction, learning, behavior, and results.

Reaction

The reaction level assesses how the participants in staff development activities feel about those activities. An evaluation of a workshop, for example, could assess participants' feelings regarding such factors as the enthusiasm of the workshop leader, the use of visual aids, the degree to which workshop objectives were accomplished, and so forth. Evaluation at the reaction level is the easiest to conduct compared to the other levels. It also yields the lowest informational value and is the most frequently used.

Learning

This level is concerned with measuring the amount of learning that resulted from a staff development activity. Kirkpatrick (1967) suggests that objectivity be stressed in measuring learning and quantitative measures be used.

Behavior

The focal point of evaluation at the behavior level is whether or not participants change their behavior as a result of a staff development intervention. Kirkpatrick

(1967) reports that evaluation at this level yields more useful information, but is more difficult to assess. He suggests gathering these data in the work setting through direct observation, using performance-based evidence.

Results

The goal at this level is to be able to measure outcomes or results as an effect of staff development. Most often this translates into determining the effects of staff development on students or the working environment. Kirkpatrick (1967) reports that assessing the results of in-service education is the most difficult of the four levels to accomplish yet yields the most valuable information. See Figure 1 for a summary of the characteristics of the four levels of evaluation.

These four levels of evaluation: reaction, learning, behavior, and results will serve as the dependent variables in this study. The meta-analysis will be directed toward answering four questions:

1. When the purpose of staff development is to achieve positive reactions, what policies, formats, materials, leadership roles and styles, and procedures are most

Figure 1
 Characteristics of the Four Evaluation Criteria

<u>Criteria</u>	<u>Value of Information</u>	<u>Frequency of Use</u>	<u>Difficulty of Assessment</u>
A. Reaction	Lowest	Relatively	Relatively
B. Learning	↓	Frequent	Easy
C. Behavior		↑	↓
D. Results	Highest	Relatively Infrequent	Relatively Difficult

(Source: Newstrom, 1978)

-
- effective in promoting the professional development of educators?
- When the purpose of staff development is increased learning, what policies, formats, materials, leadership roles and styles, and procedures are most effective in promoting the professional development of educators?
 - When the purpose of staff development is behavior change, what policies, formats, materials, leadership roles and styles, and procedures are most effective in promoting the professional development of educators?
 - When the purpose of staff development is results, what

policies, formats, materials, leadership roles and styles, and procedures are most effective in promoting the professional development of educators?

Specific Objectives

The primary objective of this study is to generate research-based evidence that can be used to improve staff development practices. Specific objectives are:

1. To establish relationships between quantitative research outcomes pertaining to staff development programs and instructional methods used routinely as part of staff development programs.
2. To determine the magnitude of effect of specific types and aspects of programs on reactions, learning, behavior change, and results.
3. To ascertain the feasibility of synthesizing effect size across studies in order to generate suggestions for overall staff development and specific inservice planning that are effective in promoting teacher growth.

Significance

The investments in and potential rewards of staff development have become too great to allow a continuation of the "head in the sand approach." As Edelfelt observed (1981, p. 113), "the volume of printed words about inservice education from 1974 to 1980 far exceeds that from 1968 to 1974." But the important questions such as: What pattern of variables increases the effectiveness of inservice training?; Has teaching been improved through inservice training?; Is the curriculum better?; are questions that generally remain unanswered. We need to go beyond affective measures that consist of teachers reporting, "This workshop was a valuable experience that will make a difference in my teaching," to look at what program policies, formats, materials, resource personnel and procedures are effective in promoting learning, behavior change, and results in the classroom. Without substantial data, we are left with observations, impressions, and imprecise generalizations. Throughout the staff development literature these kinds of casual observations are omnipresent.

There is a need to refocus our attention on the quantitative research that has been done in order to statistically analyze the findings and reach nonarbitrary conclusions. Meta-analysis provides the vehicle through which

the large mass of data on staff development can be systematically organized, sifted through, and findings reported that are research-based as well as generalizable across independent studies. There is a need to refocus our attention on some of the basic methodological issues affecting staff development research. With meta-analysis, the findings are not judged a priori or by arbitrary criteria, but are instead integrated to apply the full power of statistical analysis.

It is not uniformity in research reviewing and integrating that is desirable, rather it is clarity, explicitness, and openness--those properties that are characteristic of the scientific method more generally and which impart to inquiry its "objectivity" and trustworthiness. (Glass and Smith, 1980, p. 14)

Limitations

Social and behavioral research is a large and widely scattered enterprise of uneven quality. Obtaining knowledge through this type of research is usually a time-consuming, arduous task with limited outcomes. The social and behavioral scientist must deal with many variables simultaneously. Rigid control of these variables is difficult at best, and quantification of these variables is often hard to achieve. Because of the difficulty of controlling and quantifying these variables, and because much

of educational research is qualitative in nature, a meta-analysis of educational research on any particular topic will be inexact and incomplete. Other limitations include the fact that research techniques vary widely in approach and accuracy. Likewise, measurements vary in approach, precision, and accuracy. Populations differ in size and composition. In short, educational research by its very nature is imprecise and uneven in quality.

Beginning with the realization that meta-analysis draws its findings from the imprecise data of uneven quality that researchers report, the technique itself has some delimitations. Meta-analysis uses only quantitative data without pre-judging the quality of them. All qualitative judging is addressed a posteriori in terms of the covariance of study findings. In other words, if "good" (i.e., those with good controls and technology) and "bad" studies reported the same findings, then a large data base (all studies regardless of quality) is preferable to a small data base (only the "good" studies). However, if there is a discrepancy in findings between "good" and "bad" studies, then the "good" studies certainly are to be believed (Glass and Smith, 1980).

A caution regarding meta-analysis is the fact that it relies on summary findings from individual studies rather than using original data from individual studies. The most

desirable approach would be to pool the original data from studies, but unfortunately many informative studies do not report the raw data, only summary findings. This leaves the researcher to work with what is available.

Of all the technical criticisms of meta-analysis, Glass and Smith (1980) give most credence to the concern that

...meta-analyses are conducted on large data sets in which multiple results are derived from the study; this renders the data non-independent and gives one a mistaken impression of the reliability of the results.... For example, if Study #1 gave effects .2, .2, .2, and .2 and Study #2 gave effects .6, .6, and .6, one would have little reason to believe that he had been informed seven times about the aggregate result in question; rather the true "degrees of freedom" would seem to be somewhat closer to 2, the number of studies, than to 7, the number of effects. (p. 299)

Glass and Smith suggest in such cases that averaging all findings within a study and using "studies" as the unit of analysis could serve as a solution to this problem. However, they also concede that this solution will likely obscure many important questions that can only be addressed at the "within study" level of outcome variables (Glass and Smith, 1980).

Methodology

Meta-analysis differs from traditional research integration. In traditional research integration, quantitative and qualitative studies are gathered and judgements are made of study designs, which form the basis for arbitrary decisions to include or exclude studies. From this information, general conclusions are drawn based upon the relative weight given each study. In contrast, meta-analysis requires gathering quantitative studies that fit the topical categories under review. These studies are not prejudged in terms of research quality but rather are all statistically analyzed in order to integrate the different findings of individual studies. Glass and Smith (1980) reviewed various alternatives to meta-analysis. They concluded that meta-analysis offers an improvement over traditional methods of research integration, because it removes sources of arbitrariness to arrive at an impartial as well as representative view of existing research.

Meta-analysis begins with an exhaustive literature search. The primary aim in meta-analysis is to provide accurate, impartial, and quantitative findings on a population of studies or a specific topic (Glass and Smith, 1980). The studies should represent both published and unpublished data. Studies are not excluded due to lack of

rigor. In fact, all studies that are quantitative, have the necessary statistics presented, and meet the topical criteria are included.

After research studies are identified, each study must be described, classified and coded. This is done in order to relate the properties of the studies to the study findings.

In the proposed study, research that deals with professional development programs or activities will be gathered. These studies will no doubt have diverse objectives. These objectives will be categorized into one or more of the following categories: reaction, learning, behavior, and results. The meta-analysis will focus on how these dependent variables (reaction, learning, behavior, and results) were affected by the independent variables (policies, formats, materials, leadership roles and styles, and procedures) examined in each study.

In the studies to be examined, the researchers will have different objectives and use different measurements to determine effect. With meta-analysis, the studies can be compared by making them part of a larger class or category--effects of professional development programs. By standardizing the scores, characterizing each as the size of the effect in relation to the standard deviation, the various sets of scores can be put on the same scale.

This formula:
$$\frac{\bar{X}_E - \bar{X}_C}{S_x}$$

tells us that the effect size equals the mean of the experimental group minus the mean of the control group divided by the standard deviation of the control group (Glass and Smith, 1980, p. 136). Various strategies are employed to calculate effect size depending upon the summary statistics available.

After effect size calculations are made, statistical analyses can be used to identify the independent variables that account for the changes. From this meta-analysis, those variables which contribute the most toward program effectiveness will be identified. In other words, the independent variables that have the highest correlation with dependent variables will be selected. From these findings, suggestions for overall staff development and specific inservice practices should evolve.

Definition of Terms

In scanning the literature on staff development, one finds a wide variety of types of inservice as well as topics addressed under the umbrella of staff development. To add further confusion, staff development is called by a number of names including: inservice education, inservice

training, professional development, professional growth, teacher renewal, retraining, and personnel development to name some of the more common terms. To compare and contrast the myriad of definitions presented for these terms would be a futile exercise. In fact, no one set of definitions is right or wrong. The important task is to articulate a clear workable definition that suits the purposes of this study and can be used throughout the paper.

For the purposes of this study certain terms have been defined as follows:

Inservice education will refer to components of a staff development program such as workshops, lectures, seminars, or other similar activities that represent offerings within a program but do not characterize it as a whole.

Staff development will refer to a series of "systematically designed activities planned to increase the competencies ... knowledge, skills, and attitudes ... needed by school personnel in the performance of their assigned responsibilities" (Orrange and Van Ryn, 1975, p. 47).

Professional development will be defined as "all that a professional person does to gain a better understanding and a wider scope in his professional outlook" (Perritt, 1972, p. 324).

Meta-analysis will refer to "the integration of the

findings of many empirical research studies of a topic" (Glass and Smith, 1980, intro.). The integration has, as its purpose, quantifying outcomes from each study in order to determine the magnitude of the effect.

Effect size will be defined as the mean difference between treated and control groups divided by the standard deviation of the control group. When the necessary data are not presented, effect size may also be a conservative estimate of effectiveness based on reported statistics and assuming certain characteristics.

C H A P T E R I I
R E V I E W O F R E L A T E D L I T E R A T U R E

Introduction

The rationale for this study is based upon the need for evaluative data to indicate the efficacy of various inservice practices. An examination of the staff development literature soon reveals two generalizations that, with few exceptions, apply to existing studies of staff development: (1) carefully executed and methodologically rigorous studies are uncommon, and (2) much of what exists are laudatory reports that are not very useful (Edelfelt, 1981).

This review is organized around issues which I believe are critical considerations in the planning, conduct, and evaluation of staff development programs, if the desired results are programs with credibility, utility, and impact upon practice. The discussion that follows is divided into three categorical issues: (1) historical overview, provides the framework and perspective; (2) theoretical framework, offers conceptualizations to guide decision processes and information requirements in the evaluative process; and (3) research evidence, identifies effective practices for

the planning and conduct of staff development.

Historical Perspective

Preservice education

Prior to 1830, the public schools in the United States were staffed by teachers who in many cases were poorly educated themselves (Knight, 1922). Teaching was a temporary vocation, a part-time job not usually considered a career. For men, it was often a first-step on the way to becoming a minister; for women, it provided an acceptable position between school and marriage (Edelfelt and Johnson, 1980).

Around the year 1820, the educational philosophies and practices of Prussia and France attracted American attention. Specialized teacher training institutes were developing there that focused on the nature of the child. One American, Samuel R. Hall of Vermont, responded by establishing the first American normal school, a private one, in 1823. Slowly the idea grew until Horace Mann and others were able to help establish the first public normal school in July, 1839 in Lexington, Massachusetts. The idea of publicly supported state institutions for teacher training spread steadily throughout the United States during the 1840's and 1850's (Edelfelt and Johnson, 1980).

The early normal schools, which primarily trained elementary school teachers, offered training courses ranging from a few weeks to one year of study. Classes dealt largely with a review of elementary subjects including the Bible, orthography, and reading with an occasional emphasis on methods of teaching and general pedagogy. There were also opportunities to practice teach in model schools (Edelfelt and Johnson, 1980).

Until 1894, students entering normal schools were admitted directly from elementary school (eighth grade). Again, Massachusetts set the trend when it made high school graduation a requirement for entrance to normal schools in 1894 (Edelfelt and Johnson, 1980).

Shortly after public normal schools were well established, teacher training began to be included as a role of colleges and universities. When the first chair of education was established at the University of Iowa in 1873, the focus was on training secondary teachers. By the 1930's, however, normal schools began to become teachers' colleges and nearly 70 percent of all teachers had attended at least two years of college (Joslin, 1980).

Inservice Education

Parallel to the development of preservice education,

but trailing by a few years, was the evolution of inservice education. According to Asher (1967), teacher institutes were initiated because so many teachers were inexperienced and untrained in subject matter. The first institute was organized in 1839 by Henry Barnard, and provided a six-week session focusing on teaching methods. Massachusetts again led the way and became the first state to support teachers' institutes.

Despite the tremendous need for institutes and the teacher training they provided for nearly a century, many problems plagued the institutes. Frazier, et al. (1935) detail the difficult conditions including poor organization, lack of in-depth study, little continuity, lack of individualization, inadequate facilities, limited materials, and poor teaching methods. It is of little comfort to realize that most of these problems still exist in inservice education today.

When teacher institutes did not fulfill their promises, other approaches to inservice began to appear. Teacher reading circles were organized on a state-wide basis with reading lists and study guides provided for teachers (Richey, 1957). By the beginning of the twentieth century, inservice training had taken a new direction with normal schools, colleges, and universities becoming the major delivery vehicle for inservice education through

summer, evening, and Saturday courses.

The issue of transferring theoretical instruction into practice was of concern at least as early as 1904, when Dewey expressed the need to relate theory to practice. Dewey's concern was that lecturers in the teacher training establishments were not able to make "... real and vital theoretical instruction" (Dewey, 1904, p. 9). Asher (1967, p. 4) concurs "... lecturers in the institutes talked about pedagogical principles but violated them by preaching activity while the audience was strictly passive." This concern, with the transfer of training from inservice training sessions to classroom situations is, according to Fullan (1982), the most fundamental problem in inservice education today.

During the 1920s, inservice education began to be dominated by state regulations. In the post-World War I period, pressure mounted for higher standards in teacher education. The response was to revise and update existing subject matter, and to raise certification requirements for secondary teachers and later for elementary teachers. Between 1926 and 1937, thirty-two states stipulated from one to four years of college as a prerequisite for certification (Asher, 1967).

The 1930s brought economic difficulties for most Americans. Many students stayed in school because job oppor-

tunities for them were limited. Again, there was a need for inservice training for teachers who needed to provide a more vocationally oriented curriculum to meet the needs of the times.

The Eight Year Study, begun in 1933, focused on thirty school systems whose assigned task was to develop and implement new educational programs which could be disseminated for high school students throughout the country. Although this project was a refreshing approach to inservice training, it did not provide the curricula or new approaches that were anticipated. The most significant contribution of the Eight Year Study, according to Tyler (1971), was the educational opportunity it provided teachers for problem solving and for developing skills of educational inquiry. The Eight Year Study brought to inservice education the realization that teachers could serve as their own change agents. Although drastic changes did not come in the design of subsequent inservice programs, the concept that teachers rather than "authorities" could determine the purposes, content, and methods of inservice training had been introduced.

In the 1940s, World War II greatly diminished the number of qualified teachers. The result was unqualified teachers with emergency teaching certificates. Whatever shift in emphasis away from remediation that had taken

place now shifted back to make the focus of inservice education one of correcting teacher deficiencies. As the war ended, standards increased and teaching began to attract more able candidates. Colleges replaced almost all normal schools and inservice education became more sophisticated. The emphasis in inservice education began to shift from "direction" to "guidance." Inservice education began to be characterized by cooperative, problem-solving workshops, and action-research (Richey, 1957).

The post-Sputnik era of educational reform from the late 1950s to the early 1970s provided millions of dollars for inservice education. National curriculum projects were instituted to educate teachers in the areas of science, mathematics, social studies, English, and foreign languages (Nicholson et al., 1976). Remediation again was the focus of these programs. Outside experts were employed to help teachers improve themselves professionally. Teachers' interests, ideas, and strengths were generally overlooked.

Meanwhile colleges introduced a whole battery of graduate courses in curriculum development, teaching methodologies, human relations, and measurement and evaluation (Edelfelt and Johnson, 1980). Partially because the college courses were often unrelated to school district needs and provided theoretical rather than practical information, school administrators developed their own inservice educa-

tion opportunities. "Inservice" as it came to be labeled often meant mandatory attendance at workshops during days or portions of days when students were not in attendance. Some districts initiated voluntary inservice programs in which credits could be earned toward salary increments.

The school districts' version of inservice education differed considerably from the colleges'.

The school district inservice approach ... put considerable emphasis on how faculty members relate to and learn from each other, and on the mutual stimulation for growth that can develop when professionals work together; the idea that teacher competence does not exist in isolation is important here; the tone and the sociopsychological climate of a school are viewed as contributing factors. (Edelfelt, 1971, p. 30)

The Teacher Center concept, which originated in England, was federally funded in this country in 1971 to stimulate teacher renewal. This took inservice education one step closer to realizing the critical role the teacher plays in his/her own professional development. Four pilot teacher centers were funded, but the funding was terminated in 1975. By 1978, through teacher organizations' lobbying efforts, the program was revived. The Teacher Centers program was unique, because it required that centers be governed by a policy board made up of a majority of teachers with representation from administrators, school boards,

and institutes of higher education. Although Federal funding was terminated in 1981, over one hundred of these teacher centers still exist through local and private funds.

Summary

This history of inservice education has included only highlights of the development of inservice education in the United States. Although much evidence of significant progress and accomplishment has been presented, the professional development of educators still is characterized by many deficiencies. Some of the most commonly identified problems include:

1. Most programs are of short duration and address single, unrelated topics (Lippitt & Fox, 1971; Sobol, 1971; Havelock, 1973; Draba, 1975; Edelfelt, 1975; and Mann, 1976).
2. There is a lack of teacher input in the inservice planning process (Kinnick, et al., 1957; Sobol, 1971; Schmeider, 1972; Edelfelt & Johnson, 1975; and Joyce, et al., 1976).
3. Participation is required whether or not you have an interest in or need for the inservice training (Graubard & Rosenberg, 1974; Edelfelt & Lawrence, 1975;

Bell & Peightel, 1976; and Mann, 1976).

4. Inservice programs fail to provide adequate incentives to teachers who participate (Gardner, 1964; Sobol, 1971; Cobb, 1973; Havelock & Havelock, 1973; Barbera, 1976; Mann, 1976; Hite & Howey, 1977; and Howey, 1978).
5. Inservice activities often fail to take into account the individual needs and general stages of professional development of the teacher (Perloff et al., 1970; Rubin, 1971; Sarason, 1974; Bell, 1975; Howey, 1978; and Lieberman, 1978).
6. There has been little concern regarding the classroom application of newly acquired skills (McLaughlin & Marsh, 1978; Joyce & Showers, 1981; Brandt, 1982; and Fullan, 1982).
7. Programs have focused on remediation of teacher weaknesses rather than recognized and capitalized on teachers' strengths (Waynant, 1971; Edelfelt & Johnson, 1975; Tikunoff & Ward, 1979; and Chall, 1980).
8. Instructors have lacked appropriate skills or have had little or no recent classroom teaching experience (Lippitt & Fox, 1971; Lawrence, 1974; Edelfelt & Lawrence, 1975; Joyce et al., 1976; and Edelfelt & Johnson, 1977).
9. Many people write about staff development yet few pro-

vide methodologically sound and rigorous evaluations of programs (Peeler & Shapiro, 1974; Nicholson et al., 1976; Griffin, 1978; Watts & Hammons, 1981; Hockman, 1982; and Loucks & Melle, 1982).

10. Inadequate resources, especially time, money, and leadership are provided for staff development (Harris & Bessent, 1969; Sobol, 1971; Schmeider & Yarger, 1974; Goodlad, 1975; and Joyce et al., 1976).
11. Most programs do not focus on the "social interaction perspective" of change but instead focus on the individual teacher (Watson, 1967; Lippitt & Fox, 1971; Havelock & Havelock, 1973; and Edelfelt & Lawrence, 1975).
12. Most programs do not address actual problems teachers encounter in the classroom (Sobol, 1971; Lortie, 1975; Elliott, 1979; Tikunoff & Ward, 1979; and Fullan, 1982).

This history gives a perspective from which to view today's issues. Considering the fact that inservice education is only about 150 years old and that less than half of the teaching population held bachelor's degrees in 1930, progress appears swift. (In 1980, 99.6 percent of all public school teachers held bachelor's degrees [In Gardner, 1982].) This history serves as a reminder of significant

progress dotted with persistent problems. Identifying the problems, however, is the first step toward alleviating them.

Theoretical Framework

In this decade, staff development is continuing to receive attention and acceptance as the means to improving education. Many people are writing about what is happening in staff development. There is no shortage of suggested guidelines and claims regarding the requirements necessary for successful staff development. Edelfelt (1981) reports that many more people are writing about staff development, yet telling the whole story is a rare occurrence. Loucks and Melle (1982) report that the evaluation of staff development programs has all but stagnated. Most evaluations are reports of participant satisfaction to determine the success of programs. These perceptions yield valuable information, but are not valid indicators of whether staff development has made a difference. Hockman (1982) agrees that staff development evaluation must be taken beyond "happiness quotients" or in other words, staff development evaluation must go beyond affective measures. In order to measure the effectiveness of staff development, it is important to ascertain whether participants have learned

and/or changed their behavior, and if so, to determine the results of these changes.

To date, the response to the need for precise evaluative data has been disappointing. The mere mention of the word evaluation brings to the minds of many educators visions of computer printouts, anxiety attacks, and thick, useless reports. Evaluation, as they see it, is simply not a priority. Not only is it not a priority, it is a nuisance in many instances. Researchers are often seen as using the school and the people in it as objects of research projects. This may involve training one group and not another, breaking up large groups into small ones, and other such necessary but artificial arrangements. These research efforts serve the researcher but not the school or the teachers. It is not surprising that through this process, an aversion to research and researchers develops (Griffin, 1982).

A second contributing factor to the lack of substantive evaluative data is the fact that staff development is relatively youthful. Its promoters have not yet had time to worry about evaluation when most of their energy is going into development. Other factors include the inadequate backgrounds in evaluation techniques possessed by most staff developers, lack of a theoretical or practical literature foundation upon which to base an evaluation, and

a lack of appropriate evaluative instruments. It is, I believe, the lack of theory or conceptualizations that is the most serious threat to effective evaluation of staff development. Staff development is such a myriad of inter-connecting parts: the participants and their personal and professional needs and characteristics, the nature of the setting, the elements of the program itself, the role and nature of the leadership, the type and role of evaluation, and the interaction of all of the above (Griffin, 1982). This list could be expanded, but it serves as a reminder of the complexity of the task and an explanation of why so few people have attempted a comprehensive study of staff development.

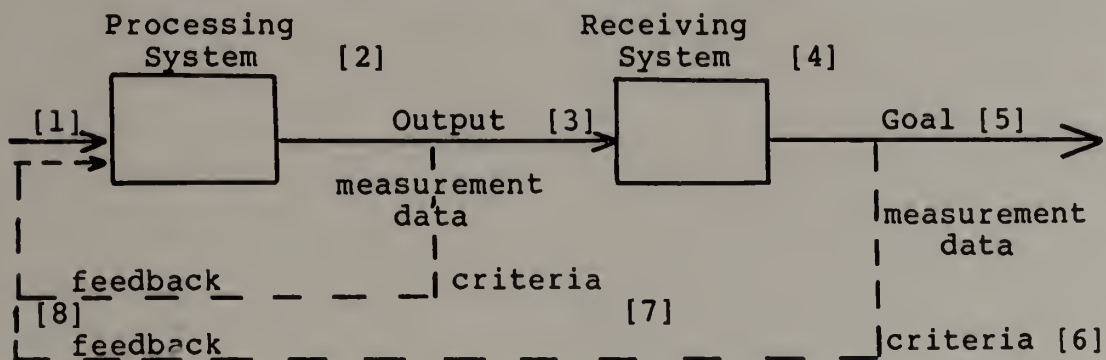
This discussion represents an attempt to provide both a conceptual base for the evaluation of staff development programs as well as a practical and operational guide for the busy practitioner. The goal is to make the evaluation of staff development as unbiased, systematic, practical, and usable as possible.

The search for a theoretical model upon which to base the evaluation of staff development programs led to a review of the training literature from business and industry which includes an abundance of relevant sources. This really came as no surprise because business and industry have had staff development programs for a long time and

profit-minded managers have insisted on evaluation of these programs. The work of Kirkpatrick (1967) and Brethower and Rummler (1977) were particularly valuable resources. Kirkpatrick was the first to conceptualize different levels of evaluation (reaction, learning, behavior, and results). Brethower and Rummler combined some of Kirkpatrick's ideas and systems theory into an evaluation system shown in Figure 2.

Figure 2

Brethower and Rummler's Evaluation System



(Source: Brethower and Rummler, 1977)

The key components of this system are the processing system, which is the staff development program and the receiving system, which consists of the specific jobs within the institution. Specifically, the components as described by Brethower and Rummler (1977) are:

1. Inputs into the system, such as instructors, administrators or secretaries.
2. The processing system, which converts inputs into outputs through such means as workshops, conferences, or seminars.
3. The outputs of the processing system, which are those same instructors, administrators, or secretaries with newly acquired skills, behavior, or knowledge.
4. The receiving system, which is the work setting into which the outputs go. (It is important to note that the processing system and the receiving system are actually sub-systems of the larger system which in most cases is the institution.)
5. The stated goal of the receiving system, such as "student drop-outs will decrease ten percent" where the receiving system is the classroom and the processing system is the instructional workshop.
6. The evaluation of the stated goal of the receiving system (e.g., do the drop-outs actually decrease by the expected ten percent?).
7. The evaluation of the outputs of the processing system. (The assessment here would focus on whether or not or to what degree the participants achieved what they were supposed to as a result of the workshop.)
8. The feedback to the processing system regarding the

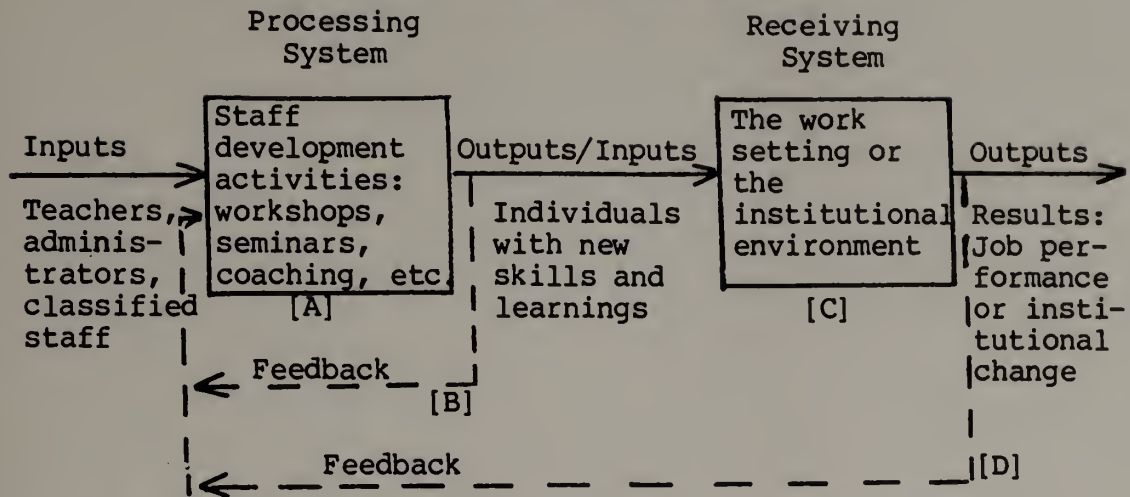
outputs of both the receiving system and the processing system.

Brethower and Rummler (1977) advocate the systems approach to staff development evaluation for several reasons. They feel that if staff development is viewed in terms of general systems theory (input, process, output), a number of alternatives or sources are identified from which evaluation data can be gathered. A second reason for a systems viewpoint is the emphasis that this approach places upon the fact that staff development can not and does not function in a vacuum. It must function as an integral part of a larger system which is the institution. Finally, the receiving system is needed as part of the model to determine the value of the staff development program as a processing system to the institution. In other words, considering only the popularity of a program or mastery of program objectives would not tell us anything about the impact of the program upon the needs of the institution.

Adapting Figure 2 into Figure 3, four sources, labeled A through D, are identical to the four levels of evaluation (reaction, learning, behavior, and results) originally described by Kirkpatrick (1967). Each level has distinct criteria for evaluating staff development and can furnish data for either formative or summative purposes.

Figure 3

Brethower and Rummler's Systems Model



(Source: Brethower and Rummler, 1977)

Levels of Evaluation

Although the systems model presented above is helpful, it really is not significantly different from systems models presented in the evaluation literature (e.g., Stufflebeam, 1969; Harris, 1980). Brethower and Rummler, however, take this systems model much further in terms of making it workable for the practitioner. Figure 3 offers several sources from which to gather evaluation data: the processing system and its outputs, and the receiving system and its outputs. As previously noted, Kirkpatrick's (1967) levels of evaluation correspond with levels A through D. The following discussion examines each of these four levels

and discusses possibilities for gathering evaluation data at each level.

Level A--Reaction

The reaction level assesses how the participants in staff development activities feel about those activities.

As shown in Chapter I, Figure 1, evaluation at the reaction level is the easiest to conduct compared to the other levels. It also yields the lowest informational value, and is the most frequently used. This level corresponds to the affective measures referred to earlier in this study (p. 30). Even though affective measures are frequently used and easiest to conduct, this does not guarantee that they are done properly. Kirkpatrick (1967) suggests the following guidelines for evaluating the reaction level:

1. Determine what facets of the activity you want to assess.
2. Develop a written assessment form to assess them.
3. Design the form so that reactions can be tabulated and quantified.
4. Maintain the anonymity of the participants for more honest reactions.
5. Allow the participants to write additional comments

not covered by the other portions of the form.

Besides questionnaires administered after a workshop, reaction level information can be obtained through unobtrusive measures such as number of registrants if registration is voluntary, attendance records, and tardiness rates.

Level B--Learning

Once data have been gathered at the reaction level, the evaluator has information regarding how well the program was received as well as information that can help to improve the program. However, an important point here is that a positive reaction to the program does not necessarily mean that the participants learned anything. A workshop may be well received because of a multitude of visual aids, numerous handouts, and a leader that commands attention, but nothing new may be learned. Kirkpatrick (1967) offers a set of guidelines for measuring learning:

1. Measure the learning of each participant so that quantitative results can be determined.
2. Utilize a pre-test and post-test approach to relate learning to the activity or program.
3. Measure the learning on an objective basis as much as possible.

4. Utilize a control group when possible for a comparison to the group that participated in the activity.
5. Analyze the results statistically so that the results have more credibility.

A number of methods could be used to accurately assess learning. Among them are: pre-test and post-tests, criterion-referenced tests, and/or norm-referenced tests. Although not a measurement of learning, the Stages of Concern Questionnaire (Hall, 1979) would also be a valuable tool to use at this stage to measure participants' developmental progress regarding their level of concern.

Pre-tests and post-tests, while easily administered, have poor external and internal validity. Criterion referenced tests can be excellent measurement tools but must be carefully checked for appropriateness, validity, reliability, and usability. Norm-referenced test, however, would not usually be available to test a specific body of knowledge such as that covered in the workshop. The Stages of Concern Questionnaire, while not particularly well validated, has proven to be reliable. It can accurately assess whether the workshop participants were helped and whether they shifted to higher level concerns than they had prior to the workshop.

Level C--Behavior

There is great difference between learning a new skill or gaining new knowledge and putting that skill or knowledge to use. Therefore, the next logical place from which to gather data is the work setting. The focal point of evaluation at the behavior level then is whether or not participants change their behavior as a result of a staff development intervention. Evaluation at this level yields more useful information, but is more difficult to assess. The guidelines that Kirkpatrick (1967) outlines for assessing behavior change are:

1. Job performance should be appraised both before and after the staff development program.
2. Job performance should be appraised by a number of people familiar with the participant's job.
3. Before and after job performance should be statistically analyzed in order to relate it to the staff development program.
4. Appraisal of job performance should take place long enough after the program for any changes to have time to be implemented.
5. A control group who does not participate in the program should be used.

The primary method for evaluating behavior change is direct observation using performance based evidence. Specific devices include recorded classroom observation (many scales and methods are presented in Beegle and Brandt, 1973), Levels of Use (LoU) interview (available through G.E. Hall at the Research and Development Center for Teacher Education, The University of Texas at Austin), and videotape or audiotape. The advantages of recorded classroom observation is the extensive sampling of behavior that can be gathered in a natural setting as well as the richness that qualitative data can offer. The disadvantage is that standardization is lacking in such an approach and the rating instrument or the rater may be inadequate.

LoU, as introduced by Hall, et al. (1975), provides a commercially available interview technique that has some advantages over direct classroom observation because: (1) interviews can get at past events, (2) interviews can reveal behavior not occurring during observation times, (3) interviews can reveal relationships that can not be observed, and (4) interviews are quick and efficient. However, in research, it is the goal to be as rigorous as possible and if one wants to measure behavior, the desired method is through direct observation.

Finally, videotape or audiotape provide a permanent record as well as information that may repeatedly be ob-

served using a variety of measures and observers. Videotape and audiotape, of course, have the same advantages mentioned above as direct observation using performance based evidence.

Another concern that should be addressed when measuring behavior change, but is often overlooked by evaluators, is long-term impact of an intervention. Initial measurement may indicate a substantial initial change in behavior and/or knowledge but it is necessary to be concerned with the impact in six months or a year following the intervention, as well.

Level D--Results

Assessing the results of a staff development program is the most difficult to accomplish yet yields the most valuable information. Unfortunately, most staff development evaluations never approach the evaluation of the results due in part to lack of prerequisite goals and objectives. As previously mentioned, staff development objectives should spell out the results that are anticipated. When the objectives are stated in specific terms, evaluation of the results becomes easier. However, determining what has happened as a result of a staff development program is still extremely difficult. Kirkpatrick (1967) cites E.C. Keachie's statement in the Journal of Industrial

Training (July-August, 1948): "Difficulties in the evaluation of training are evident at the outset in the problem technically called 'the separation of variables'; that is, how much of the improvement is due to training as compared to other factors?" Kirkpatrick concludes that studies that have attempted to penetrate such difficulties have had limited success, and our present techniques are simply not adequate.

Kirkpatrick offers no specific guidelines to follow in assessing results. He suggests that if the results or criteria have been previously stated, then evaluation should be similar to that at the behavior level. Obviously, this is the weakest area in this evaluation model, and in evaluation of staff development in general. Much work remains to be done so that interventions can be shown to have the desired results, and a cause-effect relationship between staff development and results can be proven.

Ideally, any evaluation should include data from the reaction through the results level. In practice, this rarely happens. Newstrom (1978) says that many evaluators believe there is a high sequential intercorrelation among the criteria. He suggests their reasoning goes as follows: if the reaction to a staff development activity is favorable, then participants will probably learn more; if they change their behaviors, performance will improve. A re-

verse set of conclusions would hold true for a negative reaction. Newstrom concludes that this is a dangerous and simplistic approach to evaluation. It is important to note that any given program could have favorable or unfavorable reaction, increased or no learning, desirable or undesirable behavior change, or improved or no improvement in results. The crucial point is to be cautious when assessing staff development programs not to focus on only one level, but to use a combination of all four levels.

The Evaluation Matrix

Brethower and Rummler (1977) suggest asking the following at each of the four levels of evaluation:

1. What do you want to know?
2. What can be measured to answer question 1 above?
3. What aspects of learning/performance are to be measured?
4. What instruments can be used to obtain measurement data?
5. What ways are the data to be gathered?
6. What evaluation standards are to be applied to each question?

The following matrix (Figure 4) puts into operation the model thus far described. It includes using a systems model for staff development evaluation consisting of four different levels of evaluation. By adding Brethower's and Rummel's questions to Kirkpatrick's levels of evaluation, we can determine the specifics of what and how to evaluate each particular level.

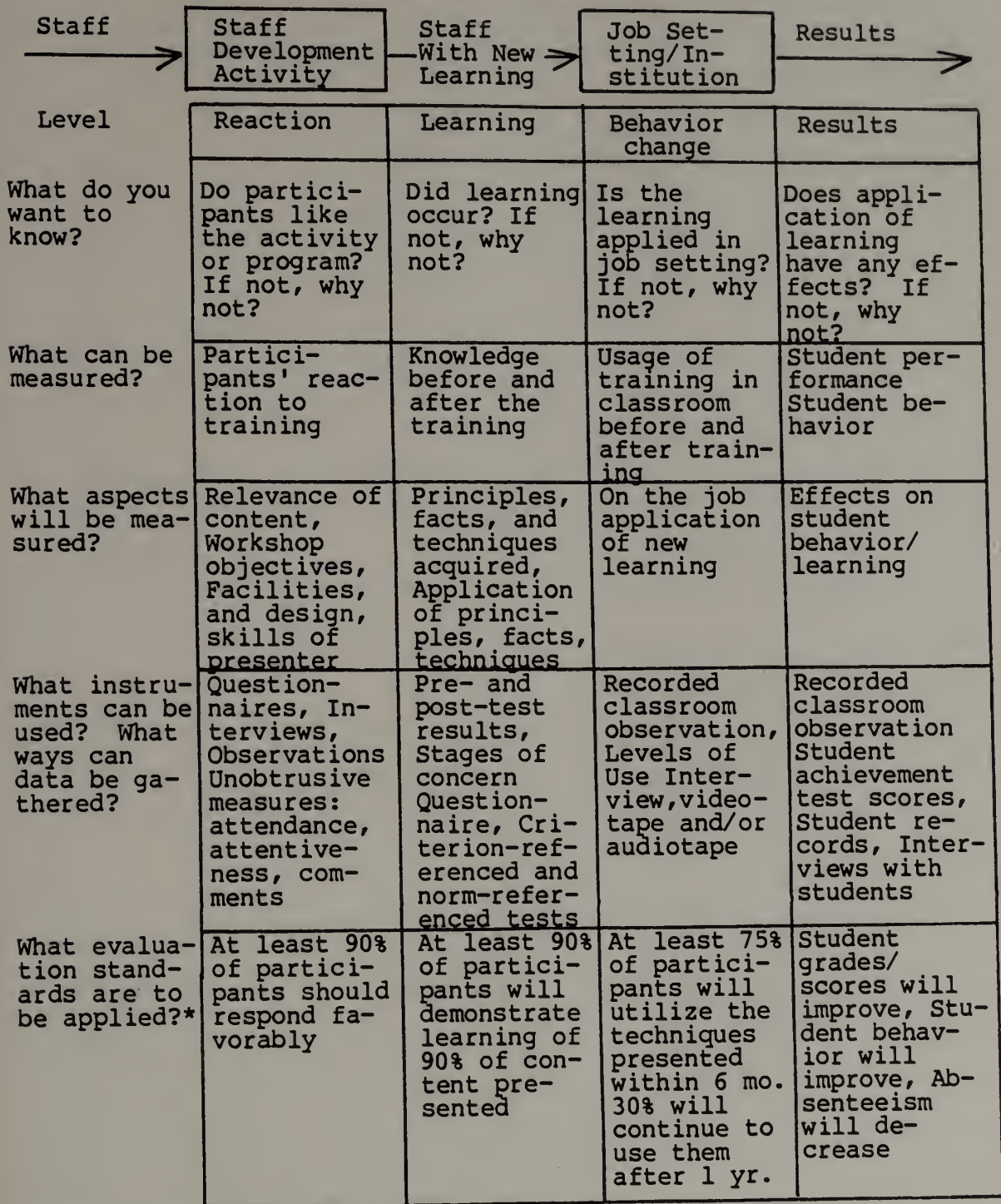
This operational model (Figure 4), provides a basis for expanding and categorizing the range of evaluation criteria applied to staff development so that the range and ambition of the program is matched by the evaluation design. In this case, the model provides the structure necessary to mount a comprehensive approach to the meta-analysis of staff development.

Research Evidence

The purpose of this section of the review is to examine selected research studies that have potential utility for improving staff development. The discussion that follows is divided into three categorical issues: (1) context, includes the physical and organizational setting; (2) design, is concerned with the ways in which inservice activities are determined, conducted, and reinforced; and

Figure 4

Operational Evaluation Model



(Source: Brethower and Rummeler, 1977)

*Levels set for success are purely arbitrary

(3) process, is concerned with the procedures involved in accepting, rejecting, or altering inservice programs.

Context

Anthropologists have known for many years the importance of and the need for an understanding of context. Spicer (1952) records a classic case of failure to recognize the importance of cultural norms. A group of Spanish American farmers in the Rio Grande Valley of New Mexico had grown poor quality corn with low yield for many years. A county extension agent introduced the farmers to hybrid corn which gave the farmers about three times the yield that was previously produced. The farmers quickly adopted the new corn but four years later, it was found that nearly all the farmers were again using the old corn. What factors had the county agent failed to consider? The answer is simple, the hybrid corn did not look like, taste like, or make tortillas like the old corn, and the farmers' wives were up in arms.

Anthropologists usually see cultures from an "etic" or an "emic" point of view. The "etic" view portrays the culture from the point of view of an outsider looking in. The "emic" point of view portrays the culture from an insiders point of view. The basic premise here is that an "emic" point of view is more likely to result in change

than an "etic" one. Heckman et al., (1983) reports that change at any given school is likely to occur in different ways. When change efforts are based on a general understanding or abstraction of schools, critical elements are missing. These elements include an understanding of the particular structures, behaviors, meanings, and belief systems that have evolved in that particular school. A likely conclusion is that an "emic" point of view is necessary in order for change to be effectively realized in schools. The "emic" point of view sees the teacher as a crucial element in the change process.

Miles (1980) reports on a four-year study by the Center for Policy Research of the planning and implementation of six innovative public schools. Despite the good intentions and strong commitments of the change agents and staff members of the schools, the efforts to change structures of the schools and behaviors of staff members were only moderately successful. Explanations of these failures can only be explained in terms of the "social architecture" a term used by Miles (1980) to talk about the degree of influence exerted by the school, its history, the characteristics of the immediate community, and the adults and their relationship within the school.

Berman and McLaughlin (1975), in their study of federal programs supporting educational change, found that pro-

jects having active support from principals were more likely to fare well. They referred to the principals as the "gatekeepers of change." In all five studies on educational knowledge diffusion, synthesized by Emrick and Peterson (1978), administrators played a crucial role in change efforts.

Wolf (1981) and Griffin and Lieberman (1974) talk about certain context-related variables that serve as predictors of change. Wolf (1981) includes four characteristics and commitments that support successful change efforts. These include: (1) prior history of successfully linking knowledge production with utilization; (2) compatibility of the leader (regarding training, experience, etc.) with the targeted audience; (3) adequate time to facilitate the change effort; and (4) technical knowledge necessary for the task. Griffin and Lieberman (1974) cite the ability to analyze and understand institutional variables as important predictors of change. This includes knowledge of the system, the subsystems, and gathering and acting upon information about the history of the organization. They conclude that a thorough understanding of organizational development and the ability to coordinate the organizational variables is essential for a successful staff development effort.

Little (1981) used a focused ethnographic methodology

to study effectiveness in six urban schools. She quickly realized the importance of contextual variables such as the principal's interaction with teachers, the character and disposition of the school (particularly adult-adult interactions and beliefs about teaching), and the overall power of the school setting characteristics to influence staff development efforts and effectiveness.

Another important contextual variable includes environmental conditions for change. Wolf (1981) identifies such variables as dissatisfaction with the current practice earmarked for change, a needs assessment that indicates a change is necessary, and the necessary resources to implement the change as important determinants of change.

In summary, context issues which have been reported to positively affect staff development efforts include:

1. An emic point of view regarding the school and its staff.
2. Supportive administrators, particularly the principal.
3. Prior history of successfully implementing innovations.
4. Project leadership that is compatible with the targeted audience.
5. Sufficient time to facilitate the change effort.
6. A project leader who is able to analyze the character-

istics of the setting and facilitate an appropriate plan.

7. An environment that includes positive adult interactions and progressive beliefs about teaching.
8. Dissatisfaction with the current practice earmarked for change.
9. A needs assessment indicating a change is desired.
10. Adequate resources needed to implement the change.
11. Positive perceptions of school personnel regarding the practice.

Design

Design issues concern and include the duration of the activity, the location and scheduling, the focus and nature of the training, and the methods of instruction employed.

Duration. As noted earlier in this review, many authors have recognized the fact that most inservice programs are of short durations and address single, unrelated topics. Although Joyce et al. (1976) reveal that forty-three percent of a nationwide sample of educators feel that professional development should be provided on a continuing basis, in fact, this rarely happens. Burk (1976) points out that administrators tend to use inservice education as a means of solving crises.

Although much is written outlining the need for long-

term approaches to training, the existing data on this question are not convincing. In fact, Lawrence and Harrison (1980) in a meta-analysis of the research on staff development actually found a higher effect level from short-term projects (lasting less than six months) than they did from long-term projects. These data, however, are questionable because the findings were based on only eleven long-term projects and on seventy-one short term projects. Duration will be examined further in the meta-analysis that follows in Chapters IV and V.

Location and Scheduling. The location of inservice programs has received considerable attention within the staff development literature. A few staff developers have argued for conducting programs away from schools in less rigid and more comfortable surroundings. Most evidence supports the opposing view, however, that the school site should serve as the primary location of inservice programs.

Havelock and Havelock (1973) have argued that the more consistent the training setting is to the implementation setting, the greater the chances for successful implementation of training. Nicholson et al., (1976) report that teachers prefer inservice training presented on-site because it is convenient and can be more closely related to the teaching role. Joslin (1980) and Lawrence and Harrison (1980) both conducted a meta-analysis including the ques-

tion of on-site versus off-site programs. Both researchers, who together analyzed 868 cases, found that on-site programs resulted in higher effect sizes. Joslin (1980) reported a mean effect size of .517 for on-site programs and a mean effect size of .352 for off-site training. Lawrence and Harrison reported a mean effect size of .828 for on-site activities and a .635 mean effect size for those activities held off-site. In short, the evidence in favor of on-site programming is quite convincing.

Another design issue, quite closely related to location of inservice programs is scheduling. Arguments can be readily found on both sides of the released time issue. Arguments supporting released time include the following: (1) teachers will be too tired to concentrate on inservice at the end of the day (Mohr, 1971); (2) teachers prefer inservice education during the school day (Joyce et al., 1976); and (3) released time is needed to plan and implement inservice programs (Sobol, 1971).

Arguments opposing released time include: (1) teachers do not like interruptions in the school day (Joyce et al., 1976); and (2) administrators will not condone the interruptions and expense created by having teachers out of their classrooms (Haines, 1973). An often spoken, but not widely quoted, objection is that teachers belong in the classroom.

Joslin (1980) and Lawrence and Harrison (1980) both looked at location in their meta-analyses. Joslin (1980) found mean effect sizes of .475 for released time activities, .538 for after hours events, and .305 for weekend/summer activities. Lawrence and Harrison (1980) report mean effect sizes of .660 for released time activities, .740 for before/after work, .810 for summer, and .180 for Saturdays. In both studies, programs held in the evenings or before/after work were more effective than those held during the work day.

Focus and Nature. The staff development literature is not dominated by findings that are directly or indirectly applicable to the formulation of sound training practices for educational personnel. The majority of the literature that does exist is highly propositional in character.

Again Joslin (1980) and Lawrence and Harrison (1980) provide the only comprehensive studies designed to identify effective practices in inservice training. Both studies found that programs that attempted to increase knowledge were more effective than those that attempted to change the behavior of participants. Lawrence and Harrison found that programs with fewer than sixty participants were more effective than larger programs. Joslin reported that programs directed at elementary school teachers were slightly more effective than those focusing on secondary school

teachers. (Berman and McLaughlin, 1975, also suggested targeting of staff development to elementary schools.) Both Joslin (1980) and Lawrence and Harrison (1980) found that college courses and institutes resulted in higher effect size means than did workshops.

Methods of Instruction. The literature on teacher training seems to indicate that mastery of teaching skills and transfer of training can be accomplished through the use of a combination of five training strategies: (1) study of theory underlying the skill, (2) opportunity to observe demonstration of the skill, (3) practice in simulated and classroom settings, (4) structured and open-ended feedback, and (5) coaching for application (Joyce & Showers, 1980).

Throughout the staff development literature, the work of Joyce and Showers is touted for its strong implications for staff developers. The message is clear: Staff development should focus squarely on skill development. More importantly, staff development should lead to classroom application. The chances for newly learned skills being applied in the classroom are greatest when there is coaching for application. Brandt (1982) sums it all up when he concludes that if staff development had less presenting and more practice, feedback, and coaching, teachers would have more professional skills and probably more self confidence.

These conclusions, however, are highly speculative.

Berman and McLaughlin (1975) report several findings related to methods of instruction. They report the following characteristics were typical of successful staff development projects: staff training focused on practical aspects of project operations, high levels of support activities (e.g., in-class assistance, visits to demonstration classrooms, observation and feedback from project leaders), and materials developed by local participants.

Although there is much speculation on the types of instruction that are most effective, no hard evidence could be found to support most of the claims. This study will attempt to provide evidence to help answer the long-neglected question: What types of instruction, individually or in combination, produce positive participant reactions, increased learning, behavior change, and results in the classroom?

Process

During the last decade, researchers have begun to address the "why" of teacher behavior change. If our goal is to improve teaching, we need to discover why teachers are affected by certain staff development efforts and not by others.

Berman and McLaughlin (1975) suggested the importance

of participative governance, inclusion of highly motivated staff who volunteered to participate, involvement of a "critical mass" of participants, a problem-solving approach, and teacher-administrator interaction in the staff development process.

Teacher-administrative teams is a theme encompassed by many theories of staff development (Griffin, 1982; see also Bentzen, 1974, and Little, 1981). Findings by Tikunoff, Ward, and Griffin (1979) in their study on Interactive Research and Development on Teaching (IR&DT) confirm the important role of teacher involvement on teams working to effect change. In IR&DT, researchers, teachers, and teacher educators work together in school-based teams to engage in systematic research and development activities. Some of the benefits of IR&DT reported by Tikunoff and Griffin (1980) include:

1. Participants increase their awareness of educational options as a consequence of the requirements of systematic inquiry into schooling.
2. Participants increase their knowledge and skill regarding educational research and development.
3. Participants are more knowledgeable about, skillful in, and sensitive to research and development issues as a result of their participation in the implementa-

tion of IR&DT.

4. Teacher isolation is ameliorated for the participants.
5. Teachers spend significant amounts of time discussing research and development issues.
6. Institutional isolation, school from school and school from university, is lessened.
7. There is an increased belief on the part of participants that the work of other team members is of value and otherwise "prestigious" in the workplace and the broader educational community.
8. Professional practices are altered as a result of research findings.
9. Teachers begin to rely less upon "instinct" and more upon research as their work proceeds.

Huling (1981) reported additional positive outcomes from IR&DT studies that include:

1. Teacher participants get increased recognition from peers.
2. Teachers are not only more likely to use findings from their own teacher-conducted research, but also are more likely to consult and use other research as well.

In short, major benefits have been realized by this

approach in which teachers and school personnel not only pose the questions, but test their own solutions as well. An IR&DT approach to change, results in an enabling condition rather than one that devalues the role of the teacher. The rise of teacher centers and advisory programs can be seen as an effort to create conditions both inside and outside schools to support continuing growth and self-renewal in teachers (Devaney, 1977). The Ford Teaching project produced a set of hypotheses sampled below which elaborate this point of view:

-The more a teacher comes to value him/herself as a potential researcher, the more open he/she will become to observer feedback.

-The more able a teacher is at self-monitoring his/her classroom practice, the more likely he/she is to bring about fundamental changes in it.

-The less financial and status rewards in the school are primarily related to administrative and pastoral roles, the more able will teachers become at self-monitoring their classroom practice. (Elliott, 1976, pp. 44-50)

Implicit throughout the IR&DT approach is the importance and value of the role of the teacher as decision-maker. Increased capacity to deal with research and development issues is a potentially powerful antidote to growing sentiments that teaching and teachers are of low status and low priority in terms of social action and reward systems

at local and national levels of authority.

A related but slightly differing finding from the work reported by Bentzen (1974) is "the peer group strategy." This phenomenon encourages teachers and administrators to not just work on problems together but to accept the fact that most, if not all, of the solutions to the problems they may face reside within themselves and their environment. With this strategy, there is no need to rely on outside technical assistance as there is in the IR&DT approach.

Berman and McLaughlin (1975) stressed the importance of "mutual adaptation" if the goal of training is to effectively implement a new idea or project and have it persist. Mutual adaptation refers to the process by which both the project design and the institutional setting change as an innovative project is introduced and implemented. Griffin (1982) reports that prior to the acceptance of this notion, a change was considered a failure unless it was implemented exactly as the developers had envisioned it. Furthermore, Griffin (1982) suggests that the staff developer who accepts mutual adaptation as a desirable outcome is more likely to plan for it and to judge his/her efforts in this light.

Adult development theory has also provided "process" guidelines for staff development. Two significant implica-

tions have been pointed out by Wood and Thompson (1980). First, it has been discovered that a higher proportion of adults are operating at a concrete level than formerly believed. This supports Berman and McLaughlin's (1975) contention that staff training should focus on the concrete with teacher-specific plans. Abstract, talk-oriented sessions are not conducive to changing behavior. Second, Ward and Thompson (1980) (see also Rapport & Rapport, 1975; and Tough, 1967) suggest that adults prefer learning in informal situations where social interactions are possible. This finding implies that inservice education may best be conducted in settings that are conducive to social interaction.

Wood and Thompson (1980) summarize their findings regarding effective staff development for adult learners into six statements:

1. Control over the "what" and "how" of learning is necessary.
2. Staff development should focus on real and important job tasks.
3. Choices and alternatives must be given in order to accommodate individual differences.
4. Opportunities to practice and apply new learnings are valuable aspects of training.

5. Small group activities provide occasions for learning from one another.
6. Opportunities for peer-participants to give each other feedback concerning performance are desirable.

From these studies, then, we can conclude that the following processes have been associated with effective staff development:

- voluntary participation
- high level support activities
- participative governance
- a problem-solving approach
- inclusion of highly motivated staff,
- involvement of a "critical mass" of teachers
- training characterized by mutual adaptation
- teacher-administrator teaming or other similar professional relationships
- opportunities for teachers to act on problems they perceive as important
- activities that are guided by adult development theory

Summary

Many research studies have been reviewed in order to examine their potential for improving staff development. Although numerous claims have been made regarding the context, design, and processes that are most conducive for effective staff development programs, little hard evidence has been presented. Much work remains to be done in order to carefully determine the most effective practices for the planning and conduct of staff development.

C H A P T E R I I I

METHODOLOGY

This chapter outlines the methods employed in the meta-analysis of the research evidence on the effects of inservice teacher education. The steps in the process included (1) identifying and collecting the studies, (2) identifying, operationally defining, and coding the variables, (3) calculating effect size, (4) computer processing and (5) analyzing the data.

Identifying and Collecting the Studies

An attempt was made to gather all relevant studies on inservice education from 1968 to 1983, or over the past fifteen years. Studies were located through the ERIC bank, Dissertation Abstracts, and references used by Joslin and Lawrence and Harrison in their 1980 meta-analyses of inservice education. Using the search words inservice teacher education, staff development, teacher improvement, instructional improvement, and elementary education or secondary education, a computer search of the ERIC files and Dissertation Abstracts was completed.

The review process began with an examination of over

300 abstracts. Approximately 190 studies were obtained and reviewed; of these, ninety-one were determined to meet the topical criteria as well as provide the necessary data from which to calculate effect size. The following criteria were used in selecting the studies included in this meta-analysis:

1. The study was quantitative rather than qualitative.
2. The data necessary for calculating effect size were presented.
3. The study examined individually or in combination the types of policies, formats, materials, leadership roles and styles, and procedures that are most effective in promoting the professional development of educators.
4. Subjects of the study were public school teachers or their students in grades K-12. While the focus was training programs for teachers, studies could include counselors, supervisors, and administrators.

Only those studies which could be read in their entirety were included due to the necessity of coding a large number of variables that simply are not presented in summaries of research. All ninety-one studies are dissertations, ERIC documents, or journal articles. Many disserta-

tions were eliminated from consideration because they were not available for loan through interlibrary loan arrangements. A complete list of research reports examined in this study is included in Appendix A. Appendix B identifies studies by an ID number.

Identifying, Operationally Defining and Coding the Variables

A review was made of several research reviews and meta-analyses of inservice education in order to determine the variables appropriate to the topic. This list of variables was then revised after reading ten studies selected for this review. The result was a list of twenty-eight independent variables and two dependent variables.

In order to assure accuracy in coding the variables, each category had to be operationally defined. Even though this precaution was taken, coding the variables was occasionally a complex task, subject to interpretation. A complete list of variables and their operational definition is included in Appendix C.

Not all studies reported information on every one of the twenty-eight selected independent variables. In the studies where data were not available, those cases were coded as missing. A complete listing of independent and

dependent variables and their values is included in Appendix D. Appendix E contains a complete list of the coded data.

Meta-Analysis

Meta-analysis uses statistical analysis for the purpose of integrating the findings of many empirical research studies on a given topic. The use of meta-analysis requires the calculation of effect size, and then enables the researcher to determine relationships between dependent and independent variables.

Although Glass is credited with developing the meta-analysis technique, Cohen (1969) and Rosenthal and Rosnow (1975) discussed the potential of using effect size as a means to make comparisons among studies. Cohen suggested criteria which can be used in the interpretation of the meaning of effect size (Table 1).

Glass and Smith (1980, abstract) describe their approach to meta-analysis as a process involving five steps: "1) defining the problem, 2) finding the research studies, 3) coding the study characteristics, 4) measuring the study findings on a common scale, and 5) analyzing the aggregation of findings and their relationship to the characteristics."

Table 1
Relative Effect Size

Classification	Value	Range
Large Negative ES	-0.8	<-0.65
Medium Negative ES	-0.5	-0.65 to -0.35
Small Negative ES	-0.2	-0.35 to 0.0
No Effect	0.0	0.0
Small Positive ES	0.2	0.0 to 0.35
Medium Positive ES	0.5	0.35 to 0.65
Large Positive ES	0.8	>0.65

Source: Cohen (1969)

Calculating Effect Size

In general terms, effect size is a standardized measure of the effectiveness of the treatment (Joslin, 1980). More specifically, it is the mean difference between treated and control groups divided by the standard deviation of the control group (Glass, 1977).

In symbolic terms,

$$ES = \frac{\bar{X}_E - \bar{X}_C}{S_x}$$

where

ES = Effect Size

\bar{X}_E = Mean of the Experimental Group

\bar{X}_C = Mean of the Control Group

S_x = Standard Deviation of the Control
Group

As shown in Table 1, a positive effect size favors the treated or experimental group while a negative one favors the control group. An effect size of .5 means that the treatment group showed one-half of a standard deviation greater change than the control group. By using Table 1, effect size can be relatively interpreted. Another way to interpret the effect size is to compare it to other effect sizes, particularly at the within study level. For example, Joslin (1980) reported a grand mean effect size of .472 for 824 data sets analyzing the effectiveness of inservice teacher education. In contrast, Lawrence and Harrison (1980) reported a grand mean effect size of .760 for eighty-two data sets measuring the effectiveness of inservice teacher training. Upon closer examination, one realizes that Joslin was conservative in determining effect sizes while Lawrence and Harrison treated studies without control groups the same as studies with control groups. In short, comparing effect sizes between the two studies is not particularly enlightening. However, when looking at

effect sizes within each study, the differences in the different treatments becomes apparent.

When research studies do not report the mean and standard deviation and/or do not employ control groups in their design, a close estimate can be obtained by making use of other reported statistics. Glass (1977) discussed these procedures in a chapter entitled "Integrating Findings: The Meta-Analysis of Research."

In this study, the following formulae as outlined by Glass (1977) and employed by Joslin (1980) were used to determine effect size:

1. Where means, standard deviation and control group figures were reported:

$$ES = \frac{\bar{X}_E - \bar{X}_C}{S_x}$$

where

\bar{X}_E = Mean of treatment group

\bar{X}_C = Mean of control group

S_x = Standard deviation of control group

2. Where only a t statistic was given to compare the experimental and control group:

$$ES = t \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

where

t = statistics as reported

n_1 = number of subjects in treatment group

n_2 = number of subjects in control group

3. Where only F values were reported for the experimental and control group:

$$ES = \sqrt{F} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

4. If only the α level of significance was reported for the experimental and control group, the most conservative t value was assumed:

$$\alpha = .05, t = 1.96$$

$$\alpha = .01, t = 2.58$$

For example if the reported α is .05 then,

$$ES = 1.96 \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

5. In cases where no control group was employed, and a correlated t, with significance given was reported:

$$\alpha = .05, t = 1.96$$

$$\alpha = .01, t = 2.58$$

$$\alpha = .005, t = 2.81$$

$$\alpha = .001, t = 3.30$$

For example, if t were significant at the .01 level of significance then,

$$ES = 2.58 \sqrt{\frac{1}{n}}$$

where

n = number of subjects in treatment group

Only the first formula can be considered an accurate measure of effect size. As noted by Joslin (1980), formulae two through five provide a conservative estimate of effect. Formulae four and five are especially conservative, because if the calculated t were available, it would be at some point beyond the alpha-level indicated.

In this study, a total of 715 effect sizes were calculated from ninety-one research reports. Many researchers examined a variety of treatments and used numerous measures in reporting their findings. Therefore, many effect sizes could result from only one research study. Care was taken to avoid overlapping effect sizes, because in many cases the same finding was reported more than once.

Computer Processing

The data from each data set were keypunched on individual IBM cards and entered for batch processing. The analysis of the data was done by computer using Statistical Package for the Social Sciences (SPSS) outlined by Nie, et al. (1975).

Data were coded using the range of values presented in

Appendix D. If no value was reported, then it was assigned a value of nine. If a variable was not applicable or not present, then a zero was assigned to it. The SPSS system permits the assignment of missing values so that cases with incomplete data can still be processed. In this study zero and nine were processed as "missing values."

Analyzing the Data

Three Statistical Package for the Social Sciences (SPSS) programs were used in the analysis of the data. Several results were expected from the analysis of the aggregated findings:

1. A grand mean effect for all data sets as well as for data sets grouped by effect level (reaction, learning, behavior, results).
2. The relationship among the various independent variables and the two dependent variables (effect size and effect level).
3. The identification of those independent variables which have the highest correlation with the two dependent variables.

The programs employed in this meta-analysis included:

1. CONDESCRIPTIVE--A program for obtaining descriptive statistics including: mean, standard error, standard deviation, variance, kurtosis, skewness, range, minimum, and maximum for a set of variables with continuous data. This program was used in order to obtain mean effect sizes and a complete description of the data base.
2. FREQUENCIES--A program for obtaining distributional characteristics of discrete variables. This program was used to obtain frequencies of selected variables.
3. ANOVA--This program performs analysis of variance for factorial designs. It was employed in order to examine the analysis of variance among the two dependent variables and the independent variables, and when possible, the interaction between the two.

C H A P T E R I V
PRESENTATION OF THE DATA

The outcomes of the meta-analysis of ninety-one studies on various aspects of inservice education are presented in this chapter. The data are reported in tables with brief narrative descriptions accompanying each one.

The data presentation is organized into five categories: (1) description of the data base, (2) grand mean effect size, (3) grand mean effect size by dependent variables, (4) analysis by independent variables, and (5) summary of the salient findings. An analysis of the results follows in Chapter V.

Description of the Data Base

A total of ninety-one studies were examined in this meta-analysis. These studies yielded a total of 715 data sets. All tables presented in this chapter are based on 715 cases for each variable. When the total N is less than 715 for a particular variable, there are missing data within that variable.

A list of studies by source, number of reports, and data sets is provided in Table 2. About half of the data

Table 2
Source of Research Studies

Source	Number of Reports	Data Sets Yielded
Dissertations	22	154
ERIC Documents	43	393
Journals	<u>26</u>	<u>168</u>
Totals	91	715

were retrieved from ERIC documents, while the other half were derived from dissertations and journal articles.

Grand Mean Effect Size Results

The grand mean effect size is .5215, a medium-sized positive effect size according to Cohen's system of classification (Table 1). This effect size can be interpreted to mean that inservice treatment groups showed .52 of a standard deviation greater change than control groups. Descriptive statistics for the grand mean effect size are presented in Table 3. While this mean effect size tells us that inservice programs are generally effective, the focus of this chapter will be on the relative effectiveness of different inservice treatments.

Table 3
Grand Mean Effect Size

Mean Effect Size	Variance	Standard Deviation	Standard Error	Minimum	Maximum
.5215*	.4633	.6806	.025	-6.333	3.318

*p<.05; N=715

Grand Effect Size by Dependent Variables

Effect level is defined as the level at which a program is evaluated. Evaluation can occur at four different effect levels: reaction, learning, behavior, and results. These four effect levels serve as the dependent variables in this study.

Substantial differences occur when studies are grouped by effect level. As indicated in Table 4, inservice programs are most effective when the training program measures outcome in terms of learning. A considerably smaller effect is found in programs that attempt to measure results by looking at the students of participants. Changes in participant learning and behavior are more easily achieved than changes in participant attitudes or student outcomes.

Table 4

Grand Mean Effect Size by Dependent Variable Measures

Effect Level	Data Sets N	Mean Effect Size	Variance	Standard Deviation
Reaction	233	.4217*	.1989	.4460
Learning	52	.9027*	.9696	.9847
Behavior	298	.6009*	.5194	.7207
Results	132	.3682*	.5042	.7101

* $p < .05$; N=715

Analysis by Independent Variables

The following discussion and tables are based on an analysis of independent variables (policies, formats, materials, leadership roles and styles, and procedures) by the effect size of the dependent variables (reaction, learning, behavior, and results). Effect sizes are reported for all independent variables in terms of the dependent variables. Total means for dependent and independent variables are also reported. An analysis of variance is used to test the significance of the effect level, the various features of the inservice program, and when possible, the two-way interactions between the effect level and the var-

ious features of the inservice program.

Instructional Focus--Tables 5 and 6

Although most instruction described in Table 5 focused on improving a specific subject or improving general teaching, those training sessions which focused on affective techniques yielded higher effect sizes.

Source--Tables 7 and 8

As reported in Table 7, studies found in journals have larger mean effect sizes than those found in ERIC documents and dissertations. The majority of studies which measure the effectiveness of inservice training are reported in ERIC documents. The ERIC documents in this study consisted mainly of papers presented at national conferences, final reports for federal grants, and reports from local education agencies.

Research results retrieved from journals resulted in significantly higher effect sizes than those found in ERIC documents or dissertations (Table 7).

Number of Participants--Tables 9 and 10

It is interesting to observe that the highest effect size reported in Table 9 is $ES=1.1999$. This effect size

Table 5

Mean Effect Size by Instructional Focus by Effect Level

Instructional Focus	Effect Levels				Total Means
	Reaction	Learning	Behavior	Results	
Improvement of Specific Subj.	.3734	.6658	.4868	.4044	.4285
(N)	(91)	(27)	(17)	(99)	(234)
Improving Gen. Teaching	.4165	.9663	.6385	.1921	.5608
(N)	(123)	(20)	(257)	(28)	(428)
Affective Techniques	.6869	1.9247	.2788	.6378	.6145
(N)	(19)	(5)	(24)	(5)	(53)
Total Means	.4217	.9027	.6009	.3682	.5215
(N)	(233)	(52)	(298)	(132)	(715)
N=715					

Table 6
 Analysis of Variance Results for Instructional Focus and
 for Effect Level

Source of Variation	df	Sum of Squares	F
Instructional			
Focus	3	14.402	11.135*
Effect Level	2	4.194	4.863*
Two-Way			
Interaction	6	11.933	4.613*

*p<.05

Table 7
Mean Effect Size by Source by Effect Level

Source	Effect Levels				Total
	Reaction	Learning	Behavior	Results	Means
Journals	.6992	2.0631	1.0021	.4664	.9166
(N)	(17)	(10)	(104)	(37)	(168)
ERIC	.3914	.4730	.3683	.4257	.3949
(N)	(133)	(24)	(151)	(85)	(393)
Dissertations	.4134	.8308	.4474	-.4837	.4134
(N)	(83)	(18)	(43)	(10)	(154)
Total Means	.4217	.9027	.6009	.3682	.5215
(N)	(233)	(52)	(298)	(132)	(715)

N=715

Table 8
Analysis of Variance Results for Source
and for Effect Level

Source of Variation	df	Sum of Squares	F
Source	2	15.140	40.519*
Effect Level	3	8.983	24.041*
Two-Way Interactions	6	3.495	9.354*

* $p < .05$

Table 9

Mean Effect Size by Number of Participants by Effect Level

Number of Participants	Effect Levels				Total Means
	Reaction	Learning	Behavior	Results	
1-20	.4668	.5725	.2737	----	.4025
(N)	(19)	(11)	(24)		(54)
21-40	.3683	.9354	.6924	.1547	.5588
(N)	(111)	(15)	(125)	(3)	(254)
41-60	.6802	.2828	.2838	.1264	.4916
(N)	(40)	(5)	(19)	(7)	(71)
>60	.3380	1.1999	.6197	.3873	.5188
(N)	(63)	(21)	(130)	(122)	(366)
Total Means	.4217	.9027	.6009	.3682	.5215
(N)	(233)	(52)	(298)	(132)	(715)

N=715

Table 10
One-Way Analysis of Variance Results
for Number of Participants

Effect		Sum of	
Level	df	Squares	F
Reaction	3	3.4706	6.207*
Learning	3	4.9915	1.796
Behavior	3	5.5724	3.573*
Results	2	.5906	.582
Combined Levels	2	1.6860	2.884

* $p < .05$

resulted when there was a large group (>60) and the effect level focus was learning.

As shown in Table 10, the number of participants in an inservice program generally appears to be unrelated to effect size. When the data are examined in terms of effect levels, however, effect size is significantly affected by the number of participants at the reaction and behavior effect levels.

Location--Tables 11 and 12

Whether programs are conducted at the participants' place of employment or elsewhere does not make a significant difference (Table 12). When the results are broken down by dependent variables, some differences appear (Table 11), but they are not significant.

Initiator--Tables 13 and 14

Only twelve cases are reported in which the initiator is a participant in the program. Most programs were initiated by university researchers. When programs are grouped by outside originators (state/federal government, university researchers, and consultant) versus in-school originators (participant, administrator/supervisor, and school), the outside originated programs number 460 versus 174 for programs initiated by the school. When effect sizes are compared, the outside originated programs result

Table 11
Mean Effect Size by Location by Effect Level

Location	Effect Levels				Total
	Reaction	Learning	Behavior	Results	Means
On-Site	.4414	1.0591	.6178	.3861	.5404
(N)	(125)	(27)	(214)	(118)	(484)
Off-Site	.3969	.7131	.5980	.2174	.4758
(N)	(99)	(24)	(47)	(14)	(184)
Total Means	.4217	.8963	.6142	.3682	.5226
(N)	(224)	(51)	(261)	(132)	(668)

N=668

Table 12
Analysis of Variance Results for Location
and for Effect Level

Source of Variation	df	Sum of Squares	F
Location	1	1.655	3.614
Effect Level	3	12.632	9.195*
Two-Way Interaction	3	1.206	.878

*p<.05

Table 13
Mean Effect Size by Initiator by Effect Level

Initiator	Effect Levels			Results	Total Means
	Reaction	Learning	Behavior		
Participant	.1470	.0000	.0788	----	.0941
(N)	(5)	(2)	(5)		(12)
Admin/Super	.3594	.5997	.2523	.4641	.4179
(N)	(18)	(3)	(12)	(54)	(87)
School	.2688	.6170	.2191	.0459	.2285
(N)	(33)	(2)	(30)	(10)	(75)
State/Fed	.6607	.9869	.5747	.0000	.6972
Government					
(N)	(39)	(11)	(3)	(2)	(55)
University	.3647	.9642	.7410	.2636	.6010
Researcher					
(N)	(106)	(33)	(180)	(36)	(355)
Consultant	1.0440	----	.5859	.2044	.5234
(N)	(4)		(33)	(13)	(50)
Other	.7326	----	.2728	.9829	.5401
(N)	(14)		(25)	(9)	(48)
Total Means	.4335	.8963	.5946	.3751	.5255
(N)	(219)	(51)	(288)	(124)	(682)

N=682

Table 14
One-Way Analysis of Variance Results for Initiator

Effect		Sum of	
Level	df	Squares	F
Reaction	6	6.6636	6.236*
Learning	4	2.2693	.554
Behavior	6	13.4153	4.585*
Results	5	5.9448	2.361*
Combined Levels	6	10.9430	4.104*

*p<.05

in ES=.6041 versus ES=.3140 for in-school originated programs.

Grade Level Taught--Tables 15 and 16

As reported in Table 15, elementary participants achieved slightly higher effect sizes than secondary participants. A surprising finding is that when the participants are made up of both elementary and secondary participants, the effect size increases for both groups to .6788.

Participation--Tables 17 and 18

Although no significant differences appear for parti-

Table 15

Mean Effect Size by Grade Level Taught by Effect Level

Grade Level Taught	Effect Levels				Total Means
	Reaction	Learning	Behavior	Results	
Elementary	.4356	.7163	.5300	.3981	.4722
(N)	(125)	(17)	(145)	(107)	(394)
Secondary	.2433	.8777	.3755	.2235	.3673
(N)	(39)	(12)	(36)	(11)	(98)
Both	.4974	1.1623	.7582	.2534	.6788
(N)	(69)	(19)	(117)	(14)	(219)
Total Means	.4217	.9332	.6009	.3682	.5214
(N)	(233)	(48)	(298)	(132)	(711)

N=711

Table 16
Analysis of Variance Results for Grade Level
Taught by Effect Level

Source of Variation	df	Sum of Squares	F
Grade Level Taught	2	2.691	3.075*
Effect Level	3	13.174	10.036*
Two-Way Interactions	6	2.873	1.094

*p<.05

Table 17
Mean Effect Size by Participation by Effect Level

Participation	Effect Levels				Total
	Reaction	Learning	Behavior	Results	Means
Mandatory	.2810	.5460	.2085	.2767	.2467
(N)	(31)	(2)	(49)	(7)	(89)
Voluntary	.4362	.7407	.5577	.1946	.4791
(N)	(189)	(43)	(217)	(71)	(520)
Total Means	.4143	.7321	.4934	.2020	.4452
(N)	(220)	(45)	(266)	(78)	(609)

N=609

Table 18
 Analysis of Variance Results for
 Participation by Effect Level

Source of Variation	df	Sum of Squares	F
Participation	1	.513	1.507
Effect Level	3	1.072	1.050
Two-Way Interactions	3	1.333	1.307

participation, effect levels, or their interactions, there is a higher effect size reported in Table 17 for voluntary versus mandatory attendance. The lack of significant differences for mandatory versus voluntary participation is surprising, because participant motivation is usually reported to have a significant impact upon program effects.

Participant Incentives--Tables 19 and 20

As shown in Table 19, college credit is the most frequently reported incentive for participation and results in a respectable effect size of .5760. The largest effect size reported for the incentive variable is status. Apparently those participants whose reward for participation is added status, either through special recognition, a competitive selection process, or a similar special designation,

Table 19

Mean Effect Size by Participant Incentive by Effect Level

Incentive	Effect Level			Results	Total Means
	Reaction	Learning	Behavior		
Pay	.3138	----	1.1820	.1475	.3104
(N)	(31)		(1)	(6)	(38)
Certificate	.0677	.4100	.3226	.1147	.2660
Renewal					
(N)	(4)	(5)	(12)	(4)	(25)
Released Time	.5267	----	.2918	.2907	.4033
(N)	(19)		(18)	(3)	(40)
Status	.6461	.4035	.9822	.5095	.7665
(N)	(32)	(2)	(26)	(4)	(64)
College	.5865	.8701	.7232	-.3712	.5760
Credit					
(N)	(25)	(8)	(53)	(11)	(97)
None	.3669	.6424	.1500	----	.3709
(N)	(8)	(5)	(6)		(19)
Total Means	.4897	.6515	.6472	.0061	.5178
(N)	(119)	(20)	(116)	(28)	(283)

N=283

Table 20
One-Way Analysis of Variance Results
for Participant Incentive

Effect		Sum of	
Level	df	Squares	F
Reaction	5	2.8344	3.354*
Learning	3	.7974	.394
Behavior	5	8.5298	5.099*
Results	4	2.9895	.395
Combined Levels	5	7.4450	3.458*

*p<.05

are more motivated to achieve.

In terms of dependent variables some significant and interesting results appear. Reaction effect size is significantly enhanced when released time, status, or college credit are offered as incentives. When the participant's incentive is college credit, status, or pay, the behavior mean effect size for the group is .8131. In terms of learning, college credit seems to increase the effect size more than other incentives, but this finding is not statistically significant. Finally, in relationship to results, status appears to be the most powerful incentive for this variable. Again, however, because the number of cases is

small, this finding is not significant.

Funding--Tables 21 and 22

The most frequent funding sources reported for inservice programs were state and federal governments. As reported in Table 21, governments funded the majority of the programs in which funding sources were reported.

The largest effect sizes were reported by programs funded through state or federal governments, a university, or the category labelled "other." These categories taken together resulted in a mean effect size of .5560. The other funding sources (participant, school, and combination) resulted in substantially lower effect sizes (mean ES=.2261).

In terms of dependent variables, it is interesting to note that the state government and university funded programs produced unusually high effect sizes in the results category (mean ES=.6007).

Instructor--Tables 23 and 24

Within the variable "inservice instructor," the largest proportion of studies are those taught by college personnel, followed by self-instruction programs. The

Table 21
Mean Effect Size by Funding by Effect Level

Funding	Effect Levels				Total
	Reaction	Learning	Behavior	Results	Means
Participant	.1948	.0000	----	----	.1623
(N)	(10)	(2)			(12)
School	.5801	.5473	.2818	-.9200	.2674
(N)	(19)	(3)	(24)	(6)	(52)
Fed. Gov't.	.4908	1.1814	.6698	.3361	.5836
(N)	(96)	(14)	(104)	(34)	(248)
State Gov't.	.2914	.6449	.3236	.6196	.4863
(N)	(18)	(9)	(29)	(51)	(107)
University	.8770	1.8240	.7900	.4084	.6916
(N)	(1)	(1)	(1)	(5)	(8)
Other	.6993	----	----	----	.6993
(N)	(7)				(7)
Combination	----	----	----	.0533	.0553
(N)				(8)	(8)
Total Means	.4709	.8900	.5481	.3843	.5056
(N)	(151)	(29)	(158)	(104)	(442)

N=442

Table 22
One-Way Analysis of Variance Results for Funding

Effect		Sum of	
Level	df	Squares	F
Reaction	5	2.1365	2.585*
Learning	4	4.5380	1.512
Behavior	3	4.7612	2.550
Results	4	13.9889	7.521*
Combined Levels	6	7.8460	2.820*

*p<.05

Table 23
Mean Effect Size by Instructor by Effect Level

Instructor	Effect Level			Results	Total Means
	Reaction	Learning	Behavior		
Teacher	.2921	.3583	.2750	----	.3010
(N)	(23)	(7)	(27)	----	(57)
Sup. Staff	.1917	----	1.3513	.6417	.6097
(N)	(9)		(3)	(48)	(60)
Administrator	.0000	----	----	----	.0000
(N)	(5)				(5)
Consultant	.2256	----	----	-1.9930	-.2864
(N)	(10)			(3)	(13)
Coll. Pers.	.4375	.7238	.4577	.1892	.4418
(N)	(143)	(34)	(135)	(44)	(356)
Self	1.2263	2.0640	.9423	.3420	.9297
(N)	(3)	(7)	(92)	(17)	(119)
State	.2111	1.0230	.0000	.0000	.2818
(N)	(13)	(2)	(1)	(1)	(17)
Other	.9181	----	----	1.4745	1.0294
(N)	(8)			(2)	(10)
Total Means	.4067	.9141	.6138	.3644	.5228
(N)	(214)	(50)	(258)	(115)	(637)

N=637

Table 24
One-Way Analysis of Variance Results for Instructor

Effect		Sum of	
Level	df	Squares	F
Reaction	7	6.6141	5.443*
Learning	3	10.9720	4.393*
Behavior	4	19.4948	9.368*
Results	5	24.3209	14.325*
Combined Levels	7	40.3660	14.043*

*p<.05

largest effect sizes are for the category "other" and for self-instruction programs. It is interesting to note that consultants as instructors achieved a negative effect size, but the small number of studies makes it difficult to draw conclusions from this sample. Teachers as instructors of other teachers are generally accepted as an effective in-service training arrangement. In this study, however, teachers as instructors achieved only a small positive effect size.

When effect levels are studied in terms of various instructional leaders, many interesting results emerge. Supervisory staff appear to cause greater behavior change and results than most other instructors. The supervisory

staff, however, achieved a relatively low effect size at the reaction level. Consultants produce a large negative effect size at the results level, but a small positive effect at the reaction level. There are so few training sessions led by consultants that these results warrant further investigation before conclusions can be drawn. Self-instruction produces effect sizes well above the grand mean effect sizes for reaction, learning, and behavior. Self-instruction produces highly positive results relative to other instructors and to effect levels.

Type of Structure--Tables 25 and 26

Staff meetings and independent study produce the highest effect sizes of the structures examined, however, there were only five cases which used the staff meeting as a structure. There do not appear to be important differences in the effect sizes among workshops, courses, mini-courses, or institutes which all report effect sizes in the medium positive range. In one-to-one structures, the effect size was considerably lower than those reported for other structures.

In terms of effect levels, participants had the most positive reaction effect levels from independent study and staff meeting arrangements and the least positive reaction effect levels from one-to-one structures. Behavior effect

Table 25

Mean Effect Size by Type of Structure by Effect Level

Structure	Effect Levels			Results	Total Means
	Reaction	Learning	Behavior		
Workshop	.3337	.7630	.4920	.3526	.4110
(N)	(128)	(24)	(87)	(96)	(355)
Course	.4507	.8278	.3147	.0000	.4223
(N)	(52)	(12)	(55)	(1)	(120)
Staff Mtg	.9912	----	----	----	.9912
(N)	(5)				(5)
Mini Course	.3440	.3175	.5736	.2100	.5035
(N)	(5)	(2)	(46)	(7)	(60)
Indep Study	1.0463	2.0640	1.0308	.4880	.9809
(N)	(4)	(7)	(80)	(24)	(115)
Instit	.6313	.5156	.4670	.3933	.5849
(N)	(32)	(7)	(2)	(4)	(45)
One-To-One	.0573	----	.2050	----	.1846
(N)	(4)		(25)		(29)
Total Means	.4237	.9027	.5933	.3682	.5190
(N)	(230)	(52)	(295)	(132)	(709)

N=709

Table 26
One-Way Analysis of Variance Results for Structure

Effect		Sum of	
Level	df	Squares	F
Reaction	6	6.1833	5.768*
Learning	4	11.7097	3.646
Behavior	5	24.2949	10.950*
Results	4	.6809	.331
Combined Levels	6	18.493	7.797*

* $p < .05$

levels appear to be most influenced by the independent study structure and least by the one-to-one structure. The learning and results effect levels did not yield important differences in terms of types of structures.

Length of Treatment--Tables 27 and 28

Length of treatment does not yield significantly different effect sizes as reported in Table 28. The mean effect sizes, however, gradually decrease as the length of treatment increases.

In terms of effect levels, short treatments (1-10 hours) result in larger reaction level effect sizes than longer treatments. The sample is small, however. Learning and results effect levels have an erratic pattern in terms of length of treatment. The behavior effect level does not appear to be influenced by the length of treatment.

Emphasis--Tables 29 and 30

The emphasis of an inservice training program does not significantly influence the effect size results. Both theoretical and practical approaches to inservice training are effective, but practical approaches are much more common than theoretical or combinations of theory and practical approaches.

Table 27

Mean Effect Size by Length of Treatment by Effect Level

Length	Effect Levels			Results	Total Means
	Reaction	Learning	Behavior		
1-10 hours	1.0705	.3936	.5035	.5934	.5514
(N)	(4)	(8)	(31)	(16)	(59)
11-20 hours	.4543	.9663	.5777	.2306	.5080
(N)	(42)	(7)	(46)	(15)	(110)
21-30 hours	.4968	.3175	.5588	.4700	.5074
(N)	(17)	(2)	(67)	(77)	(163)
>30 hours	.4943	.9128	.5101	-.1587	.4971
(N)	(122)	(20)	(93)	(14)	(249)
Total Means	.4979	.7785	.5361	.3846	.5076
(N)	(185)	(37)	(237)	(122)	(581)

N=581

Table 28
 Analysis of Variance Results for Length of Treatment
 and for Effect Level

Source of Variation	df	Sum of Squares	F
Length	3	1.332	1.261
Effect Level	3	4.990	4.724*
Two-Way Interactions	9	8.763	2.765*

*p<.05

Table 29
 Mean Effect Size by Emphasis by Effect Level

Emphasis	Effect Levels				Total Means
	Reaction	Learning	Behavior	Results	
Theory/Both	.5774	1.2466	.3633	.2587	.5097
(N)	(68)	(5)	(36)	(12)	(121)
Practical	.4447	.9603	.6524	.3947	.5716
(N)	(115)	(37)	(248)	(112)	(512)
Total Means	.4940	.9944	.6158	.3815	.5598
(N)	(183)	(42)	(284)	(124)	(633)

N=633

Table 30
 Analysis of Variance Results for Emphasis
 and for Effect Level

Source of Variation	df	Sum of Squares	F
Emphasis	1	.000	.000
Effect Level	3	7.647	5.349*
Two-Way Interactions	3	3.793	2.653*

*p<.05

Practical approaches result in considerably higher behavior effect levels than do theoretical or theory and practical approaches. Theoretical and theory and practical approaches, however, result in greater learning effect levels, although both are effective in terms of learning levels.

Schedule--Tables 31 and 32

The schedules which produced the highest effect sizes are weekends and evenings. Weekends and evenings both had too few cases, however, from which to draw conclusions. The least effective time for training appears to be a combination of times and before and after work.

Table 31
 Mean Effect Size by Schedule by Effect Level

Schedule	Effect Levels			Total Means	
	Reaction	Learning	Behavior Results		
Bef/Aft Work	.4508	.4246	.5623	.2100	.4825
(N)	(32)	(13)	(46)	(7)	(98)
Evenings	.8380	----	----	----	.8380
(N)	(4)				(4)
Summer	.5613	.5531	.4670	.3146	.5370
(N)	(46)	(8)	(2)	(5)	(61)
School Day	.2757	----	.8695	.5942	.5296
(N)	(39)		(28)	(6)	(73)
Combination	.7193	1.4630	.4219	.3192	.4654
(N)	(20)	(5)	(107)	(37)	(169)
Weekends	.6197	1.7760	.7060	----	.8855
(N)	(4)	(2)	(4)		(10)
Total Means	.4738	.7433	.5300	.3349	.4989
(N)	(145)	(28)	(187)	(55)	(415)

N=415

Table 32

One-Way Analysis of Variance Results for Schedule

Effect		Sum of	
Level	df	Squares	F
Reaction	5	3.4706	3.912*
Learning	3	6.3324	3.271*
Behavior	4	4.6577	3.959*
Results	3	.5236	1.151
Combined Levels	5	5.0880	3.974*

*p<.05

In terms of effect level, evenings, combinations, and weekends result in higher effect sizes than other schedules. Reaction effect sizes are increased when training takes place in the evening, a combination of times, or during the weekend. Learning effect sizes are positively influenced by weekends and combinations. Behavior effect sizes are most positively influenced by scheduling training during the school day and on weekends. The results effect size is positively influenced by training taking place during the school day, but there are too few cases for statistically significant conclusions.

Duration--Tables 33 and 34

Effect size is not significantly influenced by the duration of the training, just as it was not influenced by the number of hours of training. Short-term training results in a slightly higher effect size than long-term training.

Activities--Tables 35 and 36

Although training programs composed of activities pursued by the entire training group result in a higher effect size mean than individualized training programs, the differences are not significant.

Similarly, effect level differences are not significant, but individualized programs result in considerably lower effect sizes in terms of behavior effect level than do programs with common activities.

Participant Role--Tables 37 and 38

An active role by participants in a training program results in a higher mean effect size than a passive role, but these differences are not significant.

In terms of effect levels, some interesting differences appear. When participants are mostly receptive, the reaction effect level is large. In terms of the learning effect level, an active role appears to increase effect

Table 33
Mean Effect Size by Duration by Effect Level

Duration	Effect Levels			Total Means	
	Reaction	Learning	Behavior Results		
Long-Term	.6323	.6708	.3064	.4218	.4142
(N)	(29)	(5)	(76)	(75)	(185)
Short-Term	.3928	.9208	.6893	.3029	.5542
(N)	(198)	(46)	(216)	(56)	(516)
Total Means	.4234	.8963	.5896	.3710	.5173
(N)	(227)	(51)	(292)	(131)	(701)

N=701

Table 34
Analysis of Variance Results for Duration
and for Effect Level

Source of Variation	df	Sum of Squares	F
Duration	1	.243	.558
Effect Level	3	3.743	2.870*
Two-Way Interactions	3	9.045	6.936*

*p<.05

Table 35
Mean Effect Size by Activities by Effect Level

Activities	Effect Levels			Total Means
	Reaction	Learning	Behavior	
Individualized	.4354	.7620	.2872	.2623
(N)	(70)	(3)	(57)	(17)
Common	.4217	.9310	.6877	.3872
(N)	(156)	(45)	(227)	(114)
Total Means	.4260	.9204	.6073	.3710
(N)	(226)	(48)	(284)	(131)

N=689

Table 36
Analysis of Variance Results for Activities
and for Effect Level

Source of Variation	df	Sum of Squares	F
Activities	1	.995	2.232
Effect Level	3	3.041	2.274
Two-Way Interactions	3	4.072	3.044*

*p<.05

Table 37
Mean Effect Size by Participant Role by Effect Level

Participant Role	Effect Levels				Total Means
	Reaction	Learning	Behavior	Results	
Active	.4685	1.1048	.6533	.4064	.5738
(N)	(177)	(32)	(250)	(109)	(568)
Receptive	.7945	.5140	.2499	.2381	.3884
(N)	(11)	(7)	(31)	(7)	(56)
Total Means	.4876	.9988	.6088	.3962	.5572
(N)	(188)	(39)	(281)	(116)	(624)

N=624

Table 38
Analysis of Variance Results for Participant Role
and for Effect Level

Source of Variation	df	Sum of Squares	F
Role	1	1.525	3.231
Effect Level	3	3.973	2.806*
Two-Way Interactions	3	4.794	3.386*

*p<.05

size. Active roles also appear to increase effect size in terms of the behavior and results effect levels.

Goals--Tables 39 and 40

Whether goals in an inservice training program are shared/collective or personal do not significantly affect the effect size. Effect levels, however, are influenced significantly by the goals. Some interesting findings, reported in Table 39, are that learning effect levels are extremely high for personal goals and extremely low at the results level in terms of personal goals. These findings both represent too few cases, however, to draw meaningful conclusions.

Group Focus--Tables 41 and 42

Whether inservice training programs were presented to faculties as a unit or to a group of unrelated individuals did not significantly alter effect size. Mean effect sizes were higher for unrelated groups of individuals than for faculties as a unit.

Status--Tables 43 and 44

A statistically significant and somewhat surprising finding is that a higher mean effect size results from inservice training where the leader assumes a role in which

Table 39
Mean Effect Size by Goals by Effect Level

Goals	Effect Levels			Total Means
	Reaction	Learning	Behavior	
Shared	.3926	.7998	.6485	.4144
(N)	(179)	(46)	(249)	(123)
Personal	.5137	1.8164	.2937	-.2627
(N)	(53)	(5)	(45)	(9)
Total Means	.4204	.5137	.5942	.3682
(N)	(232)	(53)	(294)	(132)

N=709

Table 40
Analysis of Variance Results for Goals
and for Effect Level

Source of Variation	df	Sum of Squares	F
Goals	1	.029	.067
Effect Level	3	17.927	13.960*
Two-Way Interactions	3	12.889	10.037*

*p<.05

Table 41
Mean Effect Size by Group Focus by Effect Level

Group Focus	Effect Levels			Total Means	
	Reaction	Learning	Behavior Results		
Fac. as Unit	.3986	.7326	.4459	.3830	.4296
(N)	(88)	(19)	(107)	(102)	(316)
Individuals	.4357	1.0006	.6916	.3180	.5956
(N)	(145)	(33)	(189)	(30)	(397)
Total Means	.4217	.9027	.6028	.3682	.5220
(N)	(233)	(52)	(296)	(132)	(713)

N=713

Table 42
Analysis of Variance Results for Group Focus
and for Effect Level

Source of Variation	df	Sum of Squares	F
Group Focus	1	1.485	3.373
Effect Level	3	11.319	8.570*
Two-Way Interactions	3	2.499	1.892

*p<.05

Table 43
Mean Effect Size by Status by Effect Level

Status	Effect Levels				Total Means
	Reaction	Learning	Behavior	Results	
Part-Ldr Equal	.4220	.5747	.4791	.4357	.4545
(N)	(93)	(12)	(123)	(77)	(305)
Super-Sub	.5573	1.3540	.7363	.3065	.6620
(N)	(98)	(22)	(134)	(42)	(296)
Total Means	.4914	1.0789	.6132	.3901	.5567
(N)	(191)	(34)	(257)	(119)	(601)

N=601

Table 44
Analysis of Variance Results for Status
and for Effect Level

Source of Variation	df	Sum of Squares	F
Status	1	5.378	11.439*
Effect Level	3	10.340	7.332*
Two-Way Interactions	3	5.857	4.153*

p<.05

he/she is the major "giver of information." In classes where participants were encouraged to teach each other through classroom presentations, group work, discussion sessions, etc., a lower effect size resulted.

In terms of effect level, one of the important differences occurs in inservice activities measuring the learning level. Learning level has a significantly higher effect size when the leader assumes the role of major imparter of knowledge than when the leader and the participants both assume responsibility for teaching.

Follow-Up--Tables 45 and 46

Most inservice programs had no follow-up and these programs resulted in a statistically significant higher effect size than those with follow-up.

In terms of effect levels some big differences occur at the levels of behavior and results. At the behavior level, a higher effect size results by giving no assistance later. At the results level, a higher effect size also results by giving no assistance later. In fact, a negative effect size occurs at the results level when assistance is given after a training program.

Assignment to Groups--Tables 47 and 48

Higher effect sizes were found for groups which were

Table 45
Mean Effect Size by Follow-Up by Effect Level

Follow-up	Effect Levels			Total Means	
	Reaction	Learning	Behavior		
Assist Later	.4047	.7508	.3006	-.0913	.2958
(N)	(58)	(4)	(78)	(22)	(162)
No Assist Later	.4290	.9419	.7408	.4601	.5939
(N)	(169)	(43)	(188)	(110)	(510)
Total Means	.4228	.9256	.6117	.3682	.5220
(N)	(227)	(47)	(266)	(132)	(672)

N=672

Table 46
Analysis of Variance Results for Follow-Up
and for Effect Level

Source of Variation	df	Sum of Squares	F
Follow-Up	1	3.947	9.064*
Effect Level	3	8.734	6.686*
Two-Way Interactions	3	5.567	4.262*

p<.05

Table 47

Mean Effect Size by Assignment to Groups by Effect Level

Assignment	Effect Levels			Results	Total Means
	Reaction	Learning	Behavior		
Random	.7476	2.2119	.7250	.3148	.7502
(N)	(48)	(10)	(143)	(25)	(226)
Matching	.4179	----	.7250	.3299	.6437
(N)	(7)		(58)	(10)	(75)
Eq. Pretest	----	----	.3090	.5348	.3603
(N)			(17)	(5)	(22)
Non Equiv.	.3616	.8941	.3713	.1032	.3574
(N)	(118)	(15)	(44)	(36)	(213)
No Control	.2796	.4225	.3264	.5543	.3965
(N)	(60)	(27)	(36)	(56)	(179)
Total Means	.4217	.9027	.6009	.3682	.5215
(N)	(233)	(52)	(298)	(132)	(715)

N=715

Table 48
One-Way Analysis of Variance Results
for Assignment to Groups

Effect		Sum of	
Level	df	Squares	F
Reaction	3	6.7348	13.041*
Learning	2	23.3667	21.949*
Behavior	4	9.5781	4.849*
Results	4	4.6931	2.429
Combined Levels	4	19.5380	11.652*

*p<.05

randomly assigned to treatment groups and for those assigned to groups by matching. The other methods of assignment all resulted in substantially lower effect sizes.

Internal Validity--Tables 49 and 50

When internal validity is high (i.e., treatment groups are randomly assigned and mortality is less than fifteen percent) the effect size is highest. As internal validity decreases, so does effect size.

Reactivity Level--Tables 51 and 52

Table 49
 Mean Effect Size by Internal Validity by Effect Level

Internal Validity	Effect Levels			Results	Total Means
	Reaction	Learning	Behavior		
High	.7459	2.2119	.7160	.2140	.7372
(N)	(44)	(10)	(146)	(23)	(223)
Medium	.4821	.3890	.6106	.1604	.5038
(N)	(21)	(1)	(78)	(21)	(121)
Low	.3292	.5959	.3636	.4581	.3961
(N)	(168)	(41)	(74)	(88)	(371)
Total Means	.4217	.9027	.6009	.3682	.5215
(N)	(233)	(52)	(298)	(132)	(715)

N=715

Table 50
 Analysis of Variance Results for Internal Validity
 and for Effect Level

Source of Variation	df	Sum of Squares	F
Internal Validity	2	20.419	25.612*
Effect Level	3	8.002	6.691*
Two-Way Interactions	6	22.201	9.282*

*p<.05

The largest effect sizes occur for low reactivity levels. Low reactivity refers to evaluation procedures which use standardized tests and blind testing conditions in order to ensure that the testing situation is not biased. Medium reactivity refers to moderately controlled evaluation methods (i.e., experimenter designed tests and self-reports from participants). The mean effect size reported for medium reactivity is only slightly lower than the effect size for low reactivity. Quite surprisingly, when conditions are experimenter controlled (i.e., the experimenter has an opportunity to influence the results) the mean effect size is actually negative.

Table 51

Mean Effect Size by Reactivity Level by Effect Level

Reactivity Level	Effect Levels			Results	Total Means
	Reaction	Learning	Behavior		
Cont'l Exp	.0683	.4225	.1474	-1.9930	-.1303
(N)	(3)	(2)	(14)	(3)	(22)
Med Cont'l Exp	.4225	.9781	.6437	.4249	.5379
(N)	(226)	(29)	(240)	(107)	(602)
Low	.6405	.8442	.5120	.4142	.5707
(N)	(4)	(21)	(44)	(22)	(91)
Total Means	.4217	.9027	.6009	.3682	.5215
(N)	(233)	(52)	(298)	(132)	(715)

N=715

Table 52
 Analysis of Variance Results for Reactivity Level
 and for Effect Level

Source of Variation	df	Sum of Squares	F
Reactivity	2	11.354	13.581*
Effect Level	3	15.577	12.421*
Two-Way Interactions	6	10.299	4.106*

*p<.05

In terms of effect levels, effect sizes are scattered over a wide range. At the results effect level, an effect size of -1.9930 is reported for experimenter controlled activity. This is a substantially lower effect size than others reported at the results effect level. However, there are so few cases, that strong conclusions cannot be made.

Source of Measurement--Tables 53 and 54

When inservice programs are evaluated using a combination of measurements that are experimenter developed and published, the effect size is larger than when using only experimenter developed measurements or only published mea-

Table 53
 Mean Effect Size by Source of Measurement
 by Effect Level

Source of Measurement	Effect Levels			Total Means	
	Reaction	Learning	Behavior		
Exper. Devel.	.4127	.9585	.6058	.4253	.5464
(N)	(89)	(30)	(182)	(93)	(394)
Published	.3732	.8265	.5985	.0073	.4644
(N)	(116)	(22)	(115)	(28)	(281)
Combination	.6511	-----	-----	.8424	.6894
(N)	(28)			(7)	(35)
Other	-----	-----	.0000	.7373	.5898
(N)			(1)	(4)	(5)
Total Means	.4217	.9027	.6009	.3682	.5215
(N)	(233)	(52)	(298)	(132)	(715)

N=715

Table 54
One-Way Analysis of Variance Results
for Source of Measurement

Effect		Sum of	
Level	df	Squares	F
Reaction	2	1.7538	4.542*
Learning	1	.2210	.224
Behavior	2	.3661	.351
Results	3	6.0702	4.318*
Combined Levels	3	4.2420	3.212*

*p<.05

surement techniques.

In terms of effect levels, the reaction effect level shows the highest effect sizes when a combination of measurements are used. The results effect level also reports high effect sizes for combination measurements.

Type of Measurement--Tables 55 and 56

It is interesting to note that behavioral measurements result in higher effect sizes than other reported measurements. The lowest effect size mean results from self-report measurements.

Table 55
 Mean Effect Size by Type of Measurement
 by Effect Level

Type of Measurement	Effect Levels				Total Means
	Reaction	Learning	Behavior	Results	
Self Report	.4642	.1265	.1636	.0240	.3986
(N)	(164)	(2)	(10)	(21)	(197)
Behv Obsrv	.0683	----	.6212	.4416	.6064
(N)	(3)		(265)	(14)	(282)
Partic Tst Res	.0962	.9337	.5798	----	.5692
(N)	(39)	(50)	(21)		(110)
Std Tst Res	----	----	----	.4120	.4120
(N)				(86)	(86)
Combination	.6726		.3225	.5897	.6323
(N)	(27)		(2)	(11)	(40)
Total Means	.4217	.9027	.6009	.3682	.5215
(N)	(233)	(52)	(298)	(132)	(715)

N=715

Table 56
One-Way Analysis of Variance for Type of Measurement

Effect		Sum of	
Level	df	Squares	F
Reaction	3	6.5026	12.518*
Learning	1	1.2531	1.300
Behavior	3	2.1857	1.409
Results	3	3.2679	2.221
Combined Levels	4	5.6400	3.213*

*p<.05

Effect Formula--Tables 57 and 58

The effect formula used to find effect size had a statistically significant impact on effect size results. Use of the F test formula produced the highest effect sizes, while the alpha-level formula produced the lowest.

The original formula of dividing the differences of treatment and control means by the standard deviation of the control group was used in 210 cases. The strength of this formula is not diluted by estimation, therefore the effect size mean of .5927 is extremely significant. The total effect size mean indicates that inservice teacher education is effective.

Table 57

Mean Effect Size by Effect Formula by Effect Level

Effect Formula	Effect Levels			Results	Total Means
	Reaction	Learning	Behavior		
Mean and SD	.4988	1.7192	.6398	.0784	.5927
(N)	(27)	(11)	(140)	(32)	(210)
T Test	.5233	2.5050	.8102	.8048	.7112
(N)	(29)	(2)	(15)	(4)	(50)
F Test	.6079	1.8273	1.0285	.8146	.8183
(N)	(68)	(4)	(49)	(10)	(131)
Alpha Level	.2219	.5376	.2760	.0000	.2575
(N)	(47)	(8)	(25)	(4)	(84)
Corr w/ Signif	.2877	.4225	.2906	.4235	.3501
(N)	(62)	(27)	(69)	(82)	(240)
Total Means	.4217	.9027	.6009	.3682	.5215
(N)	(233)	(52)	(298)	(132)	(715)

N=715

Table 58
 Analysis of Variance for Effect Formula
 and for Effect Level

Source of Variation	df	Sum of Squares	F
Effect Formula	4	32.126	21.339*
Effect Level	3	22.314	19.762*
Two-Way Interactions	12	25.100	5.557*

*p<.05

Types of Instruction--Table 59

All data sets were classified according to the types of instruction used in the programs. Fifteen categories of instruction were represented. Most studies used more than one type of instruction in order to carry out the inservice activity.

Table 59 reports the mean effect sizes for those studies using a specific type of instruction (column 2) as well as a mean effect size for those studies not using that particular instructional method (column 3). The information in Table 59 presents mean effect sizes for all studies without regard to effect level. Those instructional methods which differed significantly from the mean of those

Table 59
Mean Effect Size by Type of Instruction

Type of Instruction	With (N)	Without (N)	Type of Instruction	With (N)	Without (N)
Lecture	.4396+.035 (347)	.5987+.037* (368)	Practice	.5577+.030 (589)	.3521+.039* (126)
Games/Sim.	.4077+.062 (152)	.5522+.028* (563)	Printed Mat.	.5268+.034 (379)	.5156+.038 (336)
Discussion	.4511+.034 (356)	.5913+.037* (359)	Prog. Study	.6705+.080 (23)	.5165+.026 (692)
Modeling	.5175+.042 (372)	.5258+.028 (343)	Observation	.8182+.066 (165)	.4325+.025* (550)
Coaching	.5243+.035 (225)	.5202+.033 (490)	Prod. Inst. Mat.	.4519+.105 (70)	.5290+.026 (645)
Video/Audio	.6421+.053 (228)	.4650+.027* (487)	Guid. Fld. Trip	.1027+.515 (14)	.5299+.024* (701)
Mut. Assist.	.4661+.039 (276)	.5563+.033 (439)	Film	.4650+.042 (173)	.5395+.031 (542)
Micro Teaching	.7816+.078 (133)	.4620+.025* (582)			

*p<.05

studies not using a particular instructional method are identified.

Observation followed by micro teaching were the types of instruction which produced the highest effect sizes. Other types of instruction which significantly enhanced the magnitude of the effect were video/audio feedback and practice.

Instructional methods which were associated with significantly lower effect sizes are discussion, lecture, games/simulation, and guided field trips. There were only fourteen cases which used guided field trips so caution should be observed in generalizing from this information.

Modeling and coaching, which have received much attention in the staff development literature as potentially powerful instructional tools, did not significantly alter effectiveness.

Summary

This chapter has provided an abundance of data on the effect sizes achieved through various inservice treatments. The salient findings include:

1. Inservice teacher education programs are generally effective as indicated by a .5213 grand mean effect size for 715 cases.

2. Clearly, programs which measured outcomes in terms of participant learning or behavior change had a substantially higher success rate than those which sought to measure results in terms of the participant reactions or student outcomes.
3. The number of participants in an inservice training program, the number of treatment hours, and the length of the treatment period do not significantly influence effect size results.
4. Outside-school originated programs (state and federal governments, university researchers, and consultants) were more effective than in-school originated programs (participant, administrator/supervisor, and school). Similarly, those programs funded by outside groups achieved higher effect size results than those funded by the school or the participants. Whether training sessions were held on-site or off-site made no significant difference in results.
5. When inservice training participants include both elementary and secondary participants, the effect size results are larger than for either group individually.
6. Status, (special recognition, a competitive selection process, or special designation) followed by college credit, are the incentives most likely to increase effect size.

7. Self-instruction produces the highest effect size results in terms of instruction followed by supervisory staff. Teachers, administrators, consultants, and state department of education personnel are not effective instructors in terms of effect size results.
8. Staff meetings and independent study as structures for inservice training are quite effective while the workshop or one-to-one training are considerably less effective.
9. Inservice training programs in which the leader assumes the role of "giver of information" and the participants are seen as "receivers of information" are more effective than programs where participants are seen as major contributors to the learning process.
10. Inservice training programs which use observation, micro teaching, video/audio feedback, or practice produce greater effects than those programs not using these instructional methods. Discussion, lecture, games/simulations, and guided field trips yielded lower effect sizes than the other training methods examined in this study.

C H A P T E R V
DISCUSSION AND IMPLICATIONS

This chapter will focus on four tasks:

1. Answering the four questions in Chapter I which were all related to determining the various features of inservice teacher training programs which would result in higher effect sizes.
2. Analyzing the findings of this meta-analysis in terms of the degree to which they corroborate or refute other staff development research findings.
3. Using the findings of this meta-analysis to suggest implications for staff developers.
4. Making recommendations for future study.

Discussion will focus on those findings that are most significant in terms of providing effective inservice programs for teachers.

Analysis by Dependent Variables

To determine which variables most affected the four evaluation levels (reaction, learning, behavior, and results), the data were examined in terms of effect size by effect level. When findings were significant, they were

listed as variables which elicited the most positive reactions, learning, behavior, or results.

Reaction

Those variables which elicited the most positive feelings or reactions from participants include training sessions which had the following characteristics:

1. Affective techniques as the instructional focus.
2. Participant groups ranging in size from forty-one to sixty.
3. Initiation by the state or federal government.
4. Participation by a group containing both elementary and secondary teachers.
5. Incentives for participation which enhance teachers' status.
6. Funding provided by the school.
7. Self-instruction and independent study.
8. A schedule which includes a variety of times (i.e., evenings, weekends, before/after work, etc.) to meet.
9. Instruction lasting less than six months.
10. Receptive rather than active participant roles.
11. Participant goals which were personal rather than shared.
12. A leader who takes almost exclusive responsibility for

the design and teaching of the class.

Learning

Those variables which elicited the greatest effect in terms of the learning that resulted from a staff development activity had the following characteristics:

1. An instructional focus on affective techniques.
2. Participation by a group containing both elementary and secondary teachers.
3. Self-instruction as the instructional method.
4. A schedule which includes a variety of meeting times or the weekends.
5. Instruction lasting less than six months.
6. An active rather than passive participant role.
7. A leader who takes most of the responsibility for teaching rather than allowing participants to teach.

Behavior

The variables which contributed the most to behavior change as a result of inservice teacher training had the following traits:

1. An instructional focus on improving general teaching.
2. Participant groups ranging in size from twenty-one to

forty members.

3. Initiation by a university researcher.
4. Participation by a group containing both elementary and secondary teachers.
5. Incentives for participation which enhance teachers' status.
6. Instruction through self-instructional methods.
7. An independent study structure.
8. Instruction focused on practical application.
9. Instruction scheduled during the school day.
10. Training lasting less than six months.
11. Instructional activities which were common for all participants rather than individualized.
12. An active rather than passive participant role.
13. Participant goals which were shared by all rather than personal ones.
14. Unrelated participants rather than a particular faculty unit.
15. A leader who takes almost exclusive responsibility for the design and teaching of the class.
16. No assistance is given after the initial training.

Results

The variables which elicited the most positive results from participants or their students had the following char-

acteristics:

1. An instructional focus on affective techniques.
2. Initiation by an administrator/supervisor.
3. Participation by a group containing only elementary teachers.
4. Funding provided by the state government.
5. Instruction provided by members of a school's support staff.
6. Instruction lasting one to ten hours.
7. Instruction focused on practical application.
8. Training lasting longer than six months.
9. Shared goals rather than personal ones.
10. Contributions to learning are made by both the instructor and the participants.
11. No follow-up or assistance is given after the initial training.

Conclusions

When the results of this meta-analysis are broken down by effect levels, some important differences appear. Perhaps the most important finding in this context is not the specific findings but the fact that the different levels of evaluation do sometimes produce quite different effect sizes for a given independent variable. Staff developers

would be wise to plan inservice activities which take into account the various features of staff development programs which tend to produce higher effect sizes at the evaluation level that will be employed.

Staff Development Research Findings

Some of the information coded for this meta-analysis has been examined by other researchers. As mentioned in Chapter II, Joslin (1980) and Lawrence and Harrison (1980) have conducted meta-analyses that asked some of the same questions studied here. Many other researchers have also examined questions that were studied in this meta-analysis. The discussion which follows examines findings which corroborate existing staff development literature as well as some findings which refute some published speculations.

Dependent Variables

Clearly, programs which measure outcomes in terms of participant learning produce significantly higher effectiveness ratings. The grand mean effect size for participant learning was calculated to be .9027, a large positive effect size according to Cohen's (1969) terminology. As outlined in Table 60, Joslin's (1980) work corroborates this finding as does the work of Lawrence and Harrison

Table 60

A Comparison of Mean Effect Size by Effect Level
as Determined by Wade, Joslin, and Lawrence and Harrison

Researcher	No. Cases	Reaction	Learning	Behavior	Results	Grand Mean
Wade	715	.4217	.9027	.6009	.3682	.5215
		Attitudes/ Percept.	Ach./ Knowl.	Skills/ Beh.	Student Outcomes	
Joslin	902	.184	2.101	.614	.127	.472
		Affect.	Cogn.	Perf.	Conseq.	
Lawrence & Harrison	82	.85	1.02	.74	.50	.76

(1980).

Other findings related to effect level which are noted in this study and verified by Joslin (1980) and Lawrence and Harrison (1980) are:

1. Programs which focus on behavior change result in medium to large positive effect sizes. This study and Joslin's reported that programs focusing on behavior change result in the second highest effect size by effect level. Lawrence and Harrison noted that performance-based programs ranked third in terms of effectiveness.

2. All three research studies were in agreement that results in terms of students of participants were difficult to demonstrate, yet all studies showed small positive effects in terms of results.
3. All three studies reported a grand mean effect size in the medium to large positive range.

From these three meta-analyses, the conclusion can be drawn that inservice teacher training has generally positive results. The programs which measure outcomes in terms of learning are significantly more effective than those which measure outcomes in terms of behavior change, reactions, or results. The least effective programs are those which attempt to measure changes in terms of participant or student-test results.

Duration

There has been much speculation in the staff development literature that warns against short-term training sessions. When studies in this meta-analysis were classified according to length of treatment, there were no significant differences found among treatments lasting one to ten hours, eleven to twenty hours, twenty-one to thirty hours, or more than thirty hours. Similarly, no significant differences were found between short-term and long-term effect

sizes. Lawrence and Harrison (1980) found that short-term training (less than six months) is actually more effective than long-term training, but they do not provide significance levels.

Location and Scheduling

This study was unable to corroborate findings by Joslin (1980) which reported that on-site training is significantly more effective than off-site training. Lawrence and Harrison (1980) reported similar findings but did not report levels of significance. This meta-analysis confirmed that on-site training is more effective than off-site training, but the difference was not significant.

In terms of scheduling, the data are inconclusive. Joslin (1980) and Lawrence and Harrison (1980) both concluded that programs held before or after work, including evenings, were more effective than those held during the school day. This study noted that programs held during the work day were slightly more effective than those held after working hours but the difference was not significant. None of the research points to a clear message regarding scheduling.

Process

Chapter II contained a list of processes which have

been associated with effective staff development. Those that were verified by this meta-analysis are listed below:

1. Voluntary participation is more effective than mandatory attendance but the difference is not statistically significant.
2. Involvement of a "critical mass" of teachers has been mentioned by Berman and McLaughlin (1975) as important for successful staff development activities. Although they were referring to a large group within a particular group setting, the evidence here points to the fact that larger groups in any setting may be more effective than smaller ones, but again the results are not statistically significant.
3. Berman and McLaughlin (1975) contended that staff training should focus on the concrete with teacher-specific plans rather than abstract, talk-oriented sessions. Evidence from this meta-analysis points to the fact that practical rather than theoretical instruction, with the instructor taking almost exclusive responsibility for the design and teaching of the class, results in larger effect sizes.

From this study, it appears that when sponsorship of a training program is provided by an outside group (i.e.,

university, federal government, etc.) the effect size is greater than for programs sponsored by the school or the participant. Lawrence and Harrison (1980) also looked at sponsorship and reported similar findings with the exception of school sponsored programs. School sponsored programs, they noted, produce higher effect sizes than programs sponsored by other groups. However, Lawrence and Harrison (1980) analyzed six studies sponsored by a school as opposed to this study which evaluated fifty-two cases.

Incentives for participation in inservice training activities have not been carefully examined in the staff development literature. Joslin (1980) reported that college and local district credit produced higher effect sizes than stipends. This study concurs that college credit produces higher effect sizes than stipends, but also looks at status as an incentive. When a participant was selectively chosen to participate in training, either by being designated as a representative of a particular group or selected through a competitive process, the effect sizes were greater than for all other incentives studied.

Throughout the staff development literature there is widespread acceptance that teachers should be in charge of their own inservice. This includes the belief that teachers should determine their inservice needs and also take responsibility for meeting their own needs. In short,

teachers are often the best teachers of other teachers. This meta-analysis suggests that outside initiated programs (i.e., those initiated by the state/federal government, university researcher, and consultant) produce higher effect sizes than those initiated by the participant, school, or administrator/supervisor. This conclusion is quite controversial and must be tempered by the following caveats: (1) only twelve inservice training programs were initiated by participants, (2) most studies examined (N=355) were initiated by university researchers, and (3) outside groups may have more experience in designing and carrying out training programs which can demonstrate quantitative results. Quantitative research evidence on effectiveness of teacher training is not generally a priority of schools.

Focus and Nature

Joslin (1980) reported that programs directed at elementary school teachers were more effective than those focusing on secondary school teachers. Berman and McLaughlin (1975) also suggested targeting staff development to elementary schools. The findings of this meta-analysis corroborate the fact that greater effects are evident after training a group of elementary teachers versus a group of secondary teachers. However, this study found the greatest

effects were achieved when the training group consisted of both elementary and secondary teachers.

Both Joslin (1980) and Lawrence and Harrison (1980) reported that college courses and institutes resulted in higher effect size means than did workshops. This study corroborated their findings, but found that college courses were not significantly more effective than workshops while institutes were. The structure which was significantly more effective than any other, however, was independent study.

In terms of instructional focus, this meta-analysis found that when the focus was on improving affective techniques, the effect sizes were larger than for either improving general teaching or improving a specific subject. Other researchers have not looked at instructional focus in these terms.

Instruction

Although much attention has been given to the speculative claims by Joyce and Showers (1980) that effective training should include theory, demonstration, practice, feedback, and coaching for application, little support was found for this theory. In fact, when the data were analyzed by the type(s) of instruction which yielded the highest effect sizes, the results indicated that four types

of instruction were significantly more effective than others: video/audio feedback, micro teaching, observation, and practice. Modeling and coaching, which Joyce and Showers (1982) identify as the two most powerful instructional tools, in fact, made no significant difference in effect size in this study.

Furthermore, data analyses indicated that no assistance after initial training actually produced higher effect sizes than when follow-up assistance was provided. This finding must be interpreted cautiously because no consideration was given to the timing of the effect size measurement or the amount of follow-up assistance provided. If, for example, an effect size measurement is taken six months after initial training and only minimal follow-up has been provided, the effect size probably would be lower than if taken immediately following an intensive training program. Lawrence and Harrison (1980) reported that follow-up support produced higher effect sizes than training with no follow-up. They did not disclose when effect size measurements were determined or the amount of assistance provided. Follow-up is an area which requires further study in order to determine the role it plays in effective staff development.

Other outcomes regarding instruction which do not agree with outcomes reported by Lawrence and Harrison

(1980) are: (1) a superordinate-subordinate relationship produces higher effect sizes than an equal relationship between the two parties, and (2) common activities are more effective than individualized ones. In both cases, the number of studies examined for this meta-analysis was six times greater than the cases studied by Lawrence and Harrison (1980). Thus, a superordinate-subordinate relationship and common activities do seem to be instructional strategies which would produce higher effect sizes than their counterparts.

Lawrence and Harrison's (1980) outcomes are in agreement with this meta-analysis that: (1) training activities with group goals produce higher effect sizes than when goals vary according to the personal goals set by participants, and (2) active rather than receptive participant roles produce higher effect sizes. Neither one of these findings, as reported in this study, was statistically significant. Lawrence and Harrison (1980) did not report levels of significance.

Joslin (1980), Lawrence and Harrison (1980) and this study all sought to provide an answer to the question: What leader job category produces the highest effect sizes? The answer is not clear cut, but Joslin (1980) and this study concur that self-instruction is highly effective. Other areas of agreement include: (1) Joslin (1980) and

this study agree that outside consultants produce small positive or negative effect sizes; (2) Joslin (1980) and this study concur that college instructors produce moderately positive effect size results; (3) Lawrence and Harrison (1980) and this study agree that supervisory personnel within a school system produce medium to large effect size results; and (4) Lawrence and Harrison (1980), Joslin (1980), and this study conclude that teachers as instructors produce only small positive effect sizes based upon a modest total of eighty-seven cases.

Implications for Staff Developers

This meta-analysis, along with those done by Joslin (1980) and Lawrence and Harrison (1980) have clearly demonstrated that inservice teacher education programs reported in the literature are moderately effective. Such an outcome should provide some reassurance to staff developers and other educators who wonder whether the time, money, and effort invested in staff development make a difference.

It is also clear that many factors are involved in determining the effectiveness of any given inservice activity. Considerations such as the effect level outcomes which are expected, the instructor, his or her style of instruction, as well as the type of instruction and the

structure must be taken into account. Other considerations include incentives for participation, the funding source, and composition of the training group. Some variables are easily manipulated while others are not. The task for the staff developer becomes manipulating those variables which can be readily manipulated in order to maximize the effect of inservice teacher training.

Staff developers who wish to plan programs for maximum effectiveness should contemplate the following suggestions, which are based upon outcomes of this meta-analysis:

1. Do not be overly concerned about training group size, the number of training hours, or the location of training activities.
2. Allow for initiation of training activities by a variety of people.
3. Encourage elementary and secondary participants to work together.
4. Allow for voluntary participation whenever possible.
5. Structure activities with clear objectives planned for the entire training group.
6. Offer incentives for participation, such as enhanced status or college credit whenever possible.
7. Encourage self-instruction, independent study, or training at staff meetings as alternatives to the tra-

ditional workshop format.

8. Encourage inservice trainers to use instructional techniques such as practice, video/audio feedback, micro teaching, and observation as alternatives to lecture, discussion, games/simulation, and guided field trips.

Although these eight suggestions should guide the staff developer, they are only a part of the total picture needed in order to provide effective inservice education programs. Context issues such as understanding the school climate, principal support, adequate resources, including time and an understanding of the needs, should not be underestimated. Process issues such as governance and teacher investment must not be overlooked. In short, the staff developer must weave through a labyrinthian path to effective inservice teacher education with many guidelines and considerations taken into account along the way. There is no magic formula for successful staff development, but the odds can be increased if staff developers consider the factors which appear to be related to effectiveness.

Recommendations for Future Study

This meta-analysis has provided a wealth of informa-

tion regarding the effectiveness of various features and procedures common in inservice training activities. In some cases, the findings corroborated those of other researchers, in other cases, the results refuted existing research. In some instances, the questions studied and the results found have not been previously examined.

Those findings of this meta-analysis which have not been previously studied, need to be corroborated by other researchers or need further clarification raise the following questions for further study:

1. Are the number of participants in an inservice training program, the number of training hours, and the length of the treatment period insignificant?
2. Does the time when an inservice training program is scheduled influence the outcome or are other variables responsible for the differences found between school day versus after school training programs?
3. Do programs originated outside the school produce higher effect sizes because they have a better design or are teachers truly not appropriate initiators of their own inservice training?
4. Enhanced status seems to be an effective incentive for teachers; what role might enhanced status play in increasing teacher effectiveness?

5. What effects are found when the same inservice training program is provided for a group of elementary teachers, a group of secondary teachers, and a group of both elementary and secondary teachers?
6. Were observing master teachers, practicing new skills and video/audio feedback more effective than modeling and coaching because they aren't as dependent upon the skills of a particular instructor?
7. What role does follow-up play in training effectiveness when measurements of effect are taken on two groups (one with follow-up and one without) at various times following inservice training?

Another task for future researchers is to improve the quality of research done on inservice teacher education. The technique of meta-analysis is powerful, yet when it is applied to a data base that is of uneven quality, the results must be interpreted with caution. Only twenty-nine percent of the 715 data sets provided the mean and standard deviations for experimental and control groups which were necessary for the most accurate calculations of effect size. Almost thirty-four percent of the data sets were extracted from studies which did not use control groups. Clearly, there is a need for more rigorous program evaluation standards.

This meta-analysis helped to answer some questions related to providing effective inservice teacher education, but it also raised some new ones. It is hoped that this study will provide a stimulus for future researchers to shed more light on the intricacies of staff development.

REFERENCES

- Asher, J. Inservice education: Psychological perspectives. Berkeley, California, 1967, ERIC: ED 015 891.
- Barbera, R.M. Catching the kinks in staff development programs. San Antonio, Texas: Intercultural Development Research Association, 1976, ERIC: ED 125 150.
- Beegle, C.W., & Brandt, R.M. (eds.). Observational Methods in the Classroom. Washington, DC: Association for Supervision and Curriculum Development, 1973.
- Bell, T.H. Challenges for inservice vocational education. Theory Into Practice, 1975, 14 (1), 1-4.
- Bell, H.H., Jr., & Peightel, J.W. Teacher centers and inservice education. Bloomington, Indiana: Phi Delta Kappa Education Foundation, 1976, ERIC: ED 120 171.
- Bentzen, M. Changing Organizations: The Magic Feather Principle. New York: McGraw-Hill, 1974.
- Berman, P., & McLaughlin, M. Federal programs supporting educational change (Vol. V: Executive summary). Santa Monica, California: Rand Corporation, 1975.
- Brandt, R. Overview. Educational Leadership. 1982, 40 (4), 3.
- Brethower, K.S., & Rummler, G.A. Evaluating training. Improving Human Performance Quarterly, 1977, 3 (1), 103-

120.

- Burk, J.M. The role of the superintendent in inservice development. Paper presented at the National Council of States on Inservice Education, New Orleans, 1976.
- Chall, J. Restoring dignity and self-worth to the teacher. In Kaleidoscope: Readings in Education, K. Ryan & J.M. Cooper (eds.). Boston: Houghton Mifflin Co., 1980.
- Cobb, M. Inquiry and inservice in curriculum change. In New Directions, New Dimensions, Curriculum and Staff Development, Kentucky Department of Education (ed.). Frankfurt: Kentucky Department of Education, 1973.
- Cohen, J. Statistical Power Analysis for Behavioral Sciences. New York: Academic Press, 1969.
- Devaney, K. Essays on Teachers' Centers. San Francisco: Far West Laboratory, 1977.
- Dewey, J. The relation of theory to practice in education. In The Relation of Theory to Practice in the Education of Teachers. (National Society for Scientific Study of Education, Third Yearbook). Chicago: University of Chicago Press, 1904.
- Draba, R.E. Guidelines for viable inservice education. Journal of Reading, 1975, 18 (5), 368-371.
- Edelfelt, R.A. In-service training of teachers. In The Teachers Handbook, D.W. Allen & E. Seifman (eds.).

Glenview, Illinois: Scott, Foresman, 1971.

Edelfelt, R.A. In-Service Teacher Education--Sources in the ERIC system. Washington, DC: 1975, ERIC: ED 099 308.

Edelfelt, R.A. Six years of progress in inservice education. Journal of Research and Development in Education, 1981, 14 (2), 112-119.

Edelfelt, R.A., & Johnson, M. (eds.). Rethinking In-Service Education. Washington, DC: National Education Association, 1975.

Edelfelt, R.A., & Johnson, M. Teacher designed reform in inservice education. (Final Report of the National Foundation for the Improvement of Education, published by the Department of Health, Education, and Welfare). Washington, DC: United States Government Printing Office, 1977.

Edelfelt, R.A., & Johnson, M. A history of the professional development of teachers. In 1981 Report on Educational Personnel Development. Washington, DC: Feistritz Publications, 1980, 1-76.

Edelfelt, R.A., & Lawrence, G. In-service education: The state of the art. In Rethinking In-Service Education, R.A. Edelfelt & M. Johnson (eds.). Washington, DC: National Education Association, 1975.

- Elliott, J. Developing hypotheses about classrooms from teachers' practical constructs: An account of the work of the Ford Teaching Project. Grand Forks, North Dakota: North Dakota Study Group on Evaluation, University of North Dakota, Center for Teaching and Learning, 1976.
- Elliott, J. How do teachers learn? Unpublished paper, Cambridge Institute of Education, 1979.
- Emrick, J.A., & Peterson, S.M. A synthesis of findings across five recent studies in educational dissemination and change. San Francisco: Far West Laboratory, 1978.
- Frazier, B.W., Betts, G.L., Greenleaf, W.J., Waples, D., Dearborn, N.H., Carney, M., & Alexander, T. National survey of the education of teachers (Vol. 5, Special survey studies) Bulletin 1933, No. 10. Washington, DC: United States Government Printing Office, 1935.
- Fullan, M. The Meaning of Educational Change. New York: Teachers College Press, Columbia University, 1982.
- Gardner, J.W. Self-Renewal: The Individual and the Innovative Society. New York: Harper & Row, 1964.
- Gardner, S. Status of the American Public School Teacher, 1980-81. Washington, DC: National Edu-

- cation Association Research, 1982.
- Glass, G.V. Primary, secondary, and meta-analysis of research. Educational Researcher, 1976, 5 (10), 3-8.
- Glass, G.V. Integrating findings: The meta-analysis of research. In Review of Research in Education. L.S. Shulman (ed.). Itasca, Illinois: Peacock Publishing Company, 1977.
- Glass, G.V. & Smith, M.L. Methods of integrative analysis. Final Report to the National Institute of Education for Grant No. NIE-G-78-1048. Boulder, Colorado: Laboratory of Educational Research, August, 1980.
- Goodlad, J.I. The Dynamics of Educational Change: Toward Responsive Schools. New York: McGraw-Hill, 1975.
- Graubard, P.S., & Rosenberg, H.E. Classrooms That Work: Prescriptions for Change. New York: E.P. Dutton & Co., 1974.
- Griffin, G.A. Guidelines for the evaluation of staff development programs. Teachers College Record, 1978, 80 (1), 126-139.
- Griffin, G.A. Staff development. Paper prepared for the NIE Teaching Synthesis Conference. Arlie House, Virginia, 1982.

Griffin, G.A., & Lieberman, A. Behavior of innovative personnel. Washington, DC: ERIC Clearinghouse on Teacher Education, 1974.

Haines, P.G. Career educational inservice teacher education. Columbus, Ohio: ERIC Clearinghouse on Vocational and Technical Education, 1973.

Hall, G. The concerns-based approach to facilitating change. Educational Horizons, 1979, 57 (4), 202-208.

Hall, G., Loucks, S.F., Rutherford, W.L., & Newlove, B.W. Levels of use of the innovation: A framework for analyzing innovation adoption. Journal of Teacher Education, 1975, 26 (1), 52-57.

Harris, B.M. Improving Staff Performance Through In-Service Education. Boston: Allyn & Bacon, Inc., 1980.

Harris, B.M., & Bessent, W. Inservice Education: A Guide to Better Practice. Englewood Cliffs, New Jersey: Prentice-Hall, 1969.

Havelock, R.G. The Change Agent's Guide to Innovation in Education. Englewood Cliffs, New Jersey: Educational Technology Publications, 1973.

Havelock, R.G. & Havelock, M.C. Training for Change Agents: A Guide to the Design of Training Programs in Education and Other Fields. Ann Arbor,

Michigan: University of Michigan Institute for Social Research, 1973.

Heckman, P., Oakes, J., & Sirotnik, K. Expanding the concepts of school renewal and change. Education-al Leadership, 1983, 40 (7), 26-32.

Hite, H., & Howey, K.R. Planning inservice teacher education: Promising alternatives. Washington, DC: The American Association of Colleges for Teacher Education, 1977.

Hockman, E. Staff development data banks. Taking evaluation beyond "happiness quotients." The Journal of Staff Development, 1982, 3 (1), 71-83.

Howey, K.R. A study of the perceptions of teachers, professors, and parents about current and projected practices. American Educational Research Association 1978, ERIC: ED 152 701.

Huling, L. The effects on teachers of participation in an interactive research and development project. Doctoral dissertation, Texas Tech University, 1981.

Jackson, G.B. Methods for reviewing and integrating research in the social sciences. Final Report to the National Science Foundation for Grant No. DIS 76-20309. Washington, DC: Social Research Group, George Washington University, April, 1978.

- Joslin, P.A. Inservice teacher education: A meta-analysis of the research. Doctoral dissertation, University of Minnesota, 1980.
- Joyce, B.R., Howey, K.R., Yarger, S.J., Hill, W.C., Waterman, F.T., Vance, B.A., Parker, D.W., & Baker, M.G. Issues to face. (ISTE Report 1). Palo Alto, California: National Center for Educational Statistics and the Teacher Corps, 1976.
- Joyce, B.R., McNair, K.M., Diaz, R., McKibbin, M.D., Waterman, F.T., & Baker, M.G. Interviews (ISTE Report 2). Palo Alto, California: National Center for Education Statistics and the Teacher Corps, 1976.
- Joyce, B.R., & Showers, B. Improving inservice training: The messages of research. Educational Leadership, 1980, 37 (5), 379-385.
- Joyce, B.R., & Showers, B. Transfer of training: The contribution of "coaching". Journal of Education, 1981, 163 (2), 163-172.
- Kinnick, B.J., Boney, V., Huntington, E.A., Rasmussen, R.T., Simon, S.A., Tacina, N.G., & Zimmerman, E. The teachers and the inservice education program. In Inservice Education for Teachers, Supervisors, and Administrators. The Fifty-Sixth Yearbook of the National Society for the Study of Education,

- N.B. Henry (ed.). Chicago: University of Chicago Press, 1957.
- Kirkpatrick, D.L. Techniques for evaluating training programs. Journal of the American Society of Training Directors, 1959, 13, 3-9.
- Kirkpatrick, D.L. Evaluation of training. In Training and Development Handbook, R.L. Craig & L.R. Bittel (eds.). New York: McGraw-Hill, 1967.
- Knight, E.W. Public Education in the South. Boston: Ginn, 1922.
- Lawrence, G. Patterns of effective inservice education. Report prepared for the state of Florida, Department of Education, December, 1974.
- Lawrence, G., & Harrison, D. Policy implications of the research on the professional development of education personnel: An analysis of fifty-nine studies. In 1981 Report on Educational Personnel Development. Washington, DC: Feistritz Publications, 1980.
- Lieberman, A. Staff development: New demands, new realities, new perspectives. Teachers College Record, 1978, 80 (1), 1-3.
- Lippitt, R. & Fox, R. Development and maintenance of effective classroom learning. In Improving Inservice Education, L.J. Rubin (ed.). Boston: Allyn &

Bacon, Inc., 1971.

Little, J.W. School success and staff development:

The role of staff development in urban desegregated schools. Paper adapted from final report to the National Institute of Education, 1981.

Lortie, D.C. School Teacher: A Sociological Study.

Chicago: The University of Chicago Press, 1975.

Loucks, S.F., & Melle, M. Evaluation of staff develop-

ment: How do you know it took? The Journal of Staff Development, 1982, 3, (1), 102-117.

Mann, D. The politics of training teachers in schools.

Teachers College Record, 1976, 77 (3), 323-338.

McLaughlin, M.W., & Marsh, D.D. Staff development and

school change. Teachers College Record, 1978, 80 (1), 69-94.

Miles, M. School innovation from the ground up: Some

dilemmas. New York University Education Quarterly, 1980, 11 (2), 2-9.

Mohr, P. Current research and development efforts in

inservice training and curriculum planning for teacher education. Washington, DC: Department of Health, Education, and Welfare, 1971, ERIC: ED 083 148.

Newstrom, J.W. Catch 22: The problems of incomplete

training. The Training and Development Journal,

1978, 32 (11), 22-24.

Nicholson, A.M., Joyce, B.R., Parker, D.W., & Waterman F.T. The literature on inservice teacher education (ISTE Report 3). Palo Alto, California: National Center for Educational Statistics and the Teacher Corps, 1976.

Nie, N., Hull, C.H., Jenkins, J.G., Steinbrenner, K., & Bent, D.H. Statistical Package for the Social Sciences, Second Edition. New York: McGraw-Hill, 1975.

Orrange, P., & VanRyn, M. Agency roles and responsibilities. In Rethinking In-Service Education, R.A. Edelfelt & M. Johnson (eds.). Washington, DC: National Education Association, 1975.

Peeler, T., & Shapiro, J. A focus of the cooperative inservice teacher education programs. Washington, DC: 1974, ERIC: ED 081 764.

Perloff, E., Kohne, C.F., & Gatewood, R.D. Project impact: A pilot study evaluating the NDEA summer institute program. Pittsburgh, Pennsylvania: American Institute for Research (for COMPASS), 1970.

Perritt, M.C. Professional growth in the Shelby County schools. Theory Into Practice, 1972, 11 (5), 324-328.

- Rapport, R., & Rapport, R.N. Leisure and the Family Life Cycle. London: Routledge and Kegan Paul, 1975.
- Richey, H.G. Growth of the modern conception of inservice education. In Inservice Education for Teachers, Supervisors, and Administrators. The Fifty-Sixth Yearbook of the National Society for the Study of Education, N.B. Henry (ed.). Chicago: University of Chicago Press, 1957.
- Rosenthal, R., & Rosnow, R.L. The Volunteer Subject. New York: Wiley-Interscience, 1975.
- Rubin, L.J. Teacher growth in perspective. In Improving Inservice Education, L.J. Rubin (ed.). Boston: Allyn & Bacon, Inc., 1971.
- Sarason, S.B. The Culture of the School and the Problem of Change. Boston: Allyn & Bacon, Inc., 1974.
- Schmeider, A. Teacher power and local-to-local delivery. Journal of Teacher Education, 1972, 23 (2), 115-116, 125, 144.
- Schmeider, A., & Yarger, S. Teaching centers: Toward the state of the scene. Washington, DC: ERIC Clearinghouse on Teacher Education, 1974.
- Sobol, F.T. What variables appear important in changing traditional inservice training procedures. Wash-

- ington, DC: 1971, ERIC: ED 083 146.
- Spicer, E. Human Problems in Technological Change.
New York: Russell Sage Foundation, 1952.
- Stufflebeam, D.L. Evaluation as enlightenment for
decision-making. Ohio State University Evaluation
Center (mimeo), 1969.
- Tikunoff, W., & Griffin, G. The IR&DT experience.
Paper presented at the Annual Meeting of the Amer-
ican Educational Research Association, Boston,
April, 1980.
- Tikunoff, W., & Ward, B. Interactive Research on
Teaching: An Overview of the Strategy. San Fran-
cisco: Far West Laboratory, 1979.
- Tikunoff, W., Ward, B., & Griffin, G. Interactive
research and development on teaching: Final re-
port. San Francisco: Far West Laboratory for
Educational Research and Development, 1979.
- Tough, A. Learning Without a Teacher. Toronto,
Canada: Ontario Institute for Studies in Educa-
tion, 1967.
- Tyler, R.E. Inservice education of teachers: A look at
the past and future. In Improving In-Service
Education: Proposals and Procedures for Change,
L.J. Rubin (ed.). Boston: Allyn & Bacon, Inc.,
1971.

- Watson, G. (ed.). Change in Social Systems. Washington, DC: NTL Institute for Applied Behavioral Science, National Education Association, 1967.
- Watts, .E., & Hammons, J.D. Staff development: A time for appraisal. Washington, DC:, 1981, ERIC: ED 203 928.
- Waynant, L.F. Teachers' strengths: Basis for successful in-service experiences. Educational leadership, 1971, 28 (7), 710-713.
- Wolf, W.C., Jr. The Wolf knowledge diffusion/utilization inventory. Unpublished manuscript, University of Massachusetts, 1981.
- Wood, F.H., & Thompson, S.R. Guidelines for better staff development. Educational Leadership, 1980, 37 (5), 374-378.

APPENDIX A
STUDIES INCLUDED IN THE META-ANALYSIS

- Agrawal, P.C. An analysis of individual and team effects on prescriptive and catalytic modes of staff development with specific reference to metric education. Doctoral dissertation, State University of New York at Albany, 1979.
- Allen, B.B. An in-service education model and effects of implementation on teacher perception of students. Doctoral dissertation, Brigham Young University, 1974.
- Alloway, E., Carroll, J., Emig, J., King, B., Marcotrigiano, I., Smith, J., & Spicer, W. The New Jersey Writing Project. New Brunswick, New Jersey: Rutgers University, 1980, ERIC: ED 178 943.
- Anderson, L.A., Evertson, C.A., & Brophy, J.E. An experimental study of effective teaching in first-grade reading groups. Elementary School Journal, 1979, 79 (4), 193-224.
- Askov, E.N., & Dupuis, M.M. Teacher change: Impact of an inservice program in content area reading. Paper presented at the Second Annual Meeting of the American Reading Forum, Sarasota, Florida, 1981. ERIC: ED 211 966.

- Bacon, M.A. An investigation of the effectiveness of a secondary reading inservice program on teachers' knowledge, attitudes, and teaching behavior. Doctoral dissertation, University of Massachusetts, 1978.
- Barclay, J.R. Changing the behavior of school psychologists: A training rationale and method. Hayward, California: California State College, 1968. ERIC: ED 024 993.
- Bennett, M.A. Effects of participation in the Taba Inservice Education Program on teachers' self-concept, attitude, and selected personality characteristics. Doctoral dissertation, North Texas State University, 1971.
- Bethel, L.J., & Hord, S.M. A case study of change: Inservice teachers in a National Science Foundation environmental science education program. Washington, DC: National Science Foundation, 1981. ERIC: 200 589.
- Bogut, T.L., & McFarland, B.L. The effects of instruction on elementary teachers' realistic and idealistic attitudes toward selected science related concepts. Paper presented at the National Convention of the National Science Teachers Association, Los Angeles, 1975. ERIC: ED 106 125.
- Bond, P.Y. The effects of feedback on teacher verbal behavior and attitudes toward inservice education. Doc-

toral dissertation, North Texas State University, 1969.

Borg, W.R. Protocol materials as related to teacher performance and pupil achievement. Journal of Educational Research, 1976, 69 (1), 23-30.

Borg, W.R. Changing teacher and pupil performance with protocols. Journal of Experimental Education, 1977, 45 (3), 9-18.

Breit, F., & Butts, P. A comparison of the effectiveness of an inservice program and a preservice program in developing certain teaching competencies. Paper presented at the National Association for Research in Science Teaching meeting, Pasadena, California, 1969. ERIC: ED 028 069.

Brown, J.A., & Kameen, M.C. Changing teacher competence perceptions by focused video-tape feedback. Paper presented at the Annual Meeting of the American Personnel and Guidance Association, New York, 1975. ERIC: ED 110 896.

Brown, J.A. MacDougall, M.A., & Jenkins, C.A. Changing culture perceptions of elementary school teachers. ED occasional paper No. 9. Charlottesville, Virginia: University of Virginia, School of Education, 1972. ERIC: ED 066 537.

Brown, W.J. The effects of in-service education and re-

source unit components on teacher and student learning: Research series on occupational education, No. 12. Raleigh, North Carolina: North Carolina State University, School of Education, 1968. ERIC: ED 042 885.

Cameron, W.A., & Cotrell, C.J. Remote feedback techniques for inservice education. Assessment of micro-teaching and video recording in vocational and technical teacher education: Phase X--interim report. Research and development series No. 40. Ft. Collins, Colorado: Colorado State University, Department of Vocational Education; Columbus, Ohio: Ohio State University, Center for Vocational and Technical Education, 1970.
ERIC: ED 042 901

Cantrell, P. Efficacy of in-service training of teachers in operant techniques. Doctoral dissertation, George Peabody College for Teachers, 1969.

Carline, J.L. In-service training--re-examined. Journal of Research and Development in Education, 1970, 4 (1), 103-115.

Chism, M.J.E. A study of the effect of an in-service education program on the questioning techniques employed by fifth grade and sixth grade teachers of the Memphis city schools. Doctoral dissertation, Memphis State University, 1972.

- Chow, S.H.L., Rice, C.F., & Whitmore, R.A. Effects of a mediated training course on teachers and students in mainstreaming programs. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, 1976. ERIC: ED 123 822.
- Cleveland Public Schools. Project study (curriculum for improving study skills): Project termination report. Cleveland, Ohio: Cleveland Public Schools, 1975. ERIC: ED 128 937.
- Clock, C.J., Jr. Improving the instructional use of test data through an in-service staff development program for educators. Doctoral dissertation, University of Massachusetts, 1978.
- Codwell, J.E. A demonstration of the effect of an adaptation of microteaching on the instructional behavior of rural school teachers. Final report. Atlanta, Georgia: Southern Association of Colleges and Schools, 1969. ERIC: ED 034 620.
- Constantine, Sister M., S.S.J. A curricular study concerned with the process and the product of an English-education course and its effects upon experienced teachers' abilities to think critically: Interim report. Urbana, Illinois: Illinois State-Wide Curriculum Study Center, 1968. ERIC: ED 019 288.
- Dell, H. The effects of an individualized instruction

workshop and its related follow-up program on the attitudes and behavior of selected elementary teachers and their students: Final report. San Jose, California: California State University, School of Education, 1972. ERIC: ED 073 076.

Ebmeier, H., & Good, T.L. The effects of instructing teachers about good teaching on the mathematics achievement of fourth-grade students. American Educational Research Journal, 1979, 16 (1), 1-16.

Eugene Public Schools, Oregon. Elementary career awareness project: Final report. Salem, Oregon: Oregon State Board of Education, Division of Community Colleges and Career Education, 1974. ERIC: ED 111 930.

Esposito, J. School-based staff development: Collaborative planning and preparation for change. Doctoral dissertation, University of Massachusetts, 1981.

Felđt, J.R. von, & Subtelny, J. Description and evaluation of the Webster's diacritical markings. Computer-assisted instructional program: Summary report. Rochester, New York: National Technical Institute for the Deaf, 1975. ERIC: ED 153 642.

Firth, J.L. An investigation of the effectiveness of the Bessell-Palomares inservice training program on teacher attitudes. Doctoral dissertation, Arizona State University, 1974.

- Fitzmaurice, M.D. Learning institute in-service results. Paper presented at the annual meeting of the International Reading Association, Anaheim, California, 1976. ERIC: ED 127 554
- Goldstein, H. Mischio, G.S., & Minskoff, E. A demonstration-research project in curriculum and methods of instruction for elementary level mentally retarded children: Final report. New York: Yeshiva University, 1969. ERIC: ED 058 696.
- Good, T.I., & Grouws, D.A. The Missouri mathematics effectiveness project: An experimental study in fourth grade classrooms. Journal of Educational Psychology, 1979, 71 (3), 355-362.
- Greene, J.F., Archambault, F., & Nolen, W. The effect of extended inservice training curricula upon the mathematics achievement and attitudes of elementary teachers. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, 1976. ERIC: ED 123 091.
- Guines, J.T. Improving instruction through supervision. Washington, DC: District of Columbia Public Schools, 1981. ERIC: ED 201 707.
- Guskey, T.R. The effects of change in instructional effectiveness on the relationship of teacher expectations and student achievement. Journal of Educational Re-

search, 1982, 75 (6), 345-349.

- Hall, K.A., Mitzel, H.E., Riedesel, C.A., Suydam, M.N., Trueblood, C.T., & Mechlin, S.W. Inservice mathematics education for elementary school teachers via computer-assisted instruction: Interim report. University Park, Pennsylvania: Pennsylvania State University, Computer Assisted Instruction Laboratory, 1969. ERIC: ED 076 042
- Hawkins, J., & Frechtling, J. Short-term effects of human relations training (HR-17): A pre-post comparison group evaluation study. Rockville, Maryland: Montgomery County Public Schools, 1981. ERIC: ED 200 664.
- Hawkins, J., Frechtling, J., & Brown, D. Short-term effects of human relations training (HR-18): A pre-post evaluation study. Rockville, Maryland: Montgomery County Public Schools, 1981. ERIC: ED 200 665.
- Honigman, F.K. An evaluation of the program "Supportive Training for Inexperienced and New Teachers" (STINT) in New York City schools. New York: Institute for Educational Development, 1970. ERIC: ED 051 087.
- Howell, M.M. The impact of an inservice education program on the attitude and behavior of certain teachers of the English language in Pinellas County, Florida. Doctoral dissertation, Florida State University, 1972.
- Huling, L.L. The effects on teachers of participation in

- an interactive research and development project. Doctoral dissertation, Texas Tech University, 1981.
- Hunkler, R. An evaluation of a short-term inservice mathematics program for elementary school teachers. School Science and Mathematics, 1971, 71 (1), 650-654.
- Jacobs, J. Insight learning through structured observations of classroom interaction. Journal of Research in Science Teaching, 1973, 10 (3), 213-220.
- Jaus, H.H. Using microteaching to change elementary teachers' attitudes toward science education. School Science and Mathematics, 1977, 77, 402-406.
- Kasdon, L.M., & Kelly, D. Simulation inservice education for teachers of reading. Journal of Experimental Education, 1969, 38 (1), 79-86.
- Khanna, J.L. Human relations training for educators. Final evaluation. Livingston, Tennessee: Project Upper Cumberland, 1970. ERIC: ED 043 057.
- Koran, M.L., Snow, R.E., & McDonald, F.J. Teacher aptitude and observational learning of a teaching skill. Journal of Educational Psychology, 62 (3), 219-228.
- Langer, P., & Allen, G. The minicourse as a tool for training teachers in interaction analysis. Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, 1970. ERIC:

ED 037 393.

- Lapp, B. The effects of behavior modification inservice training on teacher behavior, teacher attitudes, and knowledge of behavior modification. Doctoral dissertation, University of Connecticut, 1974.
- Lavach, J.F. Organization and evaluation of an inservice program in the history of science. Journal of Research in Science Teaching, 1969, 6 (2), 166-170.
- Lee, J.W. Effects of inservice training in reading on the attitudes and skills of secondary teachers and adult basic education teachers. Doctoral dissertation, Pennsylvania State University, 1977.
- Lee, W.S. Human relations training for teachers: The effectiveness of sensitivity training. California Journal of Educational Research, 1970, 21 (1), 28-34.
- Leonard, C.B., & Ciaglia, E.R. The effects of the instructional behavior and skills development program upon teaching behavior. Columbia, Missouri: University of Missouri, 1972. ERIC: ED 070 765.
- MacDougall, M., & Brown, J. The relationship of videotaping experiences to classroom interaction behaviors and peer acceptance among elementary pupils. Charlottesville, Virginia: University of Virginia, 1974. ERIC: ED 090 242.
- Marble, W.O., Jr. An evaluation of the professional devel-

- opment sequence on "The Slow Learner in the Secondary School". British Columbia, Canada: Burnaby School District, 1980. ERIC: ED 209 274.
- Martin, F. The effects of a creative problem-solving workshop upon the cognitive operations of verbal classroom interaction in the primary school grades. Psychology in the Schools, 1972, 9 (2), 126-130.
- Martinson, R.A., & Wiener, J. The improvement of teaching procedures with gifted elementary and secondary school students: Final report. Gardena, California: California State College, 1968. ERIC: ED 024 185.
- McKenzie, F.D. Improving instruction through supervision: Final report. Washington, DC: District of Columbia Public Schools, 1981. ERIC: ED 213 799.
- Meehan, M.L. Evaluation of the Stallings classroom management staff development demonstration project in Putnam County. Charleston, West Virginia: Appalachia Educational Laboratory, 1981. ERIC: ED 225 977.
- Merwin, W.C. Competency-based modules for in-service education. Educational Leadership, 1974, 31 (4), 329-332.
- Miller, J.P. The effects of inservice human relations training on teacher interpersonal functioning. Doctoral dissertation, University of Toronto, Canada, 1971.

- Moore, J.W., & Schaut, J.A. Stability of teaching behavior, responsiveness to training and teaching effectiveness. Journal of Educational Research, 1975, 69 (10), 360-363.
- Moore, J.W., & Schaut, J.A. Increasing instructional effectiveness through the use of a problem-solving approach to the design of instructional systems. Journal of Experimental Education, 1978, 47 (2), 156-161.
- Murray, S.L., Rassen, R.L., & Speedie, S.M. Research using problem solving: Outcome evaluation report. Improving teaching competencies program. Portland, Oregon: Northwest Regional Educational Laboratory, 1976.
ERIC: ED 121 750.
- New Hampshire University. In-service teacher education courses in art and science for the elementary teachers of New Hampshire: An evaluation report. Durham, New Hampshire: Bureau of Educational Research & Testing Services, 1968. ERIC: ED 024 645.
- Orlich, D.C., May, F.B., Harder, R.J., Ruff, T.D., & Ormsby, J. A change agent strategy: Preliminary report. The Elementary School Journal, 1972, 72 (6), 281-293.
- Ost, D.H. An evaluation of an institute for teachers of secondary school biology. American Biology Teacher, 1971, 33 (9), 546-548.

- Overline, H.M. An investigation into the effects of an inservice program in learning disabilities on elementary school teachers' attitudes and classroom instructional activities. Doctoral dissertation, Temple University, 1972.
- Porter, P.C. Systemic human relations training: The effect on levels of concern and interpersonal skills of inservice teachers. Doctoral dissertation, Brigham Young University, 1977.
- Sacco, J.M., & Cook, J. Learning hierarchies improve teachers attitudes toward accountability. Journal of Teacher Education, 1976, 27 (2), 113-115.
- Schmid, R., & Scranton, T.R. A field trial of the longitudinal inservice training model. Education, 1972, 93 (2), 195-198.
- Shaffer, W.F. , Hummel, T.J., & Rhetts, J.E. Design and evaluation of a human relations training program: Research memorandum No. 100. Minneapolis, Minnesota: University of Minnesota, Psycho-Educational Center, 1974. ERIC: ED 095 451.
- Spradlin, S.D. The evaluation of a summer institute program: Has it reached its goal? Paper presented at the annual meeting of the National Association for Research in Science Teaching, Detroit, Michigan, 1973. ERIC: ED 094 950.

- Stallings, J., Needels, M., & Stayrook, N. How to change the process of teaching basic reading skills in secondary schools. Menlo Park, California: SRI International, 1980. ERIC: ED 210 671.
- Stevens, L.P. An experiment to determine the effects of an IOTA inservice training program upon teacher-pupil verbal interaction. Doctoral dissertation, Arizona State University, 1969.
- Stone, D.E. An experimental study of in-service teacher training to promote inductive teaching and creative problem solving. Paper presented at the annual meeting of the American Educational Research Association, 1969. ERIC: ED 029 819.
- Syropoulos, M. Evaluation of the Federally Assisted Staff Training (FAST) Project. Detroit, Michigan: Detroit Public Schools, 1972. ERIC: ED 077 849.
- Tamminen, A.W. An institute for the training of participants in the ES '70 projects. Duluth, Minnesota: University of Minnesota, 1970. ERIC: ED 042 738.
- Thelen, .J., & Litsky, W. Teacher attendance at a summer institute and high school student achievement. Science Education, 1972, 56 (3), 293-303.
- Vallejo, M.E. The relationship between the effects of a career education teacher inservice preparation program and the career awareness of upper elementary school

- students. Doctoral dissertation, Kansas State University, 1975.
- Walker, L., Paddock, H., Brown, L., & Baksh, I. Nonstandard dialect and literacy: An in-service project in Newfoundland. Interchange, 1975, 6 (3), 4-10.
- Webster, W.J., & Mendro, R.L. The effects of a video tape feedback system on classroom teacher behavior. Journal of School Psychology, 1974, 12 (3), 189-198.
- Werner, E.R., Golden, Y., & Mills, S.R. The minicourse as a method for training teachers to stimulate divergent thinking. Paper presented at the annual meeting of the American Educational Research Association, Chicago, 1972. ERIC: ED 064 340.
- Whiteman, P.L. The effect of participation in an inservice training institute in diagnostic-prescriptive methodology upon teachers' knowledge of reading and the reading process. Doctoral dissertation, Indiana University, 1977.
- Williams, H.L. The effects of two spelling approaches among sixth grade students and their participating teachers: A study of spelling achievement and teachers' attitude toward staff development. Doctoral dissertation, University of Houston, 1979.
- Wood, E., Jr. The effect of an in-service training program in verbal interaction analysis on teacher behavior in

the classroom. Doctoral dissertation, University of Houston, 1968.

Yates, J.R. Model for preparing regular classroom teachers for "mainstreaming" exceptional children. Exceptional Children, 1973, 39 (6), 471-472.

APPENDIX B
STUDY IDENTIFICATION CODE

<u>ID#</u>	<u>Data Set(s)</u>	<u>Author(s)</u>
001	01-48	Cleveland Public Schools
002	01-04	Codwell, J.E.
003	01-25	Chow, S.H.L.
004	01	Brown, J.A. et al.
005	01-06	Bogut, T.L. et al.
006	01-03	Brown, J.A. et al.
007	01-02	Brown, W.J.
008	01-03	Cameron, W.A. et al.
009	01-06	Yates, J.R.
010	01-03	Webster, W.J. et al.
011	01-02	Anderson, L.A. et al.
012	01-05	Martin, F.
013	01-02	Kasdon, L.M. et al.
014	01-04	Ost, D.H.
015	01	Schmid, R. et al.
016	01	Sacco, J.M. et al.
017	01-06	Orlich, D.C. et al.
018	01	Moore, J.W. et al.
019	01-02	Moore, J.W. et al.
020	01-02	Merwin, W.C.
021	01-06	Lee, W.S.

022	01-02	Lavach, J.F.
023	01-03	Hunkler, R.
024	01-14	Carline, J.L.
025	01-12	Borg, W.R.
026	01-28	Borg, W.R.
027	01-20	Jacobs, J.
028	01-07	Ebmeier, H. et al.
029	01	Jaus, H.H.
030	01	Good, T.I. et al.
031	01-28	Koran, M.L. et al.
032	01-10	Huling, L.L.
033	01	Thelen, L.J. et al.
034	01	Feldt, J.R. von. et al.
035	01-02	Fitzmaurice, M.D.
036	01-31	Goldstein, H. et al.
037	01-08	Khanna, J.L.
038	01-09	Honigman, F.K.
039	01-02	Hall, K.A. et al.
040	01-02	Hall, K.A. et al.
041	01-04	Greene, J.F. et al.
042	01-13	Langer, P. et al.
043	01	Shaffer, W.F. et al.
044	01-02	New Hampshire University
045	01-04	Murray, S.L. et al.
046	01-06	Chism, M.J.E.

047	01-04	Martinson, R.A. et al.
048	01-18	MacDougall, M. et al.
049	01-09	Leonard, C.B. et al.
050	01	Marble, W.O. Jr.
051	01-27	Werner, E.R. et al.
052	01-03	Tamminen, A.W.
053	01-04	Syropoulos, M.
054	01-03	Stone, D.E.
055	01-06	Spradlin, S.D.
056	01-27	Barclay, J.R.
057	01-46	Breit, F. et al.
058	01-02	McKenzie, F.D.
059	01-05	Askov, E.N. et al.
060	01	Stallings, J. et al.
061	01-29	Guines, J.T.
062	01-06	Guskey, T.R.
063	01-04	Porter, P.C.
064	01-03	Lapp, B.
065	01-03	Cantrell, R.P.
066	01	Alloway, E. et al.
067	01	Bethel, L.J. et al.
068	01-10	Hawkins, J. et al.
069	01-04	Hawkins, J. et al.
070	01-04	Bacon, M.A.
071	01-17	Esposito, J.

072	01-06	Clock, C.J. Jr.
073	01-02	Constantine, Sister, S.S.J.
074	01-07	Eugene Public Schools
075	01-03	Dell, H.
076	01-06	Agrawal, P.C.
077	01-03	Bond, P.Y.
078	01-03	Lee, J.W.
079	01-14	Overline, H.M.
080	01-03	Vallejo, M.E.
081	01-04	Wood, N.E. Jr.
082	01-09	Meehan, M.L.
083	01	Howell, M.M.
084	01-03	Allen, B.B.
085	01-03	Whiteman, P.L.
086	01-06	Miller, J.P.
087	01-06	Bennett, M.A.
088	01-20	Stevens, L.P.
089	01-09	Firth, J.L.
090	01-04	Walker, L. et al.
091	01-20	Williams, H.L.

APPENDIX C
OPERATIONAL DEFINITIONS FOR DEPENDENT
AND INDEPENDENT VARIABLES

Dependent Variables

1. Reaction: Measure of how the participants feel about the staff development activities, usually subjective.
2. Learning: Objective and quantitative measures that assess how much a participant has learned as a result of inservice activities.
3. Behavior: Objective measures that document whether or not participants change their behavior as a result of a staff development intervention.
4. Results: Objectively determining the effects of staff development on students of participating teachers or on the working environment.
5. Effect Size: The mean difference between treated and control groups divided by the standard deviation of the control group or some approximation of this measurement.

Independent Variables

1. Instructional Focus
Improvement of specific subject: Pertains to any at-

tempt to upgrade teacher or student achievement in a particular subject area.

Improvement of general teaching: Pertains to any attempt to generally upgrade teacher or student achievement across subject areas.

Improvement of affective techniques: Pertains to any attempt to improve the social-emotional or human relations aspects of education..

2. Source

Journal: Refers to all studies obtained from journal articles.

ERIC: Refers to all studies obtained from the ERIC Clearinghouse on Teacher Education, Washington, DC.

Dissertation: Refers to all studies obtained from dissertations which were available locally or through interlibrary loan.

3. Number of Participants

Refers to the number of people used to calculate the effect size, including the control group. The possibilities include:

1-20

21-40

41-60

greater than 60

4. Location

On-site: Indicates that the majority of the training took place at the school.

Off-site: Indicates that the majority of the training took place at a site other than the school.

5. Initiator

Refers to the individual or group whose idea it was to pursue a training program. The possibilities include:

Participants: Those individuals who take part in the training program.

Administrator/supervisor: Those individuals who hold managerial positions.

School: Refers to a particular school site that initiates training. (This category was employed only in cases where a specific initiator was not mentioned, but a school was named as the initiator of training.)

State: Refers to person(s) and programs initiated at the state level (i.e., state departments of education, state agencies, etc.).

Federal government: Refers to programs initiated at the national level (i.e., federal agencies, national programs supported by the federal government, etc.).

University researcher: Refers to person(s) at the college level who recruit participants for training in order to pursue their own research, ensure an adequate class enrollment, etc.

Consultant: Refers to person(s) with professional expertise in some area of education who recruit participants for training. Only those individuals who work outside of the environment targeted for intervention were considered consultants. Generally, these individuals entitled themselves, "professional consultants."

Other: Refers to training initiated by any individual or group who do not belong in any of the seven previously mentioned categories.

6. Grade Level Taught

Refers to the particular grade level at which participants in a training program work. The possibilities include:

Elementary: Grades kindergarten through sixth.

Secondary: Grades seven through twelve.

Both: Elementary and secondary staff both are represented.

7. Participation

Mandatory: Participants were required to attend training.

Voluntary: Participants volunteered to attend training.

Both: Some participants volunteer to attend, while others are required to attend.

8. Participant Incentives

Pay: Participants receive financial remuneration for participating in the training.

Certificate renewal: Participants receive credit toward renewing their teaching certificate by participating in the training.

Released Time: Participants are released from their teaching duties in order to participate in the training.

Status/recognition: Participants are selectively chosen to take part in the training program or in some way given special recognition as a participant.

College credit: Participants receive college credit for participating in the training.

None: Participants receive none of the five previously mentioned rewards for their participation. In fact, no reward is mentioned as an incentive for participation in the training.

9. Funding

Refers to the individual or group who pay for the inservice training. The possibilities include:

Participant

School

Federal government

State government

University

Other

Combination10. Instructor

Refers to the job category of the person leading the inservice training.

Teacher: A person whose major job responsibility is teaching one or more of the grade levels K-12.

Supervisory staff: A person whose major job responsibility is supervisory work in a particular subject area at one of the grade levels K-12 (i.e., mathematics supervisor, language arts coordinator, department head, etc.).

Administrator: A person whose major job responsibility is general administration of a cluster of grade levels at the K-12 level (i.e., principal, assistant superintendent, instructional director, etc.)

Outside consultant: A person whose major job responsibility is educational consulting.

College personnel: A person whose major job responsibility is teaching/research at a college or university.

Self: Refers to instruction that is self-administered.

State: Refers to instruction that is lead by state department officials.

Other: Refers to instruction that is lead by someone or group other than the seven previously mentioned categories.

11. Types of Instruction

This category includes any one of the following subtopics or a combination thereof:

Lecture: Didactic instruction.

Games/simulation: Instruction in the form of games that approximate real-life situations.

Discussion: Instruction that focuses on examining a topic by argument.

Modeling: Instruction by presentation of a certain standard and/or method to be imitated by participating teachers.

Coaching: Instruction that focuses on giving advice to a teacher who is attempting to implement new learning in the classroom or school environment.

Video/Audio Feedback: Instruction that focuses on providing video and/or audio feedback to a teacher who is practicing the implementation of new learning.

Mutual Assistance: Instruction that encourages and allows for any form of reciprocal assistance among class members.

Microteaching: Instruction that focuses on analyzing and practicing particular skills related to teaching.

Practice: Instruction that provides an opportunity to practice the particular skill that is being taught.

Printed Material: Instruction that uses printed mater-

ial as the means for imparting new information.

Programmed Study: A method of instruction that presents a program of questions and answers that are used in self-study.

Observation: Instruction that provides an opportunity to watch a particular skill or skills being used in a natural environment, in this case, the classroom or school.

Production of Instructional Materials: a method of instruction that focuses on making your own materials to be used at a later time in the instruction of children.

Guided Field Trips: Instruction through investigation taking place outside the regular instructional setting.

Film: Instruction provided through film.

12. Structure

Workshop: A form of instruction focused on providing concrete learning experiences, usually of short-term duration.

Course: A form of instruction consisting of a systematized series of lectures and study, usually of longer-term duration.

Staff Meeting: A form of instruction taking place during a staff meeting, usually lasting for only one or two sessions.

Mini-Course: A form of instruction consisting of a systematized series of lectures and study, usually lasting from four to eight sessions.

Independent Study: A form of instruction consisting of self-study, often under the supervision of another individual.

Institute: A form of study, usually organized at the state or national level, providing an intensive course of study in technical subjects.

One-to-One: A form of instruction consisting of one person receiving individual instruction from another person.

Other: Refers to instruction that takes some form other than the seven previously mentioned categories.

13. Length of Treatment

Refers to the number of hours of instruction in any given inservice treatment. The possibilities include:

1-10 hours

11-20 hours

21-30 hours

greater than 30 hours

14. Emphasis

Theory: Instruction primarily focused on principles or methods rather than on actual practice.

Practical: Instruction primarily focused on actual

practice rather than on theory.

Combination: Instruction providing a fairly even emphasis on both theory and practice.

15. Schedule:

Refers to the time period during which most of the instruction takes place. The possibilities include:

Before/after work

Evenings

Saturdays

Summer

During the school day

Combination (of any of the other seven categories)

Weekends

Other

16. Duration

Long-term: Training lasting more than six months.

Short-term: Training lasting six months or less.

17. Activities:

Individualized: Activities are tailored to meet individual needs, therefore not all participants take part in the same activities.

Common: Activities are designed for the entire class based upon group needs or predetermined needs rather than the needs of any particular individual.

18. Participant Role

Active: Participants are encouraged to exert influence over the direction of the learning activities.

Passive: Participants do not seem to be encouraged to exert influence over the direction of the learning activities.

19. Goals

Shared/Collective Goals: Participants have the same end in mind as a result of taking part in the inservice training.

Personal Goal: Participants have quite different ends in mind, based upon their individual needs.

20. Group Focus

Faculty as a Unit: Instruction is focused on a particular group of teachers within a particular school building or school district. Participants are colleagues.

Individuals: Instruction is focused on a group of individuals from a variety of teaching settings. Participants often do not know one another before the training begins.

21. Status

Participant and Leader Equal: The instructional style employed suggests that participants have information to offer just as the instructor does. Participants are encouraged to give presentations, teach each other,

present opposing views, etc.

Superordinate/Subordinate: The instructional style employed suggests that the instructor is the expert and the major imparter of knowledge. Participants are not encouraged to take control of any part of the instruction.

22. Follow-Up

Assistance Later: The instructor or an aide provide assistance in some way after the actual training sequence is complete. This can take the form of an on-site visit, a phone call, or other individualized contact.

No Assistance Later: The instructor provides no assistance after the actual training sequence is complete.

23. Assignment to Groups

Random: Inferences are based upon a random sample. Each member of a given population has an equal probability of being included in a given group.

Matching: Control and experimental groups are matched on certain characteristics such as grade level taught, sex, etc.

Covariance Adjustment: Control and experimental groups are weighted in order to make correlations between pre-test groups as high as possible.

Equating Pre-test: Pre-test scores are weighted in

order to guarantee that gain scores on post-tests can be accurately compared.

Repeated Measure: The same subjects are measured under a number of different conditions.

Non-equivalent Groups: Control and experimental groups are not systematically equated and are not equal.

No Control Group: The only subjects consist of those in the experimental group.

24. Internal Validity

High: There is a less than fifteen percent mortality rate, and subjects are randomly assigned to groups.

Moderate: here is a greater than fifteen percent mortality rate or subjects are well matched.

Low: Any design not fitting the above two categories.

25. Reactivity Level

Under Control of Experimenter: Any measurement easily influenced by the experimenter (i.e., a behavioral observation done by the experimenter and then used to determine the success of an intervention).

Medium Experimenter Control: Any measurement moderately influenced by the experimenter (i.e., a test written by the experimenter and then used to determine the success of an intervention).

Low Experimenter Control: Any measurement unlikely to be influenced by the experimenter (i.e., a blind test-

ing situation, use of a standardized, published test, etc.)

26 Source of Measurement

Experimenter Developed: The experimenter has developed the testing instruments.

Published: The testing instruments used are published.

Combination: Some of the testing instruments used are experimenter developed and some are published.

Other: The testing instruments used are neither experimenter developed nor published.

27. Type of Measurement

Self-Report: The participant rates the effectiveness of the inservice training.

Behavioral Observation: The participant is observed and rated on a behavioral scale in order to determine the effectiveness of the inservice training.

Participant Test Results: The participant is measured on a written test in order to determine the effectiveness of the inservice training.

Student Test Results: Students of teachers who participate in inservice training are tested to see if the teachers' training had an effect upon the students' learning.

Combination: Any combination of the above four categories.

28. Effect Formula

Mean and Standard Deviation: Employed to determine effect size when the mean and standard deviation are presented for experimental and control groups.

t-Test: Employed to determine effect size when a t-statistic is given.

F-Test: Employed to determine effect size when an F-value is reported.

Alpha-Level: Employed to determine effect size when only the α level is reported.

Correlated t with Significance: Employed to determine effect size when only a correlated t, with significance, is given.

APPENDIX D

CODING SCHEME FOR DEPENDENT AND INDEPENDENT VARIABLES

<u>Column(s)</u>	<u>Variable names</u>	<u>Value</u>
1-3	ID	001-999
4-5	Set	01-99
7	Effect level	1=Reaction 2=Learning 3=Behavior 4=Results
8	Instruc- tional focus	1=Improvement of specific subject 2=Improving general teaching 3=Affective techniques
9	Source	1=Journal 2=Book 3=ERIC 4=Dissertation
10	Number of participants	1=1-20 2=21-40 3=41-60 4=Greater than 60
11	Location	1=On-site 2=Off-site
12	Initiator	1=Participants

- 2=Administrator/supervisor
 3=School
 4=State government
 5=Federal government
 6=University researcher
 7=Consultant
 8=Other
- 14 Grade level 1=Elementary
 taught 2=Secondary
 3=Both
- 15 Participa- 1=Mandatory
 tion 2=Voluntary
 3=Both
- 16 Participant 1=Pay
 incentives 2=Certificate renewal
 3=Released time
 4=Status/recognition
 5=College credit
 6=None
- 17 Funding 1=Participant
 2=School
 3=Federal government
 4=State government
 5=University
 6=Other

		7=Combination
18	Instructor	1=Teacher 2=Supervisory staff 3=Administrator 4=Outside consultant 5=College personnel 6=Self 7=State 8=Other
20-34	Types of instruction	1=Lecture 2=Games/simulation 3=Discussion 4=Modeling 5=Coaching 6=Video/audio feedback 7=Mutual assistance 8=Microteaching 9=Practice 10=Printed material 11=Programmed study 12=Observation 13=Production of instructional aids 14=Guided field trips 15=Film

52	Source of measurement	1=Experimenter developed 2=Published 3=Combination 4=Other
53	Type of measurement	1=Self-report 2=Behavioral observation 3=Participant test results 4=Student test 5=Combination 6=Other
54	Effect formula	1=X and S.D. 2=t-Test 3=F-Test 4= α Level 5=Correlated t-statistic with significance
56-60	Effect size	-9.999 to 99.999

003307	313217	329936	000001	000001	532621	112122	112121	00560
003308	313217	329936	000001	000001	532621	112122	112121	-0167
003309	313217	329936	000001	000001	532621	112122	112121	00175
003310	113221	329936	000001	000001	532621	112122	112111	00730
003311	113221	329936	000001	000001	532621	112122	112111	01451
003312	113221	329936	000001	000001	532621	112122	112111	01458
003313	413341	329936	000001	000001	532621	112122	112141	00227
003314	413341	329936	000001	000001	532621	112122	112141	00352
003315	413341	329936	000001	000001	532621	112122	112141	00423
003316	413341	329936	000001	000001	532621	112122	112141	00033
003317	413341	329936	000001	000001	532621	112122	112141	-0012
003318	413341	329936	000001	000001	532621	112122	112141	00024
003319	413341	329936	000001	000001	532621	112122	112141	00035
003320	413341	329936	000001	000001	532621	112122	112141	00120
003321	413341	329936	000001	000001	532621	112122	112141	00614
003322	413341	329936	000001	000001	532621	112122	112141	00219
003323	413341	329936	000001	000001	532621	112122	112141	00246
003324	413341	329936	000001	000001	532621	112122	112115	00060
003325	413341	329936	000001	000001	532621	112122	112115	00000
004001	123311	199955	000001	000001	322611	121111	732215	00674
005001	113126	125565	010000	010010	242421	122222	632212	00792
005002	113126	125565	010000	010010	242421	122222	632212	00682
005003	113126	125565	010000	010010	242421	122222	632212	00635
005004	113126	125565	010000	010010	242421	122222	632212	00445
005005	113126	125565	010000	010010	242421	122222	632212	00650
005006	113126	125565	010000	010010	242421	122222	632212	00445
006001	433341	199955	101111	101000	122622	121222	732215	00120
006002	433341	199955	101111	101000	122622	121222	732215	00120
006003	433341	199955	101111	101000	122622	121222	732215	00000
007001	213126	222247	100000	000001	292422	212222	112134	00816
007002	413342	222235	100000	000001	292422	212222	112144	00000
008001	323216	222235	001110	101001	722621	122221	112125	00745
008002	323216	222235	001110	101001	722621	122221	112125	00745
008003	323216	222235	001110	101001	722621	122221	112125	00745
009001	221299	199997	110110	101001	242992	112122	632213	01230
009002	221299	199997	110110	101001	242992	112122	632213	00845
009003	121299	199997	110110	101001	242992	112122	632213	00717
009004	121299	199997	110110	101001	242992	112122	632213	00000
009005	121299	199997	110110	101001	242992	112122	632213	00000
009006	121299	199997	110110	101001	242992	112122	632213	00000
010001	321419	129999	000000	000000	122111	122122	213223	00387
010002	321419	129999	000000	000000	122111	122122	213223	00205
010003	321419	129999	000000	000000	122111	122122	213223	00268
011001	411416	199956	001000	000011	512511	111122	423224	00739
011002	411416	199956	001000	000011	512511	111122	423224	01063
012001	321219	199999	011000	010110	122992	111222	223223	01077
012002	321219	199999	011000	010110	122992	111222	223223	01715
012003	321219	199999	011000	010110	122992	111222	223223	01050
012004	321219	199999	011000	010110	122992	111222	223223	01811
012005	321219	199999	011000	010110	122992	111222	223223	06661
013001	311419	199999	010000	000010	192922	219222	111254	00645
013002	311419	199999	010000	000010	192922	219222	111254	00000
014001	411325	224335	111000	010100	642412	112222	732215	00000
014002	211325	224335	111000	010100	642412	112222	732215	00351
014003	311325	224335	111000	010100	642412	112222	732215	00251
014004	111325	224335	111000	010100	642412	112222	732215	00351
015001	121111	129925	101010	101111	542621	121222	732115	00506
016001	121111	129925	999999	999999	149922	911191	632114	00877
017001	411419	129995	999999	999999	149612	911191	632111	-007
017002	411419	129995	999999	999999	149612	911191	632111	-180
017003	411419	129995	999999	999999	149612	911191	632111	00156
017004	411419	129995	999999	999999	149612	911191	633141	00336
017005	411419	129995	999999	999999	149612	911191	633141	00687
017006	411419	129995	999999	999999	149612	911191	633141	00301
018001	321416	324335	101010	000010	142611	121221	632125	00418
019001	321416	329335	101010	000010	122611	121221	222125	00608
019002	321416	329335	101010	000010	122611	121221	222125	00000

