This is to certify that the

dissertation entitled

ON-THE-JOB TRAINING, TEACHING PRACTICE,
AND STUDENT ACHIEVEMENT IN THAILAND:
A MULTILEVEL ANALYSIS

presented by

Ikechukwu Chukwuemeka Di-ibor

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Teacher Education

Major professor

Date 11/12/92

ON-THE-JOB TRAINING, TEACHING PRACTICE, AND STUDENT ACHIEVEMENT IN THAILAND: A MULTILEVEL ANALYSIS

By

Ikechukwu Chukwuemeka Di-ibor

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Teacher Education

1992

ABSTRACT

ON-THE-JOB TRAINING, TEACHING PRACTICE, AND STUDENT ACHIEVEMENT IN THAILAND: A MULTILEVEL ANALYSIS

BY

Ikechukwu Chukwuemeka Di-ibor

This study uses multi-level analytical techniques to examine the effectiveness of a nation's in-service teacher training programs on classroom practice and student achievement in Thailand. Findings suggest that training in measurement and evaluation was related to teacher behavior in the classroom, and the enhanced teacher behavior on this dimension of teaching was related to heightened student learning. Training in teacher-student interaction was not related to the quality of classroom instruction, but improved teacher behavior on this dimension of teaching did translate into enhanced student achievement. Training in curriculum and teaching content did not translate into improved instructional quality or student achievement. Training in the use of instructional materials also was not related to the quality of classroom instruction and student learning.

DEDICATION

Dedicated to my grand mother, Mrs. Margaret Ekenma Di-ibor, my mother, Mrs. Chinwe E. Di-ibor, my father, Chief Jonathan O. Di-ibor "OGBUEFI NZE N'EDOZI OBI", my uncle, Chief S.M.C. Di-ibor, B.A (Hons.), D.P.A., (Retired Permanent Secretary), "OGBUEFI NZE ANOCHIE", my aunt, Mrs. Maria Nwokoye, and the entire family of Chief Di-ibor for their love and the value they placed on the education of the young people.

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my gratitude to those persons who have assisted me in various ways throughout my studies at Michigan State University. Special thanks are extended to Dr. Stephen W. Raudenbush, who served as my academic advisor and the chairperson of my doctoral dissertation committee. Without his guidance and advice this study would not have been completed. Working with him contributed greatly to my knowledge, development and understanding of applied statistics and computer programming.

I am specially thankful to Dr. Michael Sedlack for his inspiration and strong support, to Dr. Charles Blackman and Dr. Eldon R. Nonnamaker for their contributions to this study.

My sincere appreciation is extended to Dr. Rex E. Ray and Dr. George Ferns for their helpful suggestions and moral support especially at the beginning of my program. Thanks are also due to Dr. Kenneth B. Ebert and the staff of the Office of International Students and Scholars for their help and cooperation.

My deep appreciation goes to Mr. John and Mrs. Gloria McMeans who treated me as part of their family during my graduate program at M.S.U. Their support, encouragement and

inspiration will always be remembered. Thanks are also extended to Mrs. Hazel V. Evans and her family for their generosity and hospitality, and to Emeka and Khalani for their love, support, patience and many sacrifices during difficult times.

I would like to express my heartfelt thanks to my beloved grandmother, my brother, sisters and the entire family of chief Di-ibor back home for their understanding, prayers, and endurance for the deprivation and hardships that my absence must have caused them.

Above all, my unqualified gratitude goes to the almighty God, who alone can take the credit of keeping me healthy and in good spirit throughout my stay in the United States.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

STATEMENT OF THE PROBLEM

Introduction

Improving teacher classroom practices by improving teaching proficiency on different dimensions of teaching is critical to the success of educational reforms which have swept across many developing nations for the past two decades (for reviews on educational reforms, see Chapman & Carrier, 1990; Wheeler et al, 1990; Chantavanich & Fry, 1988; Kelley & Lassa, 1982; Thompson & Greenland, 1981; George Psacharopoulos, 1991; Cooper Odaet, 1991; Milton Krieger, 1988).

These reforms are in the form of either Universal Free Primary Education (UPE), or changing school curriculum to meet local needs and applying various strategies to boost the quality of education. The governments of developing nations have either implemented one of these reforms or a combination of them at one time or another. For example, Guthrie (1989), noted that as increasing numbers of colonies become independent, curriculum reform is in the forefront of educational change in those developing nations. The purpose of

this reform is to help prepare local citizens to meet the demands of the new national objective of independence.

This assertion is not to understate the importance of investments aimed at providing more school buildings, upgrading management systems, hiring more teachers, and providing cost effective technology and materials for learning. Although these investments are needed, they will not achieve their objective if the teacher fails to deliver effective classroom instruction. As Gage et al. (1989) pointed out,

> Teachers have the power to improve what goes on. By influencing teachers, society can improve education. Even if improvement were sought through the influence of students, it will be the teacher through whom that student influence could be sought. (p. 261).

For example, providing universal free primary education makes sense only if pupils acquire basic skills. But these skills are acquired more easily when teachers possess curricular knowledge and instructional skill as well as the willingness to do the job thoroughly. Providing access to education without adequate planning on how best to prepare the teaching force on effective classroom instruction may be counter-productive (ie, wasteful).

Policies aimed at improving different dimensions of classroom practices are of central economic importance. In virtually all countries of the world, teacher salaries constitute the greatest portion of educational expenditure Therefore, any cost-effective policy that could be found to

improve teacher competence ought to be pursued vigorously. Such policies not only make teachers more productive on the job, but also more efficacious over the life span of their careers. An efficacious teacher is a teacher who is competent in both the curriculum content and classroom practice, and who can combine both skills for effective classroom instruction. When teachers operate at this level, there are positive results for both students and teachers. Not only does the achievement of students increase across subject domains and grade levels, but the source of self efficacy and job satisfaction improves for the teacher.

In developing countries, policies aimed at improving teacher classroom practices are particularly important, since expenditures for teaching account for upwards of seventy percent of the national educational budget. For example, Sub-saharan African teachers' salaries account for approximately 90% of the primary education budget, and 70% for secondary school recurrent expenditures (Lockheed & Komena, 1989). Thus, understanding how teachers perform their jobs, how they learn on-the-job to improve different dimensions of classroom practice, and how improved dimensions of classroom practice contribute to student learning are key to improving both educational quality and efficiency.

In some Third World countries, the teacher's role ranges beyond that of classroom instruction to include that of government agent in some rural areas. A good case in point is

Thailand, where schools in rural areas represent government agencies (Vongkek, 1982). Teachers play a critical role in both rural and urban development activities. These activities range from participating in various local community projects to initiating and helping local people understand government programs. For some teachers, the job is not over when the school day is over. They have to go into the villages to preside over local committees where they help explain government policies that affect the local people. These assignments are carried out without extra pay or overtime pay. Even in the school, the conditions under which these teachers operate are difficult. Yet the teachers work hard to educate their pupils. This assertion is well summed up in a paper presented at the Comparative International Education Conference in Pittsburgh, titled "Good Schools and Poor Schools: How They Differ" (Warwick & Reimers 1991). The authors state that

> Within the school the teacher is the single most important person, encouraging student achievement. It is the teacher who decides whether and how to follow the curriculum and the textbooks, who inspires fear, confidence or both, who helps or ignore students with problems in following the lessons, and who creates an atmosphere supporting learning in the classroom. Even with no desk, chairs and other equipment, an inspired teacher can motivate pupils to learn while with the best of facilities, a harsh or lazy teacher can so demoralize students that they drop out of the school (p. 8).

Why has the emergence of teacher competence on dimensions of classroom practice become a crucial variable in the Third

World educational policy-making process? Two historical developments in the field of Third World education will help us answer this question.

First, since the end of the Second World War, Third World governments and development agencies have emphasized school expansion, not necessarily the improvement thereof (Fuller, 1987). The policy of school expansion has had dramatic results: many countries in developing nations have implemented universal free primary education for the past two decades.

There has been rapid expansion of school buildings, facilities, equipment, and educational materials. There has also been massive credentialing and deployment of teachers. Although near universal enrollment in primary education has been attained by the vast majority of these countries, low levels of achievement and low school quality continue to persist, (Fuller, 1987).

The deterioration in educational quality has prompted Third World governments and policy makers to shift their focus from access to quality, and there is a growing concern that the majority of teachers already in place or hired during the period of expansion lack the basic education and pedagogical skill required to stimulate an acceptable level of student learning.

Second, research has begun to center on the classroom and classroom processes as important determinants of learning, with the specifics focusing on the role of teachers and

administrators as managers of student learning (Lockheed et al; 1990). The underlying premise is that enhancing student achievement by improving teacher skill and knowledge of effective classroom practice relies heavily on administrative and teaching quality.

On-the-Job Training of Teachers on Dimensions of Teaching

Considerable research attention is underway in both developed and developing nations on cost effective ways of educating teachers on the job. Teaching, like any other profession, requires constant professional development in order to keep pace with technological advances. In-service training represents one cost effective strategy of developing teachers on the job for proficiency in classroom practices. Although there are other cost effective strategies for boosting on-the-job proficiency of teachers, including teacher supervision, teacher specialization, and the use of instructional materials, none have received the same level of commitment and resources from policy makers and development agents in the Third World as in-service training.

Problem Statement

Third World countries have practiced one, or combinations, of these strategies of on-the-job training of teachers in the past. But in-service training has recently emerged as the single most widely employed strategy of

improving instructional quality in Third World countries (Chapman, 1990).

For example, some developing nations like Liberia have gone the furthest in depending exclusively on in-service training programs for the preparation of primary school teachers (Chapman & Carrier, 1991). The cost of a national in-service training program could severely strain the education budget of developing nations. The major economic implication of this type of training is that in-service training is offered on different topics across the country. Since the cost and types of training programs vary, conducting a study on comparative effectiveness of on-the-job training of teachers on different dimensions of classroom practice is essential in formulating good policy for improving educational quality in several ways.

1) It will help monitor the implementation of a training program to ensure that training is being carried out according to plan;
2) It will help to review teacher activities on different dimensions of teaching and provide feedback as to whether their actions are in keeping with the original objective and intent of the program.
3) It will help to understand the effect of training on teachers and the effects upon the students of the teachers receiving training.

4) The information obtained from this exercise could be used as the basis for immediate revision or discontinuation of a training program.

5) It could also be used for long-term planning and assessment of a program.

In both cases, a comparative study of the effectiveness of on-the-job training of teachers on dimensions of classroom practices could help policy makers guard against poorly conceived or ineffective in-service training programs being perpetuated in the past (Marshal 1988). Budget savings in this area could be directed to more promising elements of improving school quality.

Another advantage of conducting comparative effectiveness studies of on-the-job training of teachers on dimensions of classroom practices is that it will help to provide insight into the issue of which of these dimensions actually affects how much a child learns in school. If one or more dimensions are found to have a consistent positive relationship to student achievement, such a dimension may be encouraged, while the ones that do not show a consistent relationship should perhaps be discontinued. The resulting savings from this exercise could be channelled into other cost effective strategies that boost instructional quality.

Purpose of the Study

The purpose of the study is to examine the effectiveness of on-the-job training of teachers for competence on several dimensions of classroom practices and student achievement. The dimensions of classroom practice include in-service training in teacher-student interaction, measurement and evaluation, curriculum and teaching content, and the use of instructional materials. It is necessary to investigate how these dimensions of classroom practice relate to students' perception of instructional quality (sub-scales), and resulting student achievement. Specifically, the investigator will explore the total effect of training and the effect of the subscale of teaching quality on student academic achievement.

The findings of this study will be interpreted in the light of extensive field data from rural Thai primary schools. These field investigations have produced a series of case studies of highly similar topics. A series of analyses of these data have already been reported (Raudenbush & Bhumirat, 1992; Raudenbush, Bhumirat, & Kamali, 1992; Raudenbush, Easukkawat, Di-ibor, Kamali, & Wimol, 1992).

Contributions of the Study

This study contributes to the literature on research on teaching and classroom learning in several ways:

1) It extends the evidence on the effect of teaching quality and student achievement in developing

countries by analyzing data from the BRIDGES project, conducted by Michigan State University researchers and the officials of Thai government during the spring of 1988.

2) It provides a complete and systematic evaluation of a national in-service training program.

3) It seeks to determine the degree to which the dimensions of teaching practices identified as effective in developed countries are equally effective in a developing nation.

4) By applying a credible and statistically powerful new model for analysis of school effect data, this study constitutes a significant application of multi-level models.

5) Finally, it is hoped that this study can contribute to policy-making initiatives in Thailand designed to improve the quality of classroom instruction.

Research Questions

Total effect of training on student achievement:

(1) What is the effect of in-service training in teacher student interaction on student achievement?

(2) What is the effect of in-service training in measurement and evaluation on student achievement?

(3) What is the effect of in-service training in curriculum and teaching content on student achievement?

(4) What is the effect of in-service training in the use of instructional materials on student achievement?

Effect of classroom practices on student achievement:

(1) What is the effect of teacher classroom practices in teacher student interaction on student achievement?

(2) What is the effect of teacher classroom practices in measurement and evaluation on student achievement?

(3) What is the effect of teacher classroom practices in teacher fairness and concern on student achievement?

(4) What is the effect of teacher classroom practices in the use of instructional materials and student achievement?

On the issue of the comparative effectiveness of training, the question often asked by policy makers is how effective is a training program in achieving its objectives.

Effectiveness of training:

(1) What is the effect of in-service training in teacher student interaction on student perception of instructional quality (sub scale teacher student interaction)?

(2) What is the effect of in-service training in measurement and evaluation on student perception of instructional quality (sub scale measurement and evaluation)?

(3) What is the effect of in-service training in curriculum and teaching content on student perception of instructional quality (sub scale teacher fairness and concern)?

(4) What is the effect of in-service training in the use of instructional materials on student perception of instructional quality (sub scale use of instructional materials)?

Dimensions of Teaching Used in the Study

Discussions on dimensions of teaching will focus on four types of classroom practices that are commonly found in Third

World classrooms. These include teacher student interaction, measurement and evaluation, teacher fairness and concern toward students, and teacher use of instructional materials.

Teacher- student interaction. Effective in-service training on teacher student interaction should improve a teacher's competence in daily interaction with students. Teachers should realize that interacting with students as individuals should be of major importance to them.

As Sanborn (1987) noted, teachers must assume the attitude that they are not just teaching skills, but are teaching individuals in a group setting and that their primary responsibility is to help each individual progress in skill acquisition. By interacting better with their students, teachers can more effectively deliver classroom instruction to their students.

Measurement and evaluation. Effective in-service training on measurement and evaluation should improve teacher competence in the timely evaluation of student academic progress with frequent and regular feedback to the students. Research on school effectiveness suggests that feedback is a necessary condition for learning and that the best way to motivate students to learn is by giving them timely feedback on progress that they are making in school (Siedentops, 1988).

Teacher fairness and concern. Effective in-service training on curriculum and teaching content should improve teacher competence on teacher fairness and concern towards

students' welfare. Students who perceive that their teacher is fair and shows concern for their well being will learn better than they will under an insensitive teacher. A cross-national study organized by the international association for the evaluation of educational achievement (IEA) indicates that teachers were lecturing and talking at students in Nigeria and Thailand in most of the classroom sessions. This type of teaching makes teachers look authoritarian. It also may threaten students and impede their ability to grasp academic materials from the teacher, as the students may be preoccupied by feeling intimidated by the teacher.

Use of instructional materials. Effective in-service training on the use of instructional materials should improve teacher competence on the use and development of instructional materials. Evidence from research in Third World countries suggests that text books are crucial inputs for assuring high quality instruction in schools and for raising student achievement (Fuller, 1987).

The use and development of instructional materials may affect the rate at which the teachers teach and learn about their subject (Lockheed, Vail, and Fuller, 1986). As these authors have argued, textbooks and locally developed instructional materials may provide teachers with a more structured and comprehensive representation of the subject matter than would otherwise be available to them. Involving teachers in both the development of and production of teaching

materials will lead to the use of appropriate teaching materials in the classroom. The use of appropriate teaching materials may also facilitate student learning of the subject matter.

The Context

The context of the study will be the Thai primary school. Thailand, meaning "Land of the free" is a tropical country in Southeast Asian with an area of about 200,000 square miles (approximately the same size as France). It is bordered by Malaysia to the south, Burma to the west, Laos to the north and northeast, and Cambodia to the east. Thailand is one of the very few developing nations that was never colonized, thereby making its educational system comparatively free of any foreign power. Central Thai is the language of instruction in primary schools.

The country has a population of approximately 47 million. The 1950s and 1960s were characterized by population growth of slightly over 3 percent per year, resulting largely from a marked improvement in public health that started at the end of the Second World War. The 1970s, however, saw an annual growth rate of approximately 2 percent, slowing to 1.5 percent in the 1980s (Ministry of Health report, 1984). A major school mapping project initiated in 1978 found striking drops in early primary school enrollment resulting from a decline in fertility.

Primary education is free, provided universally by the government. The curricular structure emphasizes literacy, numeracy, communication skills, and abilities relevant to future occupational roles. The structure of the educational system is 6-3-3, which means six years of primary education, three years of junior secondary schools, and three years of senior secondary schools. Major problems at both primary and secondary level relate primarily to quality and equality. The focus in this study is solely on quality.

Thai primary schools provide an interesting context in which to examine the effectiveness of on-the-job training of teachers on classroom practices. Since Thailand has achieved near-universal primary access within the last 10-15 years, Thai policy makers have focused intense effort on improving the quality of primary education. Major emphasis on improving the competence of the teaching force has taken priority among educational planners. This effort was demonstrated by a variety of educational reforms taking place in this country. These include the implementation of a national in-service training program for teachers, with special emphasis on improving classroom practice, and a national program of in-service education for principals with a strong emphasis on classroom supervision, and the organization of schools into clusters each having a resource center and each encouraged to involve teachers in developing and using innovative

instructional materials (Wheeler, Raudenbush, and Pasigna, 1992).

A similarity between Thailand and other Third World nations is in the area of the implementation of universal free primary education with a large reservoir of teachers, many of whom lack the necessary subject matter knowledge and pedagogical skills. Thailand differs in its demographic history: the period of most rapid educational expansion corresponded to the period of most rapid population growth. These factors necessitated the very rapid certification and deployment of primary teachers. Shortly after 1970s, population growth slowed dramatically. The result has been that the country has had many number of untrained teachers. Because relatively few teachers have been needed, the reform of pre-service education has provided only limited opportunity to improve overall teacher competence. Thai policy makers have recognized this constraint, and have focused considerable attention on improving the productivity of already practicing teachers (Raudenbush et al., 1992).

The circumstances which have impelled Thai policy makers to concentrate on in-service teacher training are not clearly manifest in other developing nations, many of which are still struggling to achieve primary access, and many of which continue to experience rapid population growth. Nevertheless, each of these nations must be increasingly concerned with improving low levels of competence of practicing teachers. The

richness of Thailand's experience in attempting to do so, though motivated in part by its special circumstances, may provide an excellent example to other nations contemplating policy options for fostering on-the-job training of teachers.

Design and Sample

The survey procedure utilized by the Thai government officials has made it impossible to link each sixth grade teacher with that teacher's students. Thus, while it has been possible to link school characteristics with individual teachers in each school and to link school characteristics with individual students in that school, it has not been possible to link schools, teachers and students in a single analysis. Hence, the only teacher-level data available for predicting student achievement have been aggregated levels of the teacher variables. Characteristic flaws with in this type of analysis are well known (c.f., Aitkin & Longford, 1986). The central flaw in the present case is that, to the extent teachers within a school vary on the independent variable, the use of the mean teacher value on that variable to predict student outcomes produces uninterpretable results.

Recently, the investigators discovered that over 120 schools in the sample have only one sixth grade classroom and one sixth grade teacher. Therefore, in these schools there is no ambiguity about the link between the teacher and the students. Moreover, these schools are small rural schools

which provide the greatest challenge to policy aimed at improving teaching quality. Because of the availability of the national representative sample of schools, it is possible to locate these schools quite accurately in terms of national distribution of a variety of important contextual variables (Raudenbush et al., 1992). These schools will be referred to in this study as sub-sample and will constitute the sample for this study.

Limitations of the Study

The following constraints would constitute limitation to this study:

1) The findings of this study will be generalizable only to those types of in-service education programs wherein the elements described in this study are present.
2) This study is observational and not experimental, so that causal inferences should be made with caution.
3) The data to be used in this study will represent the response to a national survey of primary schools in Thailand. Generalization of the results to other Third World countries may not be appropriate when differences in cultural settings are not considered.
4) In-service training in curriculum and teaching content was used to predict the subscale of teaching quality in teacher fairness and concern. It is not known

whether the content of the training reflects this dimension of teaching.

5) The investigator did not participate in any of the training sessions, thereby judgement on the quality and effectiveness of the training program cannot be made.

6) In-service training program may vary from province to province or school district to school district.

7) Pre-service education of teachers may or may not affect the quality of classroom instruction and student achievement.

Delimitations of the Study

This study will be delimited by the following constraints:

1) It is not the primary purpose of the study to defend the need for in-service education of teachers on classroom practice.

2) The primary purpose of the study is not to defend a single in-service education program, but to evaluate a national in-service teacher education programs.

3) The study will be limited only to rural schools in Thailand.

Methodological Considerations

The primary analytical method employed in this study is the hierarchical linear model (HLM) (Raudenbush and Bryk, 1986). Such models have been referred to in sociological research as multilevel linear models (cf. Mason et al., 1983; Goldstein, 1987), in biometrics applications as "fixed and random coefficient regression models", and in the econometrics literature (cf. Rosenberg, 1973) and in the statistical literature as "covariance components models" (Dempster, Rubin & Tsutakawa, 1981; Longford, 1987).

HLM is a regression technique well suited for analyzing Multi-level data sets. This method enables one to analyze data simultaneously at different levels of the educational hierarchy, at the classroom level, school level and the district level. This means that a choice does not have to be made with respect to the appropriate level at which to analyze the data. Nor does one have to make the unrealistic assumption that there is no covariance between two students in the same class. Instead, the fact that two students within a particular class are more alike than are two randomly selected pupils is explicitly incorporated into the model. This method enables the analyst to take into account the effect of the shared membership of students in the classrooms and schools and to account for the effect of school and classroom policies measured at the group level on processes occurring within

groups, such as teaching and learning and different dimensions of classroom practices (Goldstein, 1987).

In this study, an HLM will be employed in which students are viewed as nested within schools or classrooms. (Recall that this study will focus on rural schools). These schools are comparatively small in student enrollment, with one classroom and one teacher per sixth grade level. Intervening outcomes at the teacher level will include perceived instructional quality, using four dimensions of classroom practice. Outcomes at the student level will include perceptions of instructional quality (four sub-scales), and achievement in five subject areas of the Thai curriculum. The availability of students' perception of instructional quality (sub-scales) at both levels allows the analyst to disentangle the effect of a class rating of the teacher from the individual differences in perceived instructional quality (sub-scales) among students in the same classroom.

Required Effort

The data for this study are stored in three separate files (student, teacher, school), each of which is a large-scale data base, containing hundreds of variables. Schools with a sole sixth grade classroom and one practicing teacher were selected from the school data and then matched with their students and teachers. This was accomplished by writing a computer program on the IBM mainframe computer at

the computer center, Michigan State University, using Statistical Program for Social Science (SPSSx). Within each file the scales were recreated, and their psychometric properties re-examined. Graphical and exploratory analysis were employed to understand and confirm the validity of the scales. The teacher and the school file were merged into a single school file. Again, recall that each school has only one sixth grade class and one sixth grade teacher. Separate school and student files provided the input into a merged data file labelled "a sufficient statistics file (SSM file)". The SSM file was subjected to analysis by means of the method of the hierarchical linear model as implemented by the HLM program (Bryk, Raudenbush, Seltzer, & Congdon, 1988).

Summary

In this chapter the need to improve the quality of classroom instruction by boosting on-the-job learning of teachers in the Third World has been stressed. Although educational issues may vary from developed to developing nations, all countries of the world may strive for better education, by improving the teaching learning process. For example, providing universal primary education in a Third World country or reducing the high school dropout rate in an industrialized country are crucial development issues in those societies. Developing the cadre of scientific and technical personnel to lead the world into the twenty-first century in

high tech information, biological, and material sciences is as important in the United States as it is in Mozambique. The rapid change in the world economy and the growing gap between developed and developing nations suggests that not all countries face the same educational problems. Many countries are striving to increase student learning and to equalize access to knowledge, and countries may learn from one another in using educational techniques that work.

Recall that this chapter emphasized the need to improve on-the-job learning of incumbent teachers for competence in different dimensions of classroom practice. Improving the teaching-learning process by improving the quality of teaching on dimensions of classroom practice, however, represents only one possible avenue of productive reform. Both in developed and developing nations, the need for quality education has never been felt more than now, as a result of rapid technological changes. At the center of it all are the teachers. Teachers mediate students' encounters with content or curriculum, and they control the classroom activities most directly related to learning. Thus, any policy designed to improve the competence of on the job learning of teachers will be cost effective policy. One of these policies should focus on in-service training, teaching practice and student achievement. The literature review will begin in the next chapter. Chapter Three will focus on methodology. In Chapter Four, results will be presented. Conclusions, recommendations

and implications for policy and future studies are to be found in Chapter Five.

CHAPTER II

REVIEW OF LITERATURE

This review of the related literature will present an overview of effective in-service training programs, dimensions of classroom practice and methodological considerations in the analysis of hierarchical data. The theoretical position on teacher training and student achievement is that training shapes teachers' classroom behavior on dimensions of teaching practice, which in turn affect student achievement. Thus, training and dimensions of teaching are related in this context. Therefore, in order to examine the effectiveness of training in this study it will be necessary to know the following: 1) Can training improve teacher performance in the classroom? 2) If the answer is yes, can improved teacher classroom performance translate into enhanced student achievement? and 3) Which dimensions of classroom practice are effective in improving student performance.

Section 1

Prior research on in-service training . Research on in-service training may broadly be classified into two categories: studies designed to discover whether a training program has positive effect, and studies designed to identify which types of training programs are most effective. Notably

lacking are studies that examine simultaneously the effectiveness of a national teacher training program on such dimensions of classroom practice as teacher-student interaction, measurement and evaluation, curriculum and teaching content, and the use of instructional materials.

The first type of literature review will focus on whether there is evidence that teacher participation in an in-service training program results in improved performance of classroom practice and student learning. Evidence for this type of research could be found in the United States. For example, Gage and Needles (1989) summarized the results of 13 experimental studies in the United States, in which the effect of in-service teacher education programs on teacher behavior and student achievement was evaluated. They found that, in 12 of the 13 studies, in-service training was found to have a positive effect on teacher classroom practices. In 10 of the studies, significant positive effects on student achievement were found. The usual design for this research was an experimental design, contrasting two groups of teachers: A treatment group which experienced the training, and a control group which did not. The usual scientific procedure for this type of study is an investigation in which the "independent" variable is manipulated and the subsequent values of another or "dependent" variable were measured.

An experimental study on teacher classroom practices was conducted by Ingvarson & Mackenzie (1988). In Australia, a

central component of government strategies for improving the quality of education and the implementation of new programs and policies is often the short in-service course or workshop. Ingvarson & Mackenzie have conducted a study on the impact of short in-service course in the state of Victoria, Australia. They conducted follow-up studies on 2000 teachers who had attended a week-long course on the application of micro computers in schools. Their purpose was to find out to what extent the participants in their study would act as "change agents" when they returned to their respective schools. They found out that what teachers did varied from school to school, and that the administrative and follow up assistance that teachers received in the school accounted for this variation. They concluded that

> returns to investments on in-service education by school systems will be limited if the need for such support and assistance is not anticipated, that is if planning for policy implementation goes no further than the provision of in-service courses alone, without the orchestration of follow-up support both from within the school and from external sources. (p 139).

Experimental studies are also not new in the Third World country. Nitsaisook & Anderson (1989) conducted an experimental study of in-service teacher education in Thailand. In this experiment, one hundred and fifty five fifth grade mathematics teachers were assigned at random to control and experimental groups. The experimental group participated in a six day workshop designed to improve teaching practices. Before the intervention, mathematics and attitude surveys were

administered to all students. Following the treatment, all teachers taught six month-long mathematics units, with each classroom being observed once per unit. At the end of the experiment, the mathematics and attitude surveys were re-administered to the students. The purpose of their study was to examine the effect of teacher participation in an in-service training program on teaching practices and student achievement. Significant results were reported for the experimental group on improved classroom practices and significant findings were also reported for their student achievement.

Evertson (1985) and Evertson, Weade, Green, & Crawford (1985), experimented with 16 teachers in grades 7-9 in mathematics and English in Arkansas, over one semester. The teachers were paired on the basis of their similarity in teaching experience and grade level. Within each pair, teachers were assigned randomly to either the experimental or to the control group. The experimental group participated in a one day workshop before the beginning of the school year, using a manual developed from the Texas studies of effective management. A one-day follow-up workshop that emphasized management techniques and discussion of problems was also attended by the experimental group. The control teachers did not participate in any workshop. All teachers were observed on the first class day, and on seven other occasions during the first eight weeks of school. Measures of teacher behavior were

taken from four different instruments: Component rating; Addendum component rating; and Narrative reader rating. Of the 35 point ratings, 22 showed significant differences in favor of the treatment group. To asses the effects on students, comparisons were made between the amount of disruptive behavior and inappropriate behavior and the amount of student task engagement. The experimental classroom had significantly fewer instances of student task disengagement.

Good & Grouws (1979) investigated the effectiveness of an experimental mathematics teaching program on teacher practice and student achievement in a fourth grade classroom in the United States. The treatment program was based on findings from the naturalistic study of effective mathematics teaching. Students were tested before and after the intervention with both standardized and content tests designed to approximate the actual instructional content that students had received during the treatment. Results of the study revealed that treatment teachers generally implemented the treatment, and analysis of product data showed that students of treatment teachers generally outperformed those of the control teachers on both tests (standardized and content test).

In another experimental study of the effectiveness of an in-service teacher education program for K-12 mathematics and science teachers, White, Conwell, & Passe, (1988) were determined to discover whether in-service training program significantly increased teachers' allocation of time for

problem solving, cooperative learning, use of Equal Station, and infusion of career related content during mathematics instruction. Observation of an experimental group of six teachers from three different schools and a comparable control group was conducted in North Carolina. There were no significant differences between groups prior to the training. Significant differences were reported post training for the experimental teachers, who increased mathematics instructional time to 41%, and for cooperative learning instruction to 22%. Significant differences were reported in small-group problem solving as well.

In a similar vein, Emmer, Sanford, Evertson, Clement, & Martin (1982), discussed an experimental field study in which teachers extended a 3-hour workshop and studied a manual of recommended teacher behavior as a means to promote more effective classroom management. Subsequent research revealed that the treatment group did in fact use more of the desired management behavior than the control teachers.

Not all research in education is experimental; sometimes special circumstances in field research can prevent the researcher from conducting an experimental study. For example, it is more difficult to assign students or large social groups to a treatment at random than it is to assign agricultural plots. It is also more difficult to assign students to treatment at random in a field setting than in the laboratory setting.

A quasi experiment was conducted in Israel by Fresco & Ben-chain (1985). In this quasi-experimental study, two in-service training courses designed to strengthen teachers' subject-matter competencies, while introducing them to mathematics curricula for grade 7 and 8, were investigated. In a pre and post-test, participants were given a confidence measure and knowledge test. At the end of the training, teachers manifested greater self confidence in their skills, and higher levels of confidence in their ability to teach the curriculum. Results suggest that when knowledge tests are not feasible, measurement of confidence in solving problems may be sufficient to evaluate the cognitive impact of an in-service training program.

Sometimes research on in-service training may focus on teacher outcome rather than student outcome, as in an analysis of a large-scale national survey research conducted in a developing nation (Raudenbush, et al, 1992). This is an observational study in which a hierarchical linear model (HLM), was used to study the relationship between in-service teacher training, student perception of instructional quality, and teacher sense of efficacy. Subsequent findings revealed no effect of in-service training on teacher sense of efficacy. But in-service training was found to be significant and positively related to the quality of classroom instruction.

In another attempt to examine the effectiveness of a national in-service training program in the Third World,

Chapman (1990) discussed the implication of a national in-service training program for educational reform in Liberia. According to Chapman, the Liberian government has recently relied solely on in-service training for the preparation of their primary school teachers. In-service training in Liberia occurs in four ways: (1) Teachers receive direct instruction in the summer; (2) they learn both pedagogy and content from actually using the program materials; (3) they get further on-the-job training from their instructional supervisors; and, (4) they learn from radio broadcasts aimed at reinforcing their teaching. Evaluation conducted at the pilot stage of this national effort indicated that students receiving instruction from these teachers did somewhat better than those in conventional Liberian schools (Kelly, 1984; Boothroyd & Chapman, 1988).

In some cases, investigators have designed studies in which more than two types of training are compared to a control group. Anderson & Djalil (1989) compared the effects of "intensive training and minimal training" on Indonesian fifth-grade teachers. They found that intensively trained teachers changed their behavior more often than the minimally trained group, and that their students also out-scored their counterparts in social studies tests. (The question of intensity is crucial to policy because of the obvious cost implication of intensive training). Gage & Grouws (1979) Anderson et. al, (1979) Crawford et. al, (1978) reported

results that suggested that both minimally intensive and inexpensive training can significantly boost teaching quality. Coladarcia & Cage (1984) disagreed with this finding. They argued that the minimally-trained teachers appeared to have unwittingly regarded the relatively frequent and lengthy classroom observation as a kind of supervision, or monitoring. If so, the conduct of classroom observations probably would have enhanced the compliance of the intensively trained teachers with the training recommendations.

Summary and implications I. Previous research on effective in-service training program reviewed above, suggests that in both developed and developing nations, in-service education can improve teacher classroom performance and student achievement. Therefore providing in-service teacher training programs for incumbent teachers is one strategy to improve instructional quality and student learning. The more crucial question to policy involves the kinds of programs that are effective, and how these programs can improve different dimensions of classroom practice together with the level of investment required to produce significant results.

Types of effective programs. The second question in this literature review is designed to discover which types of in-service training programs are most effective. In a review of large varieties of in-service training programs in the United States, Joyce & Showers (1980) summarized the results of 200 studies of in-service training programs. They found

that the most effective training was characterized by five components: presentation of theory or description of skill or strategy; "modeling" or demonstration of skills, or models of teaching; practice in simulated and classroom setting; structured and open ended feedback; provision of information about performance or coaching for application (hands-on in-classroom assistance with transfer of skills and strategies to the classroom).

Regarding in-class assistance, Joyce & Showers (1981) speculate from research findings that what they call "coaching" considerably increases the effectiveness of attempts to alter teacher behavior. Coaching in the classroom is analogous to coaching on athletic fields in that on-going assistance is provided, feedback and other valuative data are considered, and evaluation and refinement are a continuous process. Based on reviews of research on in-service training, Joyce and Showers offer a model that suggests that increasingly positive effects are found when a program moves from theoretical understanding to observation, to clinical practice to coaching. This inferential model could be quite useful for staff developers if positive outcomes can be substantiated.

The most typical staff development activity is a workshop, usually a one time attempt to alter the behavior, beliefs, and or thoughts of participants. In a recent report, Emmer, Sanford, Evertson, Clement, and Martin (1982), discuss

an experimental field study in which teachers extended a 3-hour workshop and studied a manual of recommended teacher behavior as a means to promote more effective classroom management. Subsequent research revealed that the treatment teachers did in fact use more of the desired management behavior than control teachers. They argued that because the workshop was only 3-hours in length, one looks for reasons other than the workshop to explain the effects. They insist that what might have been overlooked in examining the research report is the important role played by the manual given to the treatment teachers, which described real classrooms and actual teaching situations.

As the government of Thailand moves from staffing classrooms to upgrading the skills of practicing teachers, single in-service sessions were held on the specific dimensions of classroom practice. A report published by the Office of National Primary Education Commission (ONPEC (1985) concluded that such activities failed to improve instructional quality and student achievement. Thus this strategy for improving classroom instruction has been phased out.

Fullan (1982), has criticized short workshops of in-service training of teachers as one of the reasons for the inadequacies of current professional development activities in North America.

Fullan (1982) has also outlined a set of eight features which is believed to be supported by research findings and to

be related to the problems and issues of in-service training. The features are closely related, and it is apparent that in some instances they over-lap.

1) It will be designed as a consequence of a systematic problem identification by those most directly related to the problem.
2) It will be interactive.
3) It will mitigate to some degree status differences between teachers and administrators.
4) It will depend less on consultants and more on teachers and administrators for substantive and procedural guidance.
5) It will be formulated and monitored largely according to perceptions of the participants.
6) It will be formulated, in part, in terms of a careful analysis of the organization and the people for whom it is intended.
7) It will be flexible and responsive to the changes in participants and the changes in the setting.
8) It will be within reasonable limits, and be situation specific.

Summary and implications II. Fuller (1987), in an exhaustive review of school factors that raise student achievement in the Third World, commented on the scarcity

of research on the effectiveness of in-service training programs.

> Very little evidence exists on the effectiveness of in-service teacher training programs (Indicator 2). This scarcity of knowledge is in stark contrast to the increasing level of resources being invested in upgrading the skills of incumbent teachers. For instance, in the last decade, two thirds of the World Bank's education projects have included in-service teacher training components. Only four studies have examined the influence of such efforts. (p. 281).

The studies reviewed above indicate that short term courses without classroom follow-up do not constitute effective in-service training programs. There is every reason to believe that if teachers are to utilize the knowledge that in-service training programs are designed to achieve, the application of the theory must be demonstrated in classroom teaching. The teachers need an opportunity and school support to practice what they have learned and to refine pedagogical techniques and follow-up in actual classrooms, supported by classroom observation and structured feedback. Teacher involvement in the identification of course content and materials for in-service training are also important. The majority of studies cited in this study suggest a positive link between training and student achievement. One must

however be cautious about hasty generalization. Training does not have to influence achievement in order to be effective. In-service training programs may influence teacher sense of efficacy or teacher satisfaction, and may constitute effective programs. Policy makers should not rely solely on in-service training to boost the quality of education. Rather in-service training should be viewed as a component of effective educational reforms.

Policy makers must not only design relevant training programs of instruction in ways that incorporate the perspective and needs of teachers, but they must also plan for the effective administrative support of the subsequent application of knowledge in classroom teaching. Without adequate classroom coaching, or what could be termed clinical supervision, one should not expect teachers to acquire new and useful strategies.

Of major interest to policy makers are the effects of dimensions of classroom practice on student achievement. The next section of this literature review will focus on these dimensions.

Section 2

Dimensions of teaching. Research on the dimensions of classroom practices could be broadly classified into school-effect research and process-product research.

School Effects

The literature on school effects in developed nations arose by way of a critique of the Coleman et al. (1966) and Jenks et al. (1972) studies. Since those school factors which they measured, including class size, expenditure per student, number of books in the library, the quality of the school building, and nature of the curricula accounted for only a small proportion of the variance in students' achievement, it was concluded that schools make little difference on how much the child learns. Reynolds (1982 a) provides a useful summary of criticisms of these studies as a precursor to a review of more recent research that suggests that school effects had been severely under-estimated.

Despite enduring weaknesses in the methodology of statistical analysis of school effects, such as the difficulties of controlling for student intake factors, the limited nature of school outcome measures and the possible confusion of correlation with cause, it is now widely accepted

that schools do have important effects (Brophy & Good, 1986). The research suggests that such effects are related not to resource-based school input factors, but rather to school process factors that are more elusively categorized as features of school climate or school culture. For reviews of such studies, see Good & Brophy (1986); Rutter (1983); Anderson (1982); and Reynolds (1982 a). Recent literature from the United States is replete with references to effective schools being characterized by strong instructional leadership, clear goals, mission and high expectations, a sense of community and a school climate conducive to effective learning (see for example Purkey & Smith, 1983).

Process-Product Research

The research base for teacher education has been sought in part by determining how more and less effective teachers act and then by trying to get teachers to act in ways that distinguish the more effective ones from the less effective ones. Research in this field has sought objectivity, reliable descriptions and measures, public and communicable procedures, and minimization of biases. They have used devices that would make the results reflect as much as possible the way things are in the classroom. This research movement has come to be known as process-product research on teaching, the search for relations between classroom process (teaching) and products (student learning) (Brophy & Good, 1986).

Sometimes this type of research is also referred to as teacher effectiveness research. Major studies currently labelled "teacher effectiveness research" characterize large scale process-product-research on teaching. In these studies, students' performance at the end of instruction (eg., experimental treatments, teaching packages, or naturally-occurring instructional behaviors), are related to instructional variables (primarily based on classroom observations), and to the characteristics of the students (including pretests of cognitive skills) upon entry into the instructional setting.

The grade level, subject matter, sample size, and primary units of analyses of major teacher effectiveness studies are summarized in a handbook for research on teaching from 1986 (Brophy & Good, 1986). The classroom is the primary unit of analysis, and with few exceptions (e.g., the school or the student) in the earlier studies, the main statistical analyses were correlations between class mean gains (post test minus pre test) or mean residual gains (post test on pre regressions), and teaching behaviors or class aggregates of observed teacher-student interactions. Later studies (eg., Fisher, Berliner, Filby, Marliane, Cohen, Dishaw & Moore, 1978), switched to generalized regression procedures.

Dimensions of Teaching.

Use of instructional materials. The report of research into teacher use of instructional materials will concentrate on studies from Third World nations. This is due to the fact that wide-spread availability of textbooks in developed nations preceded research on the effectiveness of instructional materials. There has been little systematic study of the impact of textbooks on students' achievement in developed nations. In an experimental study of the relationship between textbook availability and the mathematics achievement of students in a Nicaraguan first grade class, Jamison et al. (1981) compared classes in which textbooks are relatively rare with those in which textbooks are used, and with radio-based instructional programs that use student work sheets but no other textual material. Classes were assigned at random to the three conditions. The control and the two treatment groups scored similarly on a pretest of mathematical readiness. Results of the study indicate that both the textbook and radio treatment had significant positive effects on achievement. The availability of textbooks increased student post-test scores by about 3.5 items correct, or approximately .33 of a standard deviation.

In a similar vein, Lockheed, Vail & Fuller (1986) demonstrated that in Thailand, teacher education did not enhance textbook use, but rather textbooks could substitute for additional years of teacher education.

Heyneman, Farrell, Sepulveda-Stuarto (1978) reviewed 14 studies that include 19 assessments of the relationship between the availability of printed materials and student outcomes in Third World nations. Although different methodologies were used and the quality of the data is far from uniform, the positive relationship between input and outcome in 16 of the 19 cases strongly suggests that textbooks are a potentially important and consistent contributor to improved quality in school.

In an exhaustive review of which school factors raise achievement in the Third World, Fuller (1987) summarized the results of the effect of teacher use of instructional materials on student achievement in developing nations. Of the 24 studies on the effect of teacher use of instructional materials, 16 confirmed positive effect to student achievement.

Less robust yet, significant, effects of textbooks were found in an experimental study in Nicaragua. Jamison et al. (1981) sampled eighty-eight first grade classrooms from rural and urban centers, including 1098 pupils. These classrooms were those recieving radio instruction, and those serving as control group. The interventions were applied at the beginning

of the school year. Post-tests were given at the end of the same academic year. Pupils who received textbooks scored 4% higher on the mathematics post-test (one third of a standard deviation).

Teacher-student interaction. Other teaching practices, such as teacher student-interaction, have received only slight attention from researchers working in Third World nations. Few studies have examined the extent to which active learning roles are created for students in classrooms. Arrigada (1983) found that self reports by teachers on the amount of time spent explaining academic materials to students held no relationship to reading or mathematics achievement in Peruvian primary schools.

In Brazil, the study of pupil performance in rural areas found a consistent effect for the number and varieties of instructional activities reported by teachers. Brazilian rural teachers were asked whether they employed nine different instructional activities, including small work groups, dramatic reading, manual work, and storytelling. The total number of instructional activities used by teachers helped to explain pupil achievement (Armitage et al., 1986). In the African country of Botswana, Loxley (1984) conducted a study on the relationship between the frequency of classroom discussion reported by teachers and achievement on the national reading and math test. The results of the study

indicated no significant relationship between frequency of classroom discussion and student achievement.

There is strong evidence that teacher-student interaction, or actively involving students in classroom activities, contributes positively to student achievement in North America. Most process-product research, and sometimes school effect research, documented evidence for this assertion in the Handbook for Research on Teaching. For example, Doyle (1986) notes that the teacher's management task in interacting with students is primarily one of establishing and maintaining work systems for classroom groups rather than spotting and punishing misbehavior, remediating behavioral disorders, or maximizing the engagement of individual students.

Measurement and evaluation. Another dimension of classroom practice most widely believed to promote student learning in the Third World is measurement and evaluation or monitoring and giving timely feedback to students. Lockheed (1989) presented results on teaching processes that have significant relationship to student achievement. In a regression analysis of cross national data in which school, teacher background, and classroom practices were held constant, the teaching practice most associated with student learning gain is frequent monitoring and giving individual feed back. She concluded that students learn less in classes in which teachers are required to spend more time maintaining order, and less in which teachers used a published workbook

frequently than in classes where teachers are required to spend less time maintaning order and using varieties of textbooks. Frequent monitoring and feedback to students on academic progress have also been identified as an effective teaching practice in Nigeria, (Ademola, 1986).

Evidence from industrialized countries suggests a strong relationship between teacher evaluation of student performance, and frequent feedback on academic progress to student (Brookover, Beady, Flood, Schweitzer, & Wisenbaker, 1979; Walberg, 1984; Bridge, Judd & Moock, 1979).

Teacher fairness and concern. This dimension of classroom practice requires more attention by researchers working in developing nations. No studies were located that specifically dealt with this particular aspect of teaching process in the Third World. One major study that helps to understand what goes on in Third World classroom is the recent IEA classroom environmental study. In Nigeria and Thailand, researchers found that in over two-thirds of the observation segments, teachers were simply lecturing at the class. In much of the remaining time, students were sitting alone on the floor or at desks working on the assigned exercises. When teachers posed questions, the utterance was directed to the entire class, not to a specific individual student. The teachers' queries usually requested a single piece of factual information, rarely requiring complex cognition (Anderson, Ryan, & Shapiro, 1987).

Even in developed countries, about two-thirds of the talk in classrooms is teacher talk. Everyone who has gone through formal schooling can attest to this. There is no reason to believe that such talk is inappropriate or that it indicates that teachers are "oppressive" or unduly "dominant." But teachers may present themselves in a humanitarian mode when conducting their daily teaching routine. Students will perhaps learn faster in classes where teachers are perceived as being fair and showing concern towards their progress than in classes where teachers are perceived as authoritarian.

Summary

Understanding the contribution of each of these dimensions of classroom practice towards student learning is crucial in formulating policies for improved instructional quality in the Third World. It is evident from the studies cited that increasing the number of textbooks or more frequent use of varieties of instructional materials, raises student achievement. Few studies have documented the relationship between teacher-student interaction and student achievement in the Third World. But most studies from developed countries confirm the positive effect of actively engaging students in classroom instruction and improved performance of students on tests of achievement.

Both in developed and developing nations, considerable evidence exists to support the conclusion that frequent

monitoring and feedback to students enhances student achievement.

Teacher fairness and concern towards students is another critical area of classroom processes that requires considerable research attention by investigators working in Third World countries.

Policy makers in Thailand have invested substantial amount of money on in-service training to improve these dimensions of classroom practice. These variables are crucial in the study of quality of education in the Third World. The effectiveness of national in-service training of teachers in improving these dimensions of classroom practice is of major interest to policy makers, and it is also the focus of this study.

Methodological Considerations in the Analysis of Hierarchical Data

Research on school effectiveness has clearly evolved to the point where careful studies are needed to directly explore the relationships between teacher training and student achievement.

Although matters of substantive concern continue to drive the research on effective schools, the " effective schools" issue has been fueled by controversy over statistical methodology, interpretations, and data (Sirotnik and Burstein, 1985). The most important statistical issue is the use of

appropriate methods to analyze multi-level data. The argument focuses on how behavior at one level (eg., classroom, school, district, state or nation) influences behavior at a different level (eg., students), and how to estimate the multi-level effect correctly (Goldstein 1987).

Educational researchers have been concerned with multi-level data for some time, but traditional research methods have not provided adequate tools with which to analyze data arising in naturally occurring hierarchies. The paradigm of educational research has been borrowed from the traditions of agriculture and psychology in which subjects are randomly assigned to each of several treatment conditions (Raudenbush & Bryk, 1988). This assures that the expected effect of confounding factors is zero.

The nature of the educational system in some Third World nations is such that students who come from different communities with various backgrounds are grouped together in classes that are located in schools administered by school districts which are located in different parts of a nation. This grouping of students is not random, but reflects on different residential patterns of various communities. Private schools usually have different admissions policies from public schools. This creates differences in the student body, shifting the group further from a random collection of students. For example, students in a school might be regarded as a homogeneous group for purposes of one study, while the

fact is ignored that students are actually grouped into a classroom unit within the school. The problem that will confront the researchers is ignoring the fact that the group under study is one of numerous social units and that each unit may exhibit different relationships among educational process.

Furthermore, a single classroom might be studied, where researchers ignore the fact that effects estimated for that class might not generalize to other classes due to differences in classroom context.

This non-random clustering of students will tend to violate two stringent assumptions of ordinary least square (OLS), regressions analysis: 1) that each case has an equal residual variance and, 2) that the covariance between the residual of any two cases is zero (Goldstein, 1987 and Raudenbush & Bryk, 1986). It is not possible using a single level model to separate out the between-level variances. The implication is dramatic.

Such mis-specification leads to inefficient parameter estimation and the too common rejection of null hypotheses. Ordinary Least Square (OLS) regression analysis will yield efficient parameter estimates only where the clustering effects at the class and school level are small (Goldstein 1987).

This problem is obtained when individual-level data are utilized; it is complicated when data are aggregated, for in such instances the within-units (classroom or school)

variability is suppressed, and there is no way of separating out the level-one variance from the level-two variance (Riddel, 1991). It has been common practice in research on school effects to use aggregated data, such as the data already collected by state boards of education or national surveys, which are cheaper to obtain largely because they are more readily available (Goldstein, 1987). The problem of using aggregated data has been exposed for some time and by authors who have criticized a multitude of research studies using these techniques (Cronbach 1976).

For example, Cronbach (1976) noted that the majority of studies of educational effects using established methods, have generated false conclusions. At the root of the problem is the misapplication of a single-level model to a reality that is clearly hierarchical.

Until recently, most discussions of multi-level analysis have remained theoretical, bounded by the costs and computational requirements of existing analytic tools. However, the recent development of new analytic tools for analyzing multi-level data has energized the debate (Aitkin & Longford, 1986; Goldstein, 1986; Mason, Wong, & Entwiste, 1984; Raudenbush & Bryk, 1986). The development of the general EM algorithm (Dempster, Laird, & Rubin, 1977), provided a theoretically satisfactory and computational manageable approach to the estimation of covariance component in multi-level data.

Rationale For The HLM Approach

The Hierarchical Linear Model (HLM) is an extension of a regression model that allows more flexible modeling of variation within classrooms and between schools. This approach has the potential to more precisely estimate the effect of in-service teacher education programs on teaching practices and student learning, after controlling for individual student level influences and school-level compositional effects. Raudenbush & Bryk (1986) noted that:

> Since the publication of Burstein's (1980) review, several groups of investigators, working independently, have developed computational methods and techniques for data analysis which essentially resolve the long standing difficulties associated with nested, multilevel data ... These methods share a common core consisting of two principles
>
> First, they require the investigator to formulate explicit structural models for processes occurring within each level of a hierarchy. Second, they enable the investigator to specify the unique random effects of each unit and to estimate the variances and the covariances of these random effects. (p. 19-20).

The use of the regression coefficient within schools as the outcome variable in a between schools regression can be represented by two sets of equations. The following discussion draws heavily on class notes taken in Dr. Raudenbush's methodology class.

The primary application of the HLM in this study will be to model total student achievement, and sub scales of teaching quality using both students and school level variables as predictors in the regression equations. To simplify the

discussion at this juncture, total student achievement measures will be used for illustration purposes. The same logic applies to sub scales of teaching quality.

In the first equation, representing the within-school model, achievement for student i in school j, Y_{ij}, is seen as a function of measured student level variables, X_{ijk}, and random error, R_{ij}. Thus, the within school model can be represented by the following equation:

$$Y_{ij} = \beta_{j0} + \beta_{j1}X_{ij1} + \ . \ . \ . \ + \beta_{jk-1}X_{ijk-1} + R_{ij} \qquad [1]$$

where the β_{jk} regression coefficients are structural relations that occur within school j which capture the effects of the independent variable on the outcomes. Later it will be clear to see that the β_{jk} coefficients can represent the distribution of achievement in school j as a function of ses, gpa or breakfast.

Raudenbush & Bryk (1986) pointed out that "a distinctive feature of HLM is that the structural relationships are presumed to vary across units" (p.21). This variation can be estimated by formulating a between-school model. In this model, the outcomes of interest are the k structural parameters, β_{jk}. These are now seen as a function of school level variables, W_{pj}, and random error, U_{jk} :

$$\beta_{jk} = \gamma_{0k} + \gamma_{1k} W_{1j} + \ . \ . \ . \ + \gamma_{p-1k}W_{p-1j} + U_{jk} \qquad [2]$$

where the γ_{pk} coefficients represent the effects of school level characteristics, W_{pj}, on the structural relations within school. Here the W_{pj}'s represent the school level variables.

The γ's reflect the effect of these school-level variables on student level variables to the distribution of achievement.

Statistical Estimation

Hierarchical Linear Modelling helps to overcome some drawbacks of the typical slopes-as-outcome approach when Ordinary Least Square (OLS) regression methods are used to estimate the outcome variable (i.e., the β_{ik}'s). The problem in OLS is that these estimates, β_{jk}'s are measured with error. Substituting this estimate into the original equation for β_{jk} a more complex error structure emerges. Neither the γ coefficients nor the covariance structure among the errors can be appropriately estimated with conventional linear regression methods (Raudenbush & Bryk 1986). HLM provides a maximum likelihood estimation of the covariance structures among this more complex error term, as well as efficient estimates for the γ's. Raudenbush & Bryk (1986) discussed important structural properties of parameter estimates generated by HLM, and summerized the advantages of this approach.

The next chapter will concentrate on the methodology to be employed in examining the effectiveness of a national in-service training program in improving teaching competence on these dimensions of classroom practice.

CHAPTER III

DESIGN OF THE STUDY

Overview

This chapter describes the research methods and procedures utilized to examine the effectiveness of a nation's in-service training program on teaching practice and student achievement. Data from the Basic Research and Implementation in Developing Educational Systems (BRIDGES) (1988) were used in the analyses. Factor analysis was applied in a confirmation mode as a check on the empirical soundness of the scales on student achievement. Through this process, it was discovered that the five subtests of academic achievement (Mathematics, Thai language, Life experience, Work experience, and character development) measure one primary factor, which has been labeled Total Student Achievement.

All outcome variables scales (total achievement and subscales of students' perception of instructional quality), as well as other continuous scales, were then subjected to a reliability analysis both at an individual and school level. Scales found to have suitable reliability were used, along with other student and school level variables (i.e., grade point average, socioeconomic status, dialect, breakfast, and repetition), and school composition variables (i.e., textbooks, infrastructure, etc), to: (1) construct a theoretical model with no predictors which yields baseline empirical evidence about the amount of variation at each level

(between and within); (2) to construct a theoretical model which identifies relevant covariates; and (3) to construct a theoretical model for the effect of each of the key predictors, controlling for the relevant covariates. These models were tested with the aid of the hierarchical linear model (HLM), as described by Raudenbush & Bryk (1986) using the computer program of Bryk, Raudenbush, Congdon, and Seltzer (1988).

Specific procedural steps taken in this analysis include data, sub sample, measures, key predictors, covariates, analysis, HLM analysis, and descriptive statistics for School, Teacher and Student variables. Result will be presented in the next chapter. All analysis was done on the IBM mainframe computer at Michigan State University.

Data

The data used in this analysis come from the BRIDGES project. The Thai government, in collaboration with the BRIDGES project (Basic Research in Developing Educational Systems), conducted a representative national survey of primary schools in Thailand in the spring of 1988, and addressed sixth grade student achievement. The BRIDGES project sample consisted of approximately 4,000 teachers and 10,000 sixth grade students in 411 schools.

Within each six grade classroom every student was administered achievement tests in the five areas of the 1978

curriculum (Thai language, Mathematics, Life experience, Work experience, and Character development).

Sub Sample

The survey procedure utilized by the Thai government allows the investigator to link school characteristics with individual pupils in that school. However it has not been possible to link schools, teachers and pupils in a single analysis. Hence the only teacher data available for predicting student outcome have been aggregated levels of the teacher variables. Characteristic flaws within this type of analysis are well known (c.f., Aitkin & Longford, 1986). The central flaw in the present case is that, to the extent teachers within a school vary on the independent variable, the use of the mean teacher value on that variable to predict student outcomes produces uninterpretable results (Raudenbush et al. 1992). For example, consider a hypothetical situation where one teacher is rated by the student as highly effective in classroom instruction, while the other teacher is rated as not effective in classroom instruction by the same students. In analysis where the mean value of instructional quality is used to predict student achievement, it is possible to predict this student achievement with a teacher that does not exist. Since the mean value of teachers in instructional quality will represent an average teacher who is neither effective nor poor in the quality of classroom instruction.

Recently, however, the investigator discovered that about 126 of the schools in the sample have only one sixth grade classroom with one sixth grade teacher. Thus, in these schools, there is no ambiguity about the link between the teachers and the students. Moreover, these schools are of special interest to policy makers. These are small rural schools, precisely those schools that provide the greatest challenge to policy aimed at improving the quality of classroom instruction. Because of the availability of the nationally representative sample of schools, it is possible to locate these schools quite accurately in terms of the national distribution of a variety of important contextual variables.

Another advantage of using these schools is that, because the goal is to evaluate the quality of instruction during the sixth grade rather than school effects over the entire course of primary education, it is now possible to utilize grade point average at fifth grade as a covariate for predicting student achievement. This is a powerful covariate, having a surprisingly large proportion of variation between schools (35%), probably because grades in Thailand are linked in part to regional and national examinations scores.

After deletion of missing data, the obtained sample for the present study includes 103 sixth grade classrooms in 103 schools, and a total of 2111 students. This compares to the national sample which has 411 schools, 3808 teachers, and 9768 students.

Measures

The measures included indicators of students' background and achievement, of material and non material input at the school and classroom levels, and of the classroom organization and teaching practices.

Below, the outcome variables are described in detail (student achievement and sub-scales of teaching quality). The key predictors (different topics in a national in-service training program), and the key covariates, that is student, teacher, and school characteristics that must be viewed as exogenous to the effects of in-service training programs are described briefly below and in more detail in the appendix.

Student achievement. Attention is restricted to a measure of overall achievement derived from subtests developed and validated by the Office of National Primary Education Commission (ONPEC), an organization responsible for the administration of primary education in Thailand. These tests include five subtests measuring Thai language achievement, six subtests measuring aspects of mathematics achievement, five subtests measuring achievement in life experiences (social sciences and social studies), six subtests measuring achievement in character development and a subtest measuring work-oriented experience. These sub tests were all significantly positively intercorrelated, each contributing to the overall internal consistency as indicated by Cronbach's alpha, which was .89 (see appendix 2).

Instructional Quality Sub-scale.

Students were asked to estimate the frequency of a variety of teaching behavior they encounter daily in their classrooms.

Teacher-student interaction. To indicate the success of the teacher in facilitating an active student role, BRIDGES researchers asked how frequently the student spoke in front of the class, how likely the student was to ask for explanations of unclear concepts, how likely the teacher assigned students to prepare reports for presentation to the class, and how frequently the teacher allowed for expression of student opinions on the conduct of a lesson.

Measurement and evaluation. To indicate the effectiveness of teachers as leaders on monitoring students' academic progress, BRIDGES researchers asked students how frequently their teachers explained the objective of a new lesson, tested student knowledge after completion of a lesson, and provided feedback on the tests in order to clarify sources of student errors.

Teacher fairness and concern. To assess the presence of fairness and personal concern, BRIDGES researchers asked students how likely their teachers were to inquire about reasons for student absence, whether their teachers were fair to every student, whether the students would approach their teachers privately about a problem, and whether their teachers provide special help for those who are behind.

Use of instructional materials. To ascertain the success of the teacher in facilitating classroom instruction by using teaching aids, BRIDGES researchers asked students how often their teachers used textbooks or workbooks or any form of teaching materials.

Reliabilities calculated separately for students in each of the subscales of instructional quality were substantially lower, with Cronbach's alpha .64 for teacher student interaction, .50 for measurement and evaluation, and .50 for teacher fairness and concern (see appendix 2). Reliability for the use of instructional material was not computed at individual level since the scale was made up of only one item. However, the reliability for the use of instructional material was computed at the school level.

The reliability of the instructional quality subscales aggregated to the school level is very high at .94 for teacher student integration, .94 for measurement and evaluation, .94 for teacher fairness and concern, and .90 for the use of instructional materials (see appendix 2). This high reliability of the school aggregate subscale of instructional quality score has two sources. First, there is a substantial variation between classrooms on the quality of instruction. Forty-two percent of the variation is between classrooms and 58% is between students within classroom for teacher student interaction. Forty-eight percent of the variation is between classrooms and 52% is between students within classroom for

measurement and evaluation. For teacher fairness and concern, 50% of the variation is between classrooms and 50% is between students within classroom. And finally, 31% of the variation is between classrooms and 69% is between students within classrooms for the use of instructional materials. These results suggest that there is a substantial agreement among the students as "raters" of instructional quality subscales within a classroom, a result that strengthens one's confidence in the instrument. Second, enough students per school/classroom were available (20 students per classroom) to stabilize the aggregated classroom subscale of instructional quality scores. If fewer students had been sampled, the reliability of the aggregate measure of these subscales of instructional quality would have been smaller.

Key predictors

In-service training

Teachers were asked first whether, during the past three years, they had received any kind of academic in-service training. Those responding affirmatively were then asked to state, for each instance of in-service training, the topic: teacher student interaction, measurement and evaluation, curriculum and teaching content, and the use of instructional materials. In each case they were asked to name the provider (staff within the school or cluster, district or regional office, or teachers' college or university), and the duration

in days. Based on these responses, it was possible to construct for each teacher a measure of total coverage, number of topics and total duration of in-service training.

For the present study, attention is restricted to four different topics of in-service training. Exposure to in-service training will not be dealt with here. The povider of the training was excluded as a covariate, since preliminary descriptive results indicate very small variation on this variable. More than 80% of the training was provided by local organizers, (staff within school, or cluster, district or regional office).

After constructing a measure of total coverage in differing numbers of topics in the in-service training for each teacher in the subsample, each topic was coded into a dummy variable with (0) indicating that the teacher did not receive any form of in-service training on each topic, and a (1) if the teacher received in-service training on each topic. For example, In-service training in teacher student interaction, has a value of (1) if the teacher had, within the past three years received any form of academic training on teacher student interaction, and (0) if the teacher had not. In-service training on measurement and evaluation, has a value of (1) if the teacher had, within the past three years received any form of academic training on measurement and evaluation and a (0) if the teacher had not. In-service training in curriculum and teaching content, has a value of

(1) if the teacher had, within the past three years received any form of academic training on teacher fairness and concern and a (0) if the teacher had not. In-service training on the use of instructional materials, has a value of (1) if the teacher had received any form of academic training on the use of instructional materials and a (0) if the teacher had not.

Covariates

Measures of Student, Community and School Background

Student level variables include socio-economic status (ses). This was derived from measures of the father's education, the mother's education and the natural logarithm of the amount of pocket money the child typically brings to school (alpha =.73), as reported by the students. Sex or gender is an indicator variable on a value of 0 if female, and 1 if male. Linquistic background (dialect) was coded as "Central Thai vs others" as reported by the students. Student nutrition was measured by students' reports about how often they ate breakfast. An indicator variable was constructed: "Breakfast" (1 = daily, 0 = not at all). Pre-primary experience was also coded dichotomously (1 = one or more years of pre-primary experience) based on student report. Time needed to travel from home to school (in hours) was based on student report. Prior repetition is an indicator variable (0 = never repeated a grade; 1 = ever repeated a grade), and grade point average (gpa) is a seven category variable (1 =

1.00 or lower; 2 = 1.50; 3 = 1.51-2.00; 4 = 2.01-2.50; 5 = 2.51-3.00; 6 = 3.01-3.50; 7 = 3.51-4.00).

School level background variables included a scale measuring modernity of community infrastructure which consists the sum of nine items including the presence in the local community of drinking water, paved roads, irrigation, telephone service, electricity, hospital, a commercial bank, a factory, and a post office as reported by the principal. Each item was coded dichotomously; these dichotomies were then summed up, and the resulting sum had an internal consistency of alpha = .81.

Remoteness scale is the mean of three items: distance to the district capital, distance to the market and distance to the highway (all in Kilometers). There were five geographic regions: "North," "South," "Central," "Bangkok," and "Northeast." Indicator variables were constructed for the first four so that the Northeast region constituted the reference group.

School resources. These included enrollment, pupil teacher ratio, facilities and equipment, availability of textbooks, and teaching materials. Enrollment is listed under school resources because prior research (c.f., Raudenbush and Bhumirat, in press) has indicated that larger schools tend to have a critical mass of resources not typically available in smaller schools. Other measures of resources, including scales measuring the physical condition of classrooms and schools

buildings, were not included because preliminary analysis showed no indication that these were related to the outcome.

Teacher background. Teacher background variables exogenous to in-service training in different dimensions of classroom practice included pre-service education (coded 0 if less than bachelors degree and a 1 if more than a bachelors degree). Sex or gender, teacher age, experience, and school experience were not included as predictors because preliminary analysis revealed that these were unrelated to the outcomes.

Transformation. Prior to their use in the analytical models, one variable was transformed into a logarithmic metric. This variable had positively skewed distribution, and scatter plots revealed that it tended to relate non-linearly to the outcomes. Infrastructure was the only variable transformed to a logarithmic metric. Linear relations between log-transformed variable and outcomes imply that the relation between the original variable and the outcome are non linear, suggesting a diminishing effect of the predictor variable as its value increases. For example, a positive effect of log infrastructure implies that the benefit associated with an increment to infrastructure is greatest when little or no infrastructure is currently available and its smallest when a large amount of infrastructure are already available (Raudenbush and Bhumirat 1992). Such interpretations are often substantively more reasonable than are interpretations implying that the value added to an outcome by investing in a

resource is constant regardless of the current availability of that resource.

Analysis

The hierarchical structure of the data, with pupils nested within classrooms, requires a form of regression analysis that takes into account the two separate sources of variation attributable to pupils and to school/classroom. This is of substantive interest because the latter is a measure of the size of unexplained difference among schools/classrooms caused by school/classroom characteristics.

The analysis proceeded through three stages for each of the outcome variables: a) specification of models with no predictors, which yields baseline empirical evidence about the amount of variation at each level; b) specification of a model which identifies the relevant covariates; c) specification of models for each of the key predictors.

Specification of models with no predictors.

The goal here is to estimate how much of the variation in total student achievement lies between and within schools. The model can be represented by the equation:

$$Y_{ij} = \beta_{oj} + R_{ij} \text{ where } R_{ij} \sim N\ (0,\sigma^2) \qquad [3]$$

This equation states that the achievement test score for student i in school j are assumed to vary around the school

mean, β_{oj}, within school variance σ^2. The associated between school model is:

$$\beta_{oj}=\gamma_{oo} + U_{oj} \text{ where } U_{oj} \sim N(0,T_{oo}) \quad [4]$$

Equation [4] shows that the school means are assumed to be normally distributed around the grand mean, γ_{oo}, with variance T_{oo}. Substituting equation [4] into equation [3] yields the combined model:

$$Y_{ij}=\gamma_{oo}+U_{oj}+R_{ij} \quad [5]$$

Where γ_{oo} is the grand mean, U_{oj} is the between school variance and R_{ij} is the within school variance. This model could also be called a random effect model because the group effects are construed as random. The variance of Y_{ij} is:

$$Var(Y_{ij})=Var(U_{oj}+R_{ij})=T_{oo}+\sigma^2 \quad [6]$$

The HLM analysis provides maximum likelihood estimates of the within and between school variances. With this information, intraclass correlation P could be computed, using the following formula:

$$P=T_{oo}/(T_{oo}+\sigma^2) \quad [7]$$

This statistic measures the degree of dependence among observations in school, and it contains information about the estimated proportions of the total variation in achievement which is between schools (with no control for the covariates). The HLM computer software output displays a chi-square test of significance to evaluate the hypothesis of no difference in the mean achievement among schools in the sample.

The analysis also gives an estimate of the reliability of the school's sample means as estimate of their true means. This reliability increases as within school sample size and between school variation increases (Bryk & Raudenbush, 1992).

Specification of models with relevant covariates (student level). In the second stage of the HLM analysis, student-level variables are added to the within-school model. The goal is to test whether the effects of these student level variables can be assumed to be constant across schools in the sample. This "unconditional" model provides a baseline for the analysis that follows. In this analysis the expanded within school model becomes:

$$Y_{ij}=\beta_{oj}+\beta_{1j}X_{ij1}+\ .\ .\ .+\beta_{5j}X_{ij5}+R_{ij} \qquad [8]$$

where the X_{ij}'s are the student level variables that were used to classify students according to their sex, dialect, breakfast, repetition (see appendix 1 for coding). Other student level continuous variables have been centered around their means. The X_{ij}'s are actually the "dummy" values deviated from the grand mean for student i in school j. Thus β_{oj} is now an adjusted school mean (i.e., the raw school mean minus adjusted for the mean of student level independent variables). The β_{jk}'s now represent the effect of student level variables on achievement test score in school j. The between models for this analysis stage are:

$$\beta_{oj}=\gamma_{oo}+U_{oj} \qquad [9]$$

and

$$\beta_{jk}=\gamma_{k0}+U_{1jk} \qquad [10]$$

Where γ_{00} is the grand mean for the outcome measure and γ_{k0} represent the mean within school regression coefficient for variable k across schools in the sample. The result of this stage of analysis provides a test of the hypothesis of no variation across schools in their regression coefficients.

$$H_0: \text{Variance } (\beta_{jk})=0 \qquad [11]$$

If this null hypothesis is rejected, the variable should remain in the model as a random effect. If the null hypothesis is retained and the variable has an effect on the outcome, then the variable should be retained in the model as fixed effect. Here a determination has to be made on which student level variables are significantly related to the achievement measures and whether the relationship varies across schools.

Specification for Key Predictors

Policy makers and educational researchers will want to know whether teacher participation in an in-service training program has any demonstrable effect on perceived instructional quality and student learning. But, before this final stage of HLM analysis is reached, it is necessary to know whether the composition, the location, and the resources of the school in which a student is enrolled will affect his or her achievement beyond the influences already exerted by the student background variables. To examine this possibility, the between

school model can be extended further to include variables that represent the location and the school resources.

In this analysis up to four variables are available for modelling the " base" (i.e., mean student achievement), within each school. The between school model for this analysis is therefore

$$\beta_{oj}=\gamma_{o1}+\gamma_{j1}C_{j1}+\ldots+\gamma_5 C_{j5}+U_{ojk} \qquad [12]$$

were γ_{jk} is the effect of school contextual variable K on "base" level student achievement. C_{jk} is contextual variable K for school j (i.e., textbook availability in school j).

To model the effect of each of the key predictors which is the main focus of this study, student background and school contextual variables which have no effect on the outcome are purged out of the model. Now the between-school model is elaborated to include the effects of the key predictors.

$$\beta_{oj}=\gamma_{oo}+\gamma_{o1}C_{j1}+\gamma_{o2}W_{1j}+\ldots+\gamma_{o6}W_{5j}+U_{oj} \qquad [13]$$

Equation [13], for example indicates that four school level contextual variable and one key predictor variable for a total of five predictor variables are being used in the model. To avoid problems associated with multicollinearity in regression equations, it is necessary to run separate models for each of the key predictors.

In this analysis, students were viewed as nested within schools/classrooms. Because the focus is on schools with one classroom per sixth grade and one teacher, school and classroom variation are indistinguishable. Predictor variables

at both levels were rescaled to have zero means so that the intercept would have meaning. Hence, the intercept for each model may be interpreted as the expected score of an otherwise "typical" student having a zero value on all the indicator variables and attending a classroom having zero value on classroom level indicators.

Missing data were treated with pairwise deletion at the student level and listwise deletion at the classroom level. The final sample therefore involved 103 classrooms and 103 schools. The distribution of the predictor variables in the complete sample (n=125) and the reduced sample were compared and found to be highly similar. The results of this analysis will be presented in the next chapter.

HLM Analysis

The central focus of this study involves the use of Hierarchical Linear Modeling to examine the effects of in-service education of teachers on students perception of instructional quality (sub scales) and student achievement (total achievement). Here the three major sections of the research questions (total effect of training, effect of classroom practices on student achievement, and comparative effectiveness of training on dimensions of classroom practice are explored in theoretical terms. In this section the four stages of HLM analysis discussed earlier are illustrated in terms of actual data from Thailand. In an effort to simplify

the presentation of the results for this analysis in all the models, student achievement is the first step, and then sub scales of instructional quality is the second step. The models will be discussed theoretically in this section and the results and technical discussions will follow in chapter four.

Apportioning parameter variance for total achievement.

Using the simple model represented by Equation 3 and 4 of this chapter, several important pieces of information could be derived regarding total achievement and subscales of teaching quality outcome measures. First using the formula for intraclass correlation, an estimate for the amount of variation that lies between schools on student achievement and each of the subscales of teaching quality (teacher student interaction, measurement and evaluation, teacher fairness and concern, and the use of instructional material) could be found. Both T_{∞} and σ^2 are provided by the HLM output. The Intraclass correlation, P, results for student achievement and each of the subscales of teaching quality is reported in chapter 4 on between school variation in all the outcome measures.

Another piece of useful information derived from this analysis involves a test of the significance of the between school variations on both total achievement and each of the subscale of teaching quality. The null hypothesis is represented symbolically by H_o:Too=o. The chi-square test provided by HLM indicates that the null hypothesis can be

rejected, and infers that there are significant differences among schools in their mean total achievement and the mean of each of the four subscales of teaching quality. For total student achievement (χ^2=3785, df=103, p<.0005) (see column 1 table 2,3,4,5,6,7,8,9,); for teacher student interaction (χ^2=4211, df=103, p<.0005), (see column 1 table 10); for measurement and evaluation (χ^2=4758, df=103, p<.0005), (see column 1 table 11); for teacher fairness and concern (χ^2=4729, df=103, p<.0005) (see column 1 table 12); and for the use of instructional materials (χ^2=3187, df=103, p=.0005), (see column 1 table 13).

Finally, HLM provides an estimate of the reliability of the school sample means as estimators of their true means. The point estimates of these reliabilities are given in the appendix for both total achievement and each of the subscales of teaching quality.

Specification of models with relevant covariates (student level)

At this stage of analysis only student level-variables are added to the-within school model as depicted in Equation 8. Some student-level variables were dropped from the model when they failed to approach significance and some were retained even when they did not reach significant level for comparison purposes. This leaves the regression model described in figure 1

Unconditional regression models for total achievement

Within school model:

(TOTAL STUDENT ACHIEVEMENT) $Y_{ij} = \beta_{oj} + \beta_{1j}(GPA) + \beta_{2j}(DIALECT) + \beta_{3j}(BREAKFAST) + \beta_{4j}(REPETITION) + R_{ij}$

Between-school model:

$\beta_{oj} = \gamma_{oo} + \gamma_{o1}(NORTH) + \gamma_{o2}(TEXTBOOKS) + \gamma_{o3}(REMOTENESS) + \gamma_{o4}(TEACHER\ STUDENT\ INTERACTION) + U_{oj}$

Combined Model (Teacher student interaction):

(TOTAL STUDENT ACHIEVEMENT) $Y_{ij} = \gamma_{oo} + \gamma_{o1}(NORTH) + \gamma_{o2}(TEXTBOOKS) + \gamma_{o4}(TEACHER\ STUDENT\ INTERACTION) + \gamma_{10}(GPA) + \gamma_{20}(DIALECT) + \gamma_{30}(BREAKFAST) + \gamma_{40}(REPETITION) + U_{oj} + R_{ij}$ (see table 2).

Note. γ_{oo} is the grand mean of the outcome Y_{ij}, and the γ_{jk}'s are the mean within-school regression coefficients for variable K across all schools in the sample (103) (K=1,2,3,4). The result of this combined model is presented in the next chapter. The γ_{ok}'s are the GAMMA's following Base in Table 2,3,4,5,6,7,8,9.

The equations are similar for the other outcomes.

Unconditional regression models for sub scale of teaching quality

Figure 2

Within school model:

(MEAN SUBSCALE OF TEACHING QUALITY) $Y_{ij} = \beta_{oj} + \beta_{1j}(GPA) + \beta_{2j}(SES) + \beta_{3j}(BREAKFAST) + \beta_{4j}(PREPRIMARY\ ED) + R_{ij}$

Between-school model:

$\beta_{oj} = \gamma_{oo} + \gamma_{o1}(EDUCATION) + \gamma_{o2}(FACILITIES) + \gamma_{o3}(TEACHER\ STUDENT\ INTERACTION) + U_{oj}$

Combined Model (Teacher student interaction):

(MEAN TEACHER STUDENT INTERACTION) $Y_{ij} = \gamma_{oo} + \gamma_{o1}(EDUCATION) + \gamma_{o2}(FACILITIES) + \gamma_{o3}(TEACHER\ STUDENT\ INTERACTION) + \gamma_{1o}(GPA) + \gamma_{2o}(SES) + \gamma_{3o}(BREAKFAST) + \gamma_{4o}(PREPRIMARY\ ED) + U_{oj} + R_{ij}$ (see table 10).

Unconditional regression models for subscale of teaching quality

$Yij=\beta_{jo}+\beta_{j1}(GPA)+\beta_{j2}(SES)+\beta_{j3}(BREAKFAST)+\beta_{j4}(PREPRIMARY\ EDUCATION)+Rij$

Between-School Model:

INTERCEPT	$\beta_{oj}=\gamma_{oo}+U_{oj}$
GPA	$\beta_{1j}=\gamma_{1o}+U_{1j}$
SES	$\beta_{2j}=\gamma_{2o}+U_{2j}$
BREAKFAST	$\beta_{3j}=\gamma_{3o}+U_{3j}$
PREPRIMARY EDUCATION	$\beta_{4j}=\gamma_{4o}+U_{4j}$

NOTE. γ_{oo} is the grand mean of the outcome measure Y_{ij} (each of the subscale of teaching quality). The γ_{jk}'s are the mean within-school regression coefficients for variable K across all schools in the sample (103), (K=1,2,3,4). The result of this unconditional model is presented with the full model in chapter four (Tables 10,11,12,13).

Testing for school level covariates on achievement and subscale of teaching quality

This is the third stage of HLM analysis, the between school model of mean total achievement and the mean of each of the subscales of instructional quality are extended to account for the effects of school level variables. These school level variables represent the school level covariates which have

been described earlier and the full process involved will be explained in detail in chapter four.

Assessing the effect of key predictors on achievement and subscale of teaching quality

It is now appropriate to address the three research questions, especially the total effect of training on achievement, the effect of classroom practices on achievement, and the comparative effect of training on dimensions of classroom practices.

Figure 3.

Final explanatory model for total achievement

Within-School Model:

$$Y_{ij}=\beta_{jo}+\beta_{j1}(GPA)+\beta_{j2}(DIALECT)+\beta_{j3}(BREAKFAST)+\beta_{j4}(REPETITION)+R_{ij}$$

Between-School Model:

$$\text{INTERCEPT}\quad \beta_{jo}=\gamma_{oo}+\gamma_{o1}(NORTH)+\gamma_{o2}(TEXTBOOKS)+\gamma_{o3}\quad (REMOTENESS)+\gamma_{o4}$$

$$(\text{TEACHER-STUDENT INTERACTION})+U_{jo}\qquad\qquad (\text{Table 2})$$

Note γ_{oo} is the grand mean of the outcome (total achievement), Y_{ij}, and the γ_{jk}'s are the mean within-school regression coefficients for variable K across all schools in the sample (K=,2,3,4,5).

The other key predictors are substituted for teacher-student interaction in the model.

Figure 4.

Final exploratory model for the subscale of teaching quality

Within School Model:

$$Y_{ij}=\beta_{jo}+\beta_{j1}(GPA)+\beta_{j2}(SES)+\beta_{j3}(BREAKFAST)+\beta_{j4}(PREPRIMARY\,EDUCATION)+R_{ij}$$

Between School Model:

$$INTERCEPT\quad \beta_{jo}=\gamma_{oo}+\gamma_{o1}(NORTH)+\gamma_{02}(PRESERVICE\quad EDUCATION)+\gamma_{o3}$$

$$(EQUIPMENT\ \&\ FACILITIES)+\gamma_{o4}(INSERVICE\ TRAINING\ TEACHER\ STUDENT$$

$$INTERACTION)+U_{jo}$$

(Table 10)

Descriptive Statistics for School, Teacher and Student Variables

Table 1: Descriptive statistics for National Sample and Sub-Sample.

Variable	National Sample m	National Sample sd	Sub-sample m	Sub-sample sd
(a) School Level				
Infrastructure	3.09	2.42	1.89	2.23
North	0.19	0.40	0.20	0.40
Central	0.18	0.39	0.16	0.37
South	0.13	0.12	0.07	0.25
Bangkok	0.15	0.13	0.08	0.27
Northeast	0.33	0.22	0.49	0.50
Pupil Teacher Ratio	21.32	6.78	20.49	7.93
Enrollment	425.23	515.91	164.77	85.15
Remoteness	7.77	2.70	8.49	2.23
Principal Age	44.78	8.75	41.13	8.50
Equipment and Facilities	0.00	0.19	0.00	0.42
Textbooks	26.26	8.41	20.49	7.93
Teaching Materials	1.94	1.60	1.47	1.33
(b) Teacher Level				
Sex	0.37	0.48	0.40	0.49

Table 1 (continued)

Variable	National Sample m	National Sample sd	Sub-sample m	Sub-sample sd
Education	0.59	0.49	0.73	0.44
Inservice Training	1.61	0.89	1.84	0.73
Inservice Training (Teacher Student Interaction)	.60	.50	.65	.48
Inservice Training (Use of Instructional Material)	.50	.50	.43	.50
Inservice Training (Curriculum &Teaching Content)	.56	.50	.61	.50
Inservice Training (Measurement and Evaluation)	.30	.50	.30	.50
Utilization of Training	.40	.50	.40	.50
Internal Supervision	9.16	11.19	7.57	7.51
(c) Student Level				
Dialect	0.49	0.50	0.32	0.47
SES	0.00	0.69	-0.30	0.42
Breakfast	0.83	0.37	0.90	0.30
Time to School	18.77	15.08	16.48	12.60
Pre-primary education	0.49	0.50	0.38	0.48
GPA	4.81	1.39	4.49	1.41
Sex	0.50	0.50	0.49	0.50
Repetition	0.15	0.36	0.20	0.40
Total Achievement	-0.12	11.83	-2.77	8.88
Instructional Quality (Sub-scale Teacher Student Interaction)	12.99	2.73	12.65	2.66
Instructional Quality (Sub-scale Use of Instructional Materials)	3.01	.79	2.94	.78
Instructional Quality (Sub-scale Measurement and Evaluation)	10.84	2.19	10.65	2.18
Instructional Quality (Sub-scale Teacher Fairness and Concern)	12.41	2.59	12.42	2.56

Comparison of the National Sample and the Subsample

Descriptive statistics for selected variables are provided in Table 1. The table shows a comparison between the obtained sample of small rural primary schools with a single sixth grade classroom and a sole teacher and a nationally representative sample of Thai primary schools.

Table 1 indicates that in many ways, the schools in the sub-sample are disadvantaged relative to the schools in the population as a whole. They are substantially smaller and more remote, the infrastructure of the communities in which they are located is less modern, and fewer textbooks, and teaching materials are available. They are likely to be located in the Northeast, the most impoverished section of Thailand. Their students are of lower socio-economic status, are less likely to speak the central Thai dialect, less likely to have experienced pre-primary education, more likely to have repeated a grade, are likely to spend less time to get to school, considering the fact that they do not have to encounter heavy traffic on their way to school, and they exhibit lower grade point average. Their principals are younger, which reflects the tendency of principals to move to larger schools as they are promoted. Their sixth grade achievement levels are lower. Such disadvantaged schools are targets for improvement efforts according to Thailand's most recent national educational development plan.

Their teachers, however, are not disadvantaged. They have more education on the average than their counterparts in the nation as a whole, reflecting the policy of the Thai government to place new graduates of teacher training colleges in relatively remote, rural schools. They have, on the average, experienced more in-service training on teacher student interaction and curriculum and teaching content. They are likely to experience less in-service training on the use of instructional material, and are also less likely to have received regular internal supervision.

CHAPTER IV

ANALYSIS AND RESULTS

Outline

The results of the study are presented in three sections. The first section contains the results of the total effect of in-service training on student achievement. The second section will address the effect of classroom practices on student achievement (here subscales of instructional quality will be used to predict student learning). The last section presents the results of the effectiveness of training on dimensions of classroom practices.

Section 1

Total effect of training on student achievement

The objective of this section was to test the effect of each of the four topics covered during the in-service training sessions. These four topics include, teacher-student interaction, measurement and evaluation, curriculum and teaching content, and the use of instructional materials. For each topic, one HLM model was presented to facilitate simplicity. These four models addressed the first four research questions outlined in chapter one.

Results

Table 2: Inservice Training as a Predictor of Total Achievement.

Predictors	Base-only Model			Inservice Training Exposure to Teacher Student Interaction		
	Coeff	SE	t	Coeff	SE	t
(a) School/Classroom Level						
Intercept	-24.45	6.25	-3.92	-35.49	10.21	-3.48
North				50.59	12.97	3.90
Textbooks				8.67	2.75	3.16
Remoteness				7.40	2.64	2.80
Infrastructure				29.20	15.11	1.93
Inservice Training (Exposure to Teacher Student Interaction)				-1.13	10.87	-0.10
(b) Student Level						
GPA				25.68	1.14	22.44
Dialect				12.92	7.98	1.62
Breakfast				12.04	4.49	2.68
Repetition				-23.83	3.74	-6.37
Variances						
Between class	3785			2424		
% Explained	0.0			36.0		
Within class	4099			3079		
% Explained	0.0			25.0		

Note This is the result of the final explanatory model for in-service training in teacher-student interaction.

Table 3: Inservice Training as a Predictor of Total Achievement.

Predictor	Base-only Model			Inservice Training Exposure to Measurement and Evaluation		
	Coeff	SE	t	Coeff	SE	t
(a) School/classroom Level						
Intercept	-24.45	6.25	-3.92	-40.17	8.25	-4.87
North				50.65	12.88	3.93
Textbooks				8.73	2.71	3.22
Remoteness				7.35	3.62	3.22
Infrastructure				29.75	14.97	1.99
Inservice Training (Measurement and Evaluation)				11.42	10.92	1.05
(b) Student Level						
GPA				25.68	1.44	22.44
Dialect				13.82	8.01	1.73
Breakfast				12.06	4.49	2.69
Repetition				-23.75	3.74	-6.34
Variances						
Between class	3785			2398		
% Explained	0.0			36.7		
Within class	4099			3079		
% Explained	0.0			24.9		

Note This is the result of the final explanatory model for in-service training in measurement and evaluation.

Table 4: Inservice Training as a Predictor of Total Achievement.

Predictors	Base-only Model			Inservice Training Exposure to Curriculum Teaching content		
	Coeff	SE	t	Coeff	SE	t
(a) School/Classroom Level						
Intercept	-24.45	6.25	-3.92	-39.46	9.88	-3.99
North				50.32	12.93	3.89
Textbooks				8.75	2.72	3.22
Remoteness				7.36	2.64	2.79
Infrastructure				28.85	15.01	1.92
Inservice Training (Exposure to Curriculum and Teaching Content)				5.11	10.41	0.62
(b) Student Level						
GPA				25.68	1.14	22.44
dialect				13.30	8.02	1.66
Breakfast				12.01	4.49	2.68
Repetition				-23.81	3.74	-6.36
Variances						
Between class	3785			2437		
% Explained	0.0			36.6		
Within class	4099			3069		
% Explained	0.0			25.1		

Note This is the result of the final explanatory model for in-service training in curriculum and teaching content.

Table 5: Inservice Training as a Predictor of Total Achievement.

Predictors	Base-only Model			Inservice Training Exposure to the Use Instructional Material		
	Coeff	SE	t	Coeff	SE	t
(a) School/Classroom Level						
Intercept	-24.45	6.25	-3.92	- 33.62	8.84	-3.80
North				50.31	12.92	3.90
Textbooks				8.83	2.72	3.24
Remoteness				7.17	2.66	2.69
Infrastructure				28.13	15.09	1.87
Inservice Training (Exposure to Use of Instructional Materials)				5.55	10.46	-.53
(b) Student Level						
GPA				25.68	1.14	22.44
dialect				12.18	8.04	1.51
Breakfast				12.08	4.49	2.69
Repetition				-23.82	3.74	-6.36
Variances						
Between class	3785			2412		
% Explained	0.0			36.3		
Within class	4099			3080		
% Explained	0.0			25.1		

Note This is the result of the final explanatory model for in-service training in the use of instructional materials.

Predictors of Achievement

Base-only model. A remarkable proportion of the variation in total student achievement was found to lie between classrooms and between students in the same classroom. As Tables 2 through 4 show, the between classroom variance with no predictor in the model was estimated to be 3785/(3785+4099)=.48 or 48% of the variation was between classrooms, implying that classrooms vary substantially, either in their effectiveness or in their contextual or compositional characteristics.

Covariates and key predictors. Specification of covariates proceeded through four stages. First, the "best set" of student covariates was identified (gpa, dialect, breakfast, and repetition). Second, school composition variables were added to the model. Those significantly related to the outcome at a nominal ten percent level of significance were retained. Next, school resource indicators were included in the model, with significant predictors retained. At each stage, residuals were regressed on predictors that had previously been dropped from the model to guarantee that the initial decision to drop a predictor remained correct in light of subsequent additions to the model.

Once a set of useful covariates had been identified, key predictors were added. Indicators of in-service training on teacher student interaction, measurement and evaluation, curriculum and teaching content, and the use of instructional

materials, were added singly in each of the four models with the same set of covariates retained for comparison purposes. Non-significant key predictors were retained so that the first four research questions could be answered.

The rather painstaking procedure described above was necessary because the number of potential predictors was quite large relative to the number of classrooms. The goal was to achieve a theoretically plausible set of steps towards the identification of a parsimonious model. A large nominal significance level worked against Type II errors in this study having a restricted degree of freedom between classrooms.

Student-level predictors. Results for the final model for each of the four topics on in-service training appear in Tables 2 to 5. Among the student level covariates, GPA was strongly and positively related to achievement, β=25.68, t=22.44; prior repetition was negatively related to achievement, β=-23.83, t=6.37. Positive effects were found for speaking central Thai, β=12.92; t=1.62, and eating breakfast daily, β=12.04; t=2.68. Once these variables were controlled, no significant effects of sex, ses or pre-primary educational experience were present. Therefore, these variables were dropped from the model. Each effect is adjusted for other effects in the model, and the importance of variables reflecting linguistic, social, nutritional, and pre-primary educational background are clearly underestimated, given that they are exogenous to gpa and repetition.

Standardized effect sizes for the dummy variables are easily computed. Given that the outcome has a standard deviation of 89, regression coefficient can be divided by 89 to obtain standardized mean differences between the group labelled one and the group labelled zero. The resulting standardized mean differences are -.27 for repetition, .15 for central Thai, and .14 for eating breakfast daily. Standardized regression coefficients for the continuous predictors may be computed by multiplying each raw regression coefficient by the ratio of standard deviation of the predictor to the standard deviation of the outcome (Table 2 to 5). Results are .41 for gpa. Note that the residual within school variance is estimated to be 3072. Given the estimated within-school variance of 4099 based on the base-only model (i.e., the model having no predictors, column 1 of tables 2 through 4), it is possible to infer that the student-level predictors accounts for (4099-3079)/4099=.25 or about 25% of the within school variation.

School-level predictors

Only three school level predictors remained in the final model (tables 2 through 4). It should be emphasized that other school-level variables have significant zero-order relations with the outcome, and some undoubtedly operate indirectly through those predictors in this model. For example, school-level background and resource variables may predict

student gpa and internal supervision. The present focus, however, is on the effects of policy-relevant predictors rather than identifying all pathways to achievement.

Results indicate that the community infrastructure (log metric) is significantly and positively related to achievement, β=29.20, t=1.93, as do textbooks, β=8.67, t=3.16, and remoteness, β=7.40, t=2.80, and the largest effect for North, β=50.59, t=3.90.

Nothing approaching any significant effect was found for each of the four topics, for example, Teacher student interaction, β=-1.13, t=-.10, (Table 2), Measurement and evaluation, β=11.42, t=1.05, (Table 3), Curriculum and Teaching content, β=5.11, t=0.62, (Table 4), and the Use of instructional materials, β=-5.55, t=-.53, (Table 5).

Standardized effect sizes for the continuously-distributed predictors may be obtained by multiplying the raw regression coefficient by the ratio of the predictor to the outcome standard deviation (Raudenbush et al., 1991). The investigator has a choice of standard deviation between the school-level standard deviation of 62 (square root of 3785), or the overall standard deviation of 89. It is preferable to use the school-level standard deviation because school-level variables can predict only school-level variations. The standardized regression coefficient for the infrastructure (.20), remoteness (.27), and textbooks (.11). The residual variance between schools is 2424 (column two of table 2), 2393

(column two of table 3), 2437 (column two of table 4) and 2412 (column two of table 5). This is a 36% reduction from the between school variance of 3785 associated with the base-only model (column one of table 2 to 5).

Section 2

Effect of classroom practices on student achievement.

To investigate whether the effect of in-service training could plausibly be viewed as working through the dimensions of classroom practices, a model would have been specified in which both training and dimensions of classroom practices were included as predictors, along with the necessary covariates.

Recall that the results in section 1 indicate no significant effect of any of the in-service training topics on student achievement. Thus, it is necessary to exclude training in these models and use only dimensions of classroom practices as the key predictors in the models. Here the mean level of sub-scales of instructional quality was used to predict student achievement. For each sub-scale, there is one model with the necessary covariates controlled, to a total of four models.

These four models provided the answers to the second research questions outlined in chapter one.

Results

Table 6: Sub-scale of Teaching quality as a Predictor of Total Achievement.

Predictor	Base-only Model			Sub-scale of Teaching Quality (Teacher-Student Interaction)		
	Coeff	SE	t	Coeff	SE	t
(a) School/Classroom Level						
Intercept	-24.45	6.25	-3.92	-35.37	7.31	-4.84
North				47.12	12.92	3.65
Textbooks				8.26	2.69	3.07
Remoteness				7.00	2.61	2.68
Infrastructure				26.33	14.91	1.77
Sub-scale of Teaching Quality (teacher student interaction)				0.13	0.08	1.64
(b) Student Level						
GPA				25.67	1.14	22.44
Dialect				12.20	7.93	1.54
Breakfast				11.98	4.49	2.67
Repetition				- 23.82	3.74	-6.37
Variances						
Between class	3785			2349		
% Explained	0.0			38.0		
Within class	4099			3080		
% Explained	0.0			25.0		

Note This is the result of the final explanatory model for subscale of teaching quality (teacher-student interaction).

Table 7: Sub-scale of Teaching Quality as a Predictor of Total Achievement.

Predictor	Base-only Model			Sub-scale of Teaching Quality (Measurement and Evaluation)		
	Coeff	SE	t	Coeff	SE	t
(a) School/classroom Level						
Intercept	-24.45	6.25	-3.92	-35.46	7.31	-4.85
North				47.48	12.89	3.69
Textbooks				8.66	2.68	3.23
Remoteness				7.47	2.60	2.87
Infrastructure				30.26	14.83	2.04
Sub-scale (Measurement and Evaluation)				0.19	0.07	1.66
(b) Student Level						
GPA				25.70	1.14	22.47
Dialect				12.23	7.93	1.54
Breakfast				11.96	4.49	2.67
Repetition				-23.75	3.74	-6.33
Variances						
Between class	3785			2398		
% Explained	0.0			37.8		
Within class	4099			3079		
% Explained	0.0			25.1		

Note This is the result of the final explanatory model for subscale of teaching quality (measurement and evaluation).

Table 8: Sub-scale of Teaching Quality as a Predictor of Total Achievement.

Predictor	Base-only Model			Sub-scale of Teaching Quality (Teacher Fairness and Concern)		
	Coeff	SE	t	Coeff	SE	t
(a) School/Classroom Level						
Intercept	-24.45	6.25	-3.92	-36.46	7.35	-4.96
North				51.39	12.93	3.98
Textbooks				8.89	2.73	3.27
Remoteness				7.39	2.63	2.82
Infrastructure				28.33	14.98	1.90
Sub-scale of Teaching Quality (teacher fairness and concern)				0.06	0.07	0.90
(b) Student Level						
GPA				25.70	1.14	22.46
Dialect				12.89	7.96	1.62
Breakfast				12.02	4.49	2.70
Repetition				- 23.78	3.74	-6.35
Variances						
Between class	3785			2403		
% Explained	0.0			36.5		
Within class	4099			3080		
% Explained	0.0			24.9		

Note This is the result of the final explanatory model for subscale of teaching quality (teacher fairness and concern).

Table 9: Sub-scale of Teaching Quality as a Predictor of Total Achievement.

Predictor	Base-only Model			Sub-scale of Teaching Quality (Use of Instructional Materials)		
	Coeff	SE	t	Coeff	SE	t
(a) School/Classroom Level						
Intercept	-25.45	6.25	-3.92	- 36.32	7.36	-4.93
North				50.83	12.95	3.93
Textbooks				8.71	2.72	3.21
Remoteness				7.44	2.64	2.82
Infrastructure				29.78	15.10	1.97
Sub-scale of Teaching Quality (use of instructional materials				0.04	0.09	0.44
(b) Student Level						
GPA				25.68	1.14	22.43
Dialect				12.81	7.97	1.61
Breakfast				12.03	4.49	2.68
Repetition				-23.83	3.74	-6.37
Variances						
Between class	3785			2418		
% Explained	0.0			36.1		
Within class	4099			3079		
% Explained	0.0			24.9		

Note This is the result of the final explanatory model for subscale of teaching quality (use of instructional materials).

Nothing approaching any significant effects were found for sub-scales of teaching quality in teacher fairness and concern, β=0.06, t=0.09 (Table 8), and the use of instructional materials, β=0.04, t=0.44 (Table 9). Modest positive effects were found for teacher student interaction,

β=0.13, t=1.64 (Table 6), and measurement and evaluation, β=0.19, t=1.66 (Table 7). This is a 38% reduction from the between-school variance of 3785 associated with the Base-only model (column one of table 6), 36% reduction for Table 9, and 37% for Table 8.

Section 3

Comparative effectiveness of training

If in-service training were to influence student achievement, it ought to do so through the mechanism of improving dimensions of classroom practices, which in turn boosts student learning. The implication is that an effective in-service training program should at least meet its objective. Otherwise investment in such programs may become counter productive. How effective are the national in-service training programs on four dimensions of classroom practice in Thailand in meeting their objectives? This section will deal with the last part of the research questions outlined in chapter 1.

To accomplish this, the analysis again proceeded through three stages. A "base-only" model made possible the partition of the variation within and between schools on each of the dimensions of classroom practices (teacher student interaction, measurement and evaluation, teacher fairness and concern, and the use of instructional materials).

Next the "best" set of covariates for predicting each of the four dimensions of classroom practices were identified.

Results

Table 10: Inservice Training as a predictor of Sub-scale of Teaching Quality (Teacher Student Interaction).

Predictor	Base-only Model			Inservice Training (Exposure to Teacher Student Interaction)		
	Coeff	SE	t	Coeff	SE	t
(a) School/Classroom Level						
Intercept	-8.45	6.58	-1.28	-34.29	17.01	-2.07
Education				13.87	14.67	0.95
Equipment and Facilities				37.46	15.55	2.47
Inservice Training (Teacher-student interaction)				11.35	13.75	0.83
(b) Student Level						
GPA				7.96	1.34	5.91
SES				6.38	3.81	1.67
Breakfast				8.82	5.32	1.66
Pre-primary eduction				11.53	4.64	2.48
Variances						
Between class	4211			3815		
% Explained	0.0			53%		
Within	4456			4342		
% Explained	0.0			51%		

Note This is the result of the final explanatory model for in-service training in teacher-student interaction predicting subscale of teaching quality (teacher-student interaction).

Table 11: Inservice Training as a Predictor of Sub-scale of Teaching Quality (Measurement and Evaluation).

Predictor	Base-only Model			Inservice Training (Exposure to Measurement and Evaluation)		
	Coeff	SE	t	Coeff	SE	t
(a) School/Classroom Level						
Intercept	-9.55	6.99	-1.37	-42.56	16.68	-2.55
Education				28.12	16.24	1.73
Equipment and Facilities				12.10	16.66	0.73
Inservice Training (Measurement and evaluation)				25.24	15.44	1.63
(b)						
GPA				4.05	1.42	2.84
SES				-2.09	4.02	-0.52
Breakfast				7.93	5.60	1.42
Pre-primary education				-5.55	4.89	-1.14
Variances						
Between class	4758			4618		
% Explained	0.0			50%		
Within	4826			4810		
% Explained	0.0			50%		

Note This is the result of the final explanatory model for in-service training in measurement and evaluation predicting subscale of teaching quality (measurement and evaluation).

Table 12: Inservice Training as a Predictor of Sub-scale of Teaching Quality (Teacher Fairness and Concern).

Predictor	Base-only Model			Inservice Training (Exposure to Curriculum and Teaching Content)		
	Coeff	SE	t	Coeff	SE	t
(a) School/Classroom Level						
Intercept	1.42	6.97	0.20	-12.56	17.72	-0.71
Education				14.11	16.28	0.86
Equipment and Facilities				4.37	17.36	0.25
Inservice Training (Curriculum and Teaching Content)				1.09	14.81	0.07
(b)						
GPA				3.81	1.44	2.65
SES				0.34	4.10	0.08
Breakfast				0.76	5.71	0.13
Pre-primary education				9.26	4.99	1.85
Variances						
Between class	4729			4617		
% Explained	0.0			51%		
Within class	5030			5002		
% Explained	0.0			50%		

Note This is the result of the final explanatory model for in-service training in curriculum and teaching content predicting subscale of teaching quality (teacher fairness and concern).

Table 13: Inservice Training as a Predictor of Sub-scale of Teaching Quality (Use of Instructional Materials)

Predictor	Base-only Model			Inservice Training (Exposure to the Use of Instructional Materials)		
	Coeff	SE	t	Coeff	SE	t
(a) School/Classroom Level						
Intercept	-2.62	5.74	-0.46	-2.43	14.14	-0.17
Education				8.63	13.24	0.65
Equipment and Facilities				2.25	14.10	0.16
Inservice Training (Use of instructional materials)				-15.05	11.85	-1.27
(b)						
GPA				3.15	1.51	2.07
SES				8.35	4.24	1.98
Breakfast				3.69	5.93	0.62
Pre-primary education				0.27	5.09	0.05
Variances						
Between class	3090			3187		
% Explained	0.0			49%		
Within class	5429			5409		
% Explained	0.0			50%		

Note This is the result of the final explanatory model for in-service training in the use of instructional materials predicting subscale of teaching (use of instructional material).

Base-only model. The results indicate that 42% of the variation in teacher student interaction (Table 10), 47% variation in both measurement and evaluation and teacher fairness and concern (Table 11 and 12), and 30% variation for the use of instructional material were between schools (Table

13), implying that substantial agreement existed among students sharing a classroom regarding the quality of instruction. Based on this level of agreement, and a sample size of about 20 students per classroom, the reliability of the classroom mean on each of these dimensions of instructional quality was .94 for teacher student interaction, .94 for measurement and evaluation, .94 for teacher fairness and concern, and .90 for the use of instructional materials. (see Appendix 2-1).

Because the classroom mean levels of each of the four dimensions of classroom practices were highly reliable and because the means were utilized as predictors in the analysis of total achievement in the second section of this chapter, each mean of each dimension of classroom practice was used as the outcome in the subsequent analysis. Hence, the analysis was no longer multi-level, but rather was a single-level analysis at the classroom level.

Specifications of covariates. A series of regressions was estimated. First, school background indicators were used to predict each of these dimensions of classroom practices and the non significant predictors at a nominal 10 percent level were dropped from the model. It was necessary to keep ses and breakfast in the model even when they fail to approach significance. This decision was taken because research from Thailand indicates a strong relationship between these variables and student achievement. Next, school resource

variables were added, with non significant ones dropped and residuals again regressed on previously excluded predictors.

Specification of in-service training . Next predictors indicating exposure to each topic in the in-service training program were added singularly to the models. The results appear in Tables 10 through 13. The model is remarkably parsimonious, including three predictors at the school level: preservice education, equipment and facilities, and in-service training. At the student level, predictors include, gpa, ses, breakfast, and pre-primary eduction. Although some of the predictors reached significance level in some models and failed to reach significance in other models, it was necessary to keep all the predictors in the model for comparative purposes. Note that gpa reached a significant level in all the models.

For comparative purposes, Table 11 presents the results of in-service training in measurement and evaluation. Note that the effect of in-service training in measurement and evaluation reached a modest significance at β=25.24, t=1.63, and preservice education β=28.12, t=1.73. Nothing approaching any significant level was found for in-service training in teacher-student interaction (Table 10), curriculum and teaching content (Table 12), and the use of instructional material (Table 13).

CHAPTER V

SUMMARY AND CONCLUSIONS

Introduction

The purpose of the analysis conducted during the course of this investigation was to answer a series of questions regarding the effectiveness of a nation's in-service training programs for incumbent teachers: 1) What are the total effects of training on student achievement? 2) What are the effects of classroom practices on student achievement? 3) What is the comparative effectiveness of the four different topics covered in the training program? This chapter will focus on the findings of this study and the presentation of some caveats concerning their interpretations.

Summary

Total effect of training. The first analysis in this study examined the extent to which each of the four topics covered in a national in-service teacher training program affected the overall academic achievement of students. For this analysis, four models of HLM were created to answer these four questions. No significant effect of the training was found for each of the four topics on student achievement for the 103 teachers in 103 schools in the sample in Thailand.

Effect of classroom practices. In the second analysis, it is possible to hypothesize that if training is going to be effective in raising student achievement, it ought to do so

through each of the dimensions of classroom practices. Four models were created in which each of the four dimensions of classroom practices (sub scales of teaching quality) predicted the overall student achievement. Note that only modest positive effect for teacher student interaction β=0.13, t=1.64; and measurement and evaluation β=0.19, t=1.66, were found. No significant effects were found for curriculum and teaching content and the use of instructional materials. It is not known at this point why the use of the instructional materials did not predict achievement contrary to results of similar studies on school effects. Fuller (1987), in an exhaustive review of factors that raise student achievement in the Third World, found a consistent positive impact of instructional materials across several studies. Of the 24 studies that looked into the relationship between availability of instructional material (measured, for instance, in terms of the number of textbooks available per student) and student achievement, 16 confirmed a significant effect of textbooks on achievement.

Comparative effectiveness of training. In the final stage of this analysis, it was necessary to know to what extent each training predicts its own dimension of classroom practices. The point here is that an effective training program designed to improve the quality of instruction ought to achieve that objective. Otherwise it is not an effective training program. For example, an effective training program on teacher-student

interaction should improve a teacher's competence in actively involving their students during classroom instruction, and so on.

Four models were created in which each topic of the training program predicted each dimension of classroom practice. For example, training in teacher-student interaction predicted teacher-student interaction; Training in measurement and evaluation predicted effective use of measurement and evaluation; Training in curriculum and teaching content predicted teacher fairness and concern; and Training in the use of instructional materials predicted the use of instructional materials.

For this analysis, it is necessary to compare the result of each topic to the others. The result from the final analysis indicates that training in measurement and evaluation was more effective than the remaining training programs β=25.2, t=1.63. The remaining three programs did not predict their respective dimensions of classroom practice.

Again it is not clear at this point why training in measurement and evaluation was more effective than others. It would have been informative to add the organization of training variable into the model to see whether it will help to explain this variation. But there was no variation in that variable, since over 80% of the training was provided by the Office of the National Primary Education Commission (ONPEC), the office responsible for the administration of primary

education in Thailand. The only thing that could be said at this time is that the result on measurement and evaluation is consistent with the theory and results of past research of school effects in Third World countries.

The results for the use of instructional materials are discouraging and not consistent with the theory and results of past research. Evidence from case studies done on effective schools in Thailand indicates that effective schools were those where instructional materials are effectively used. The key issue is not the mere availability of instructional materials. Materials are a necessary condition for effective instruction. Furthermore, materials need not be costly. Rather, the processes by which educators develop, distribute and use, or fail to use instructional materials are key indicators of the organizational health of the school (Wheeler et al, 1990).

Results from research conducted in rural schools in Brazil, indicates that students learn less in classes in which seat work and workbook use is more frequent, and learn more in classes in which students receive frequent individual feedback (Lockheed et al, 1990). The Lockheed result is encouraging in helping us to understand why the use of instructional materials did not work. Since most of the schools in this sample are rural schools, it is possible to assume that workbooks, as opposed to textbooks, may have been used in these schools. If the answer is yes, this result is consistent

with the findings of past research in rural schools in the Third World. Case studies or on-site visitations to these rural schools in the sample are needed to ascertain the type of instructional materials used during classroom instruction. It is also necessary to confirm that workbooks are used as instructional material in these schools before this assumptions could be validated.

Improving Analytic Models.

The persistent "unit of analysis" problems that generally plague research on school effectiveness have been addressed in this study with the application of HLM as the primary analytical tool. In this study students were viewed as nested within classrooms and classrooms nested within schools. The problem associated with the "unit of analysis" in multilevel data of this type would have been difficult to solve with traditional research methods.

This study represents the first time HLM was used in an analysis that examines the relationship between a national in-service training program, teacher classroom practices and student achievement in the Third World.

In this process, it was necessary to study the effect of teacher training on teacher classroom practices and student achievement simultaneously. This is a relatively new approach and has been termed "the third wave" of research for effective schools (Riddel, 1989). This methodology has the potential to

more accurately model the process that affects student achievement at multiple organizational levels. At the very least the investigator has reduced errors of interpretation that are introduced as a result of inherent biases associated with more traditional analytical procedures.

Discussion

The need to improve the on-the-job competence of incumbent teachers is greater now that near-universal enrollment in primary education has been attained by the vast majority of developing nations. Policy makers and educational leaders can now concentrate more intensely on improving educational quality and efficiency while providing access. Of central concern to policy makers and development specialists are the types of strategies to improve schools and teacher effectiveness. When educators and researchers have considered these strategy, they have typically regarded in-service training programs as the primary, or even the sole, policy option for improving the productivity of incumbent teachers.

Recall that the government of Liberia has relied solely on in-service training programs for the preparation of its teaching force (Thiagaran, 1990).

Moreover, considerable research is now available suggesting that in-service training can significantly improve teacher classroom performance and student outcomes, and the characteristics of successful and unsuccessful programs are

gradually coming into focus. However, the present study is based on two apparent weaknesses in research and thinking about improving the effectiveness of teachers on different dimensions of classroom practice.

First, few studies have sought to assess the summative effect of a nation's investment in-service training on dimensions of teaching quality. Here the question is not whether in-service training can have an effect on classroom practices and student achievement, though that question is important. Rather, the issue is whether empirical evidence can be found to indicate that a nation's investment on in-service training has paid off.

If such evidence of the effect of in-service training is scanty, many would ask how in-service training programs could be improved. That is certainly an important question, and one that has absorbed the interest of numerous researchers, some of whose work were reviewed in chapter 2 of this study. The results are consistent, and they indicate that in-service training programs must be more intensive than short workshop programs, must have a follow-up, must meet the needs of teachers and students, and that knowledge gained out of the training must be practiced in a classroom setting to produce demonstrable results.

However, improving in-service training by increasing its intensity, providing for follow-up, and increasing the amount of financial support represent only one set of several options

available to policy makers in improving education quality and efficiency in the Third World. Hence, it is appropriate at this time to turn to a second question that has rarely been asked: What promise do policy options other than in-service training hold for improving the on-the-job competence of teachers on different dimensions of classroom practice.

Thailand's experience is relevant to both questions, perhaps because of its special circumstances, including a sharply reduced birth rate, the recent termination of a rapid expansion of primary school access, and a long-standing commitment to improving the quality of primary education. That country has implemented a wide variety of programs aimed at improving the performance of incumbent teachers. Among these are a variety of in-service training programs for improvement in classroom practices, and a campaign to improve the principal's and district supervisor's skill in providing supervision of classroom teaching.

The findings of this study, which apply to small schools with only one sixth grade classroom and one sixth grade teacher, are quite clear. There is empirical evidence to indicate that a teacher's experience in in-service training on measurement and evaluation courses predicts improved instructional practice on the dimension of teaching relating to monitoring student academic progress, and that the improved instructional practice in measurement and evaluation also predicts students' achievement. The evidence here is

consistent with the theory holding that in-service training improves student achievement by improving instructional quality. Of the four topics covered in the national in-service training program which were the focus of this study, training in measurement and evaluation seemed to be most effective. Thus, investment on this topic in a national in-service training program in Thailand may have produced a return.

In-service training in teacher-student interaction did not predict instructional quality on this dimension of teaching. However, the mean value of teacher-student interaction predicted student achievement. This simply indicates that teacher-student interaction is important in boosting academic learning, but is unaffected by a national in-service training program in Thailand. The results are consistent with past research findings on school effects in both developed and developing nations (active student involvement in classroom discussion raises student achievement). Thai policymakers should find ways to improve in-service training on this dimension of teaching in order to expect a return from their investment on this topic.

There is little empirical evidence to indicate that a teacher's experience in the national in-service training courses on curriculum and teaching content predicts improved instructional quality (teacher fairness and concern towards students) or student achievement. Investment in training on

this topic may not have paid off. The investigator does not know why this is so.

In-service training on the use of instructional materials did not predict improved instructional quality or student achievement. This is a surprising result considering the huge sums of money the government spends every year in providing instructional materials to schools. Thailand's investment in this topic may not have produced a return. The best guess is that Thai rural students must have perceived the use of instructional materials as workbooks. Evidence from past research findings in rural schools in developing nations indicate that workbooks do not enhance student achievement.

Suggestions for Future Research

A great deal could be learned from both the strengths and limitations of this study. It is argued that the HLM approach is the most appropriate tool at this time for investigating multi-level phenomena. Researchers who deal with such nested data, and that means the majority of researchers in social and behavioral sciences including biology, biochemistry, and education should seriously consider employing these techniques in their work, or be able to justify not using them on the grounds other than expediency.

HLM, like any regression approach, requires that the researcher make certain assumptions regarding the properties of the data. Student learning and the quality of classroom

instruction have been the focus of this research, as well as other research on classroom processes, teacher effectiveness, and school effects. The objections of those who warn educational researchers about the dangers of too narrow a focus should be heeded. For example, Stevenson (1987) has argued that the measures of student achievement involved in school effect research have been narrowly confined to basic skill performance, which is not generally agreed to be a central goal of secondary schooling.

Important work remains to be done with teacher behavior in the classroom and student learning, especially in the areas of improvement. But this should not prevent educational researchers from including other outcomes, such as teacher satisfaction, teacher knowledge, and teacher sense of efficacy.

In the light of the findings regarding in-service training on curriculum and teaching content, more studies will be needed to understand why students' perceptions of the teacher as being fair and showing concern toward their welfare during classroom instruction did not translate into heightened learning. More studies are also needed to indicate how in-service training in this particular dimension of teaching should be improved in order to expect a return from the investment on this topic. More studies in this area are needed especially with the urban students and students of high

socioeconomic status to see how their perceptions of teacher fairness and concern translates into achievement.

On the use of instructional materials, more studies are needed, especially with small rural schools in Third World countries, to validate the results of this study. Studies on cost effective strategies to improve on-the-job training of teachers on dimensions of classroom practice should be explored. Cross national studies on the effect of each of the four dimensions of classroom practice used in this study will contribute to the improvement of teaching-learning processes in both developed and developing nations.

A Final Note

The recommendations for future studies outlined here may seem to assume that researchers have unlimited resources. But the issue is not how to raise huge sums of money to support elaborate research projects. The challenge lies in finding ways to incorporate what is known to be the best procedure of analyzing multi level data inherent in social and behavioral science into existing research on school effects, teacher effectiveness and school improvement efforts. The bottom line is that research on classroom processes, teacher effectiveness, and school effects could be of great benefit to policy, if researchers in these areas could work more closely with each other. Generating guidelines for what works in

school is the easy part. Faithfully implementing them and evaluating their true impact is where the focus should be.

The dilemma and tensions inherent in stimulating a change while managing the overall educational system performance in the Third World countries presents challenges of major proportions, once an educational system moves from the stage of universal free primary education to the need to improve the quality of education.

While the results of Thailand's data do not provide answers to other Third World nations contemplating the path to educational reforms, they do illuminate the issue and suggest the kinds of questions which may be considered when a country decides to embark on educational reform. Thailand's experience, based on the research conducted here, could provide guidance for direction which might prove fruitful to other nations.

APPENDICES

APPENDIX 1-1

Predictor Variables

Student, School and Community Background

1. Student socio-economic status (ses) was derived from measures of the father's education, the mother's education and the natural logarithm of the amount of pocket money the child typically brings to school (alpha = .73), as reported by the students.

2. Students' linguistics background ("Dialect") was coded as "central Thai" vs "others" as reported by the students.

3. Student nutrition was measured by students' reports about how often they ate breakfast. An indicator variable was constructed: "Breakfast" (1 = daily, 0 = not daily).

4. Time needed to travel from home to school (in hours) was based on student reports.

5. Children's pre-primary experience was also coded dichotomously (1 = one or more years of pre-primary experience) based on student reports.

6. GPA (the student's grade point average) is a seven category variable (1 = 1.00 or lower; 2 = 1.01-1.50; 3 = 1.51-2.00; 4 = 2.01-2.50; 5 = 2.51 - 300; 6 = 3.01 - 3.50; 7 = 3.51 - 400).

APPENDIX 1-1 (cont'd.)

7. Repetition is an indicator variable (0 = never repeated a grade; 1 = ever repeated a grade).

8. Gender is an indicator taking on a value of 0 if female, and 1 if male.

School level variables indicating community and school background were measured as follows:

1. The modernity of community infrastructure was indicated by a scale consisting of the sum of nine items including the presence in the local community of drinking water, paved roads, irrigation, telephone service, electricity, hospital, a commercial bank, a factory and a post office as reported by the principal. Each item was coded dichotomously; these dichotomies were then summed up, and the resulting sum had an internal consistency of alpha = .81.

2. Student SES was aggregated to the school level to create the contextual variable, "Mean SES."

3. There were five geographic regions: "North," "South," "Central," "Bangkok," and "Northeast." Indicator variables were constructed for the first four so that the Northeast region constituted the reference group.

4. Remoteness is the mean of three items: distance to the district capital, distance to the market and distance to the highway (all in kilometers).

APPENDIX 1-1 (cont'd.)

Availability of Resources

1. Enrollment is the count of students attending the school as reported by the principal.

2. Class size is the enrollment in the school's sole sixth grade classroom as reported by the principal.

3. Equipment was based on an 18-item scale including equipment used for instruction ("hard technologies"), but also including some equipment that could be used for administration. These were highly intercorrelated and seemed best viewed as a single factor. Items included the presence or absence of a Thai typewriter, English typewriter, copying machine, slide projector, overhead projector, amplifier, radio cassette, radio, tape module, television, video cassette player, sewing machine, microscope, metronome, scale (for weighing), Thai musical instrument, international musical instrument, and water tank. The overall alpha was .77.

4. The availability of textbooks and workbooks, as indicated by the sum of the texts and workbooks available to student use across the five areas of the curriculum: (alpha = .75: student report). Maximum total is 10.

5. The availability of instructional material is based on 12 items indicating the teachers' assessment of each resource as sufficient or insufficient. Resources include a teaching manual, textbooks, and books to broaden the teachers' knowledge.

APPENDIX 2-1

RELIABILITY ESTIMATES FOR THE OUTCOME VARIABLE

Variable	# of Items	Alpha	Rho	TAU/B_AR
Total Achievement	219	.89	.48	.94
Teacher-Student Interaction	4	.64	.42	.94
Measurement and Evaluation	3	.49	.47	.94
Curriculum and Teaching Content	4	.45	.47	.94
Use of Instructional Materials	1		.30	.90

Note. Cronbach's Alpha was provided by SPSSx. Intraclass correlations (Rho) and school level reliability (TAU/D_BAR) are based on the results of HLM analyses.

APPENDIX 3-1

CORRELATION MATRIX FOR SCHOOL LEVEL VARIABLES BY THE SPSSX PROGRAM

Correlations:	NORTH	DTCTRAIN	DTPTRAIN	DTATRAIN	DTETRAIN	NTQUALD
NORTH	1.0000	.0926	.0927	.0923	.0924	.0839
DTCTRAIN	.0926	1.0000	1.0000**	1.0000**	1.0000**	.1009
DTPTRAIN	.0927	1.0000**	1.0000	1.0000**	1.0000**	.1007
DTATRAIN	.0923	1.0000**	1.0000**	1.0000	1.0000**	.1007
DTETRAIN	.0924	1.0000**	1.0000**	1.0000**	1.0000	.1003
NTQUALD	.0839	.1009	.1007	.1007	.1003	1.0000
REMOTEC	-.0508	-.0548	-.0548	-.0551	-.0547	-.0409
L_INFRAC	-.0111	-.0498	-.0496	-.0500	-.0500	.0838
ΣFACTOTC	.0176	-.0182	-.0183	-.0186	-.0184	.0783
MTXTBKC	-.1721	.1043	.1041	.1046	.1044	.1636
LTOINSPC	-.0802	.0429	.0427	.0425	.0428	-.2131
LTOINUTC	-.0816	.0199	.0197	.0196	.0198	-.1748
MTEQPEDC	.1355	.0096	.0094	.0096	.0096	.1007
MTEQMEAC	.1370	-.0239	-.0238	-.0237	-.0236	.1382
MTEQCURC	-.0876	-.1085	-.1085	-.1085	-.1081	.0794
MTEQMATC	-.0637	-.0192	-.0195	-.0196	-.0194	.0979

1-tailed Signif: * - .01 ** - .001

Note: NORTH is the geographical region. DTCTRAIN is in-service training in curriculum and teaching content. DTPTRAIN is in-service training in teacher-student interaction. DTATRAIN is in-service training in the use of instructional materials. DTETRAIN is in-service training in measurement and evaluation. NTQUALD is pre-service education.

APPENDIX 3-2

CORRELATION MATRIX FOR SCHOOL LEVEL VARIABLES BY THE SPSSX PROGRAM

Correlations:	REMOTEC	L_INFRAC	ZFACTOTC	MTXTBKC	LTOINSPC	LTOINUTC
NORTH	-.0508	-.0111	.0176	-.1721	-.0802	-.0816
DTCTRAIN	-.0548	-.0498	-.0182	.1043	.0429	.0199
DTPTRAIN	-.0548	-.0496	-.0183	.1041	.0427	.0197
DTATRAIN	-.0551	-.0500	-.0186	.1046	.0425	.0196
DTETRAIN	-.0547	-.0500	-.0184	.1044	.0428	.0198
NTQUALD	-.0409	.0838	.0783	.1636	-.2131	-.1748
REMOTEC	1.0000	-.5070**	-.1498	-.1545	.1520	.1433
L_INFRAC	-.5070**	1.0000	.2276*	.2318*	-.2657*	-.2318*
ZFACTOTC	-.1498	.2276*	1.0000	.3000**	.0344	.0207
MTXTBKC	-.1545	.2318*	.3000**	1.0000	.0725	.0151
LTOINSPC	.1520	-.2657*	.0344	.0725	1.0000	.9146**
LTOINUTC	.1433	-.2318*	.0207	.0151	.9146**	1.0000
MTEQPEDC	.0133	.1215	.2679*	.0922	.0966	.1004
MTEQMEAC	.0166	-.0428	.0834	-.0293	.1373	.1943
MTEQCURC	-.0027	.0485	.0473	-.0772	.1340	.2337*
MTEQMATC	.0114	-.0958	.0620	-.0251	.0365	.0853

1-tailed Signif: * - .01 ** - .001

Note: REMOTEC is remoteness centered around its grand mean. L_INFRA is the natural log for infrastructure. ZFACTOTC is the Z score for equipments and facilities centered around its grand mean. MTXTBOOKC is the aggregated students textbooks centered around its grand mean. LTOINSPC is the natural log for internal supervision. LTOINUTC is the natural log for the utilization of of internal supervision.

APPENDIX 3-3

CORRELATION MATRIX FOR SCHOOL LEVEL VARIABLES BY THE SPSSX PROGRAM.

Correlations:	MTEQPEDC	MTEQMEAC	MTEQCURC	MTEQMATC
NORTH	.1355	.1370	-.0876	-.0637
DTCTRAIN	.0096	-.0239	-.1085	-.0192
DTPTRAIN	.0094	-.0238	-.1085	-.0195
DTATRAIN	.0096	-.0237	-.1085	-.0196
DTETRAIN	.0096	-.0236	-.1081	-.0194
NTQUALD	.1007	.1382	.0794	.0979
REMOTEC	.0133	.0166	-.0027	.0114
L_INFRAC	.1215	-.0428	.0485	-.0958
ZFACTOTC	.2679*	.0834	.0473	.0620
MTXTBKC	.0922	-.0293	-.0772	-.0251
LTOINSPC	.0966	.1373	.1340	.0365
LTOINUTC	.1004	.1943	.2337*	.0853
MTEQPEDC	1.0000	.5645**	.5148**	.2551*
MTEQMEAC	.5645**	1.0000	.6867**	.2903**
MTEQCURC	.5148**	.6867**	1.0000	.3126**
MTEQMATC	.2551*	.2903**	.3126**	1.0000

1-tailed Signif: * - .01 ** - .001

NOTE MTEQPEDC is the aggregated subscale for perceived instructional quality (teacher-student interaction) centered around its grand mean. MTEQMEAC is the aggregated subscale of perceived instructional quality (measurement and evaluation) centered around its grand mean. MTEQCURC is the aggregated subscale of perceived instructional quality (teacher fairness and concern) centered around its grand mean. MTEQMATC is the aggregated subscale for perceived instructional quality (use of instructional materials) centered around its grand mean.

APPENDIX 4-1

CORRELATION MATRIX FOR STUDENT LEVEL VARIABLES BY THE SPSSX PROGRAM

Correlations:	TOTLACHC	GPAC	SESC	PPED1D	BRF1	REP1
TOTLACHC	1.0000	.4836**	.1676**	.1213**	.0534*	-.1400**
GPAC	.4836**	1.0000	.1764**	.1308**	-.0125	-.1942**
SESC	.1676**	.1764**	1.0000	.2150**	-.0422	-.1020**
PPED1D	.1213**	.1308**	.2150**	1.0000	-.0614*	-.0699**
BRF1	.0534*	-.0125	-.0422	-.0614*	1.0000	-.0248
REP1	-.1400**	-.1942**	-.1020**	-.0699**	-.0248	1.0000
DIALCT1	.1449**	.1361**	.2744**	.0935**	-.1023**	-.0769**
SAGEC	-.0659*	-.1429**	-.2300**	-.1125**	.0107	.3993**
SAGEC	-.0659*	-.1429**	-.2300**	-.1125**	.0107	.3993**

1-tailed Signif: * - .01 ** - .001

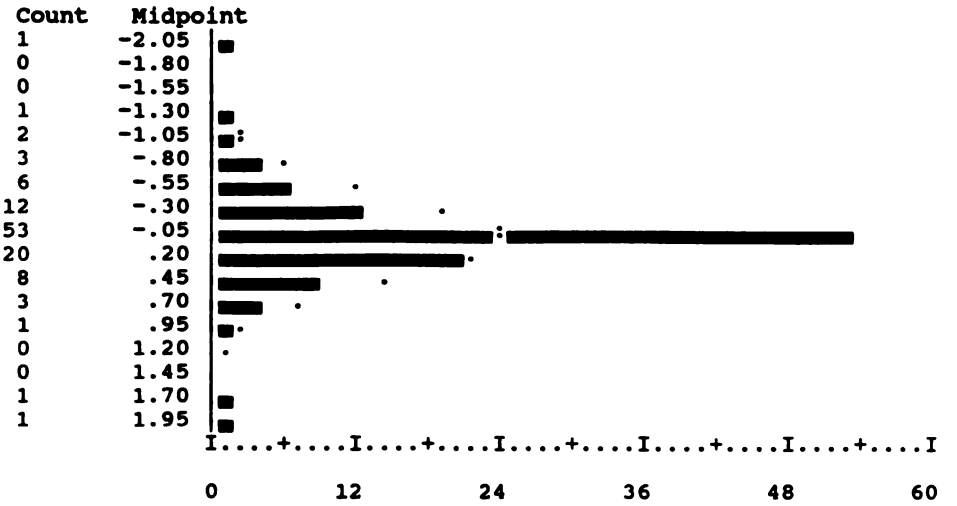
Correlations:	DIALCT1
TOTLACHC	.1449**
GPAC	.1361**
SESC	.2744**
PPED1D	.0935**
BRF1	-.1023**
REP1	-.0769**
DIALCT1	1.0000

1-tailed Signif: * - .01 ** - .001

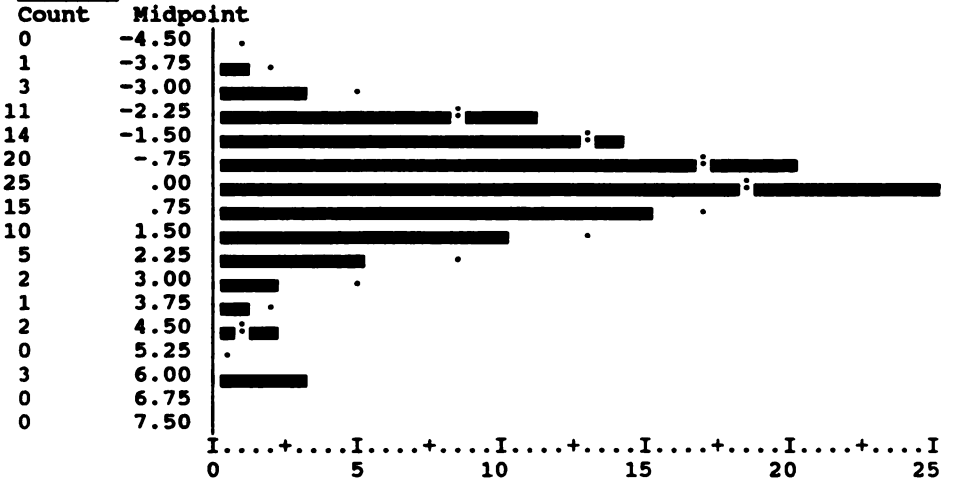
Note: TOTLACHC is total student achievement centered around its grand mean. GPAC is student grade point average centered around its grand mean. SESC is the student socioeconomic status centered around its grand mean. PPED1D is preprimary education. BRF1 is breakfast. REP1 is student repetition. DIALCT1 is student dialect.

APPENDIX 5-1

HISTOGRAM FREQUENCY FOR USE OF INSTRUCTIONAL MATERIAL BY THE SPSSX PROGRAM

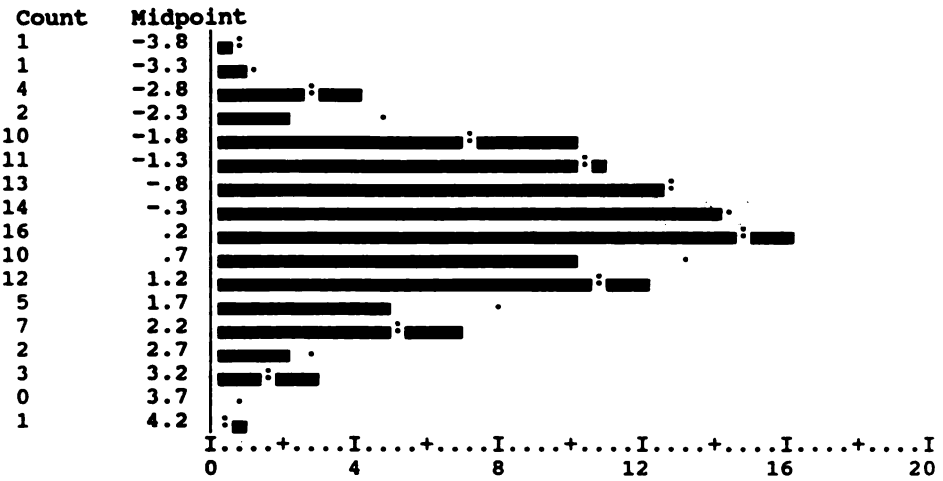


HISTOGRAM FREQUENCY FOR TEACHER FAIRNESS AND CONCERN BY THE SPSSX PROGRAM

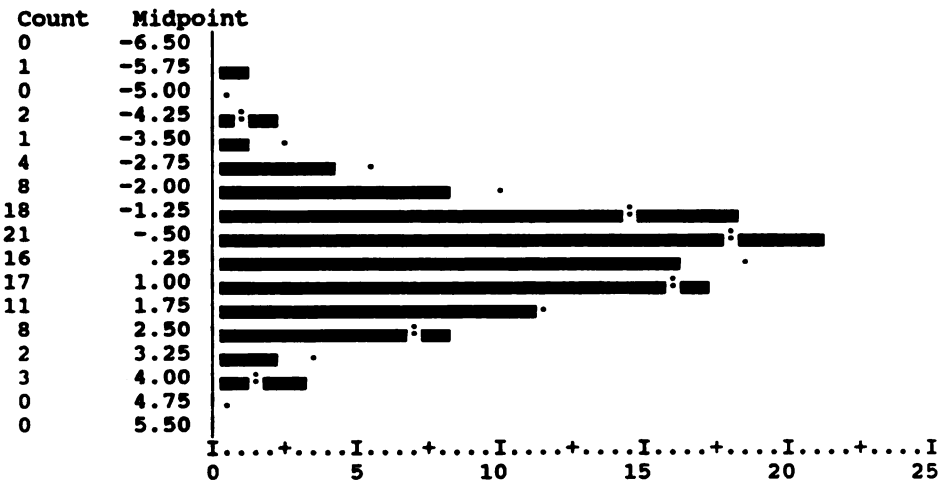


APPENDIX 5-2

HISTOGRAM FREQUENCY FOR MEASUREMENT AND EVALUATION BY THE SPSSX PROGRAM

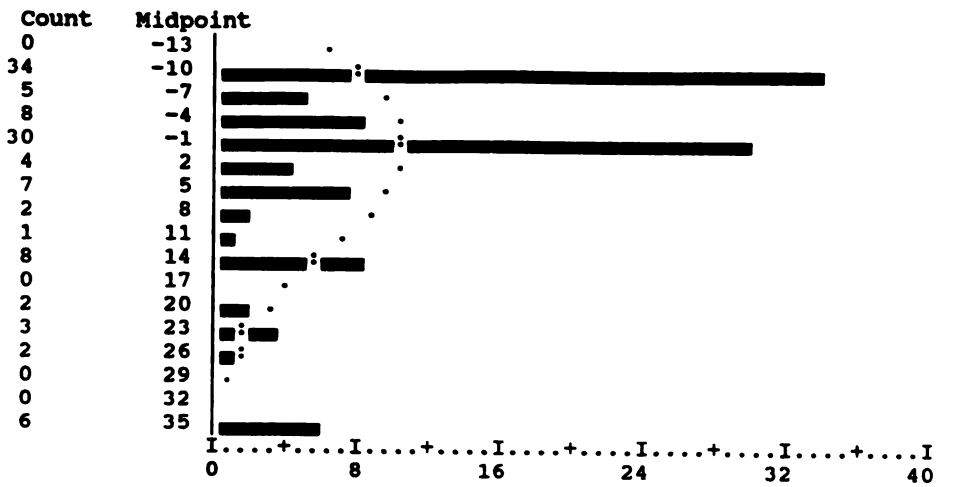


HISTOGRAM FREQUENCY FOR TEACHER STUDENT INTERACTION BY THE SPSSX PROGRAM

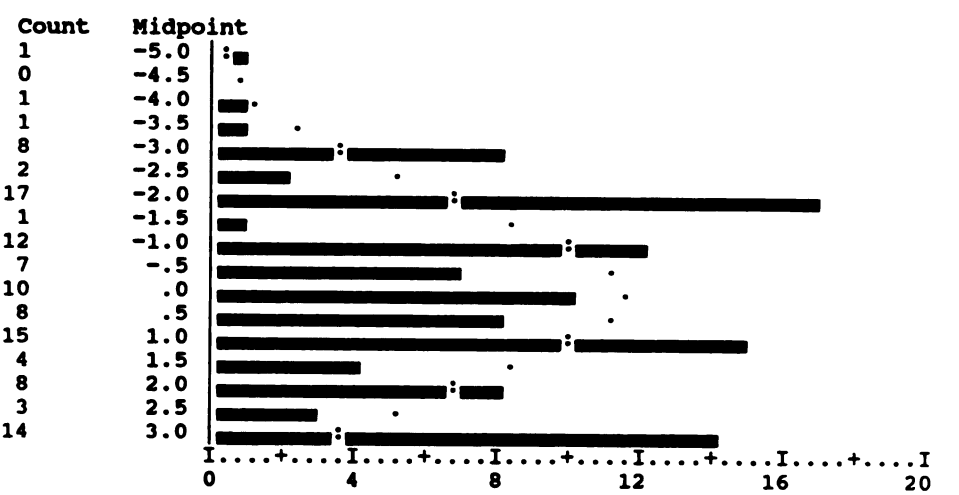


APPENDIX 5-3

HISTOGRAM FREQUENCY FOR LOG UTILIZATION OF INTERNAL SUPERVISION BY THE SPSSX PROGRAM

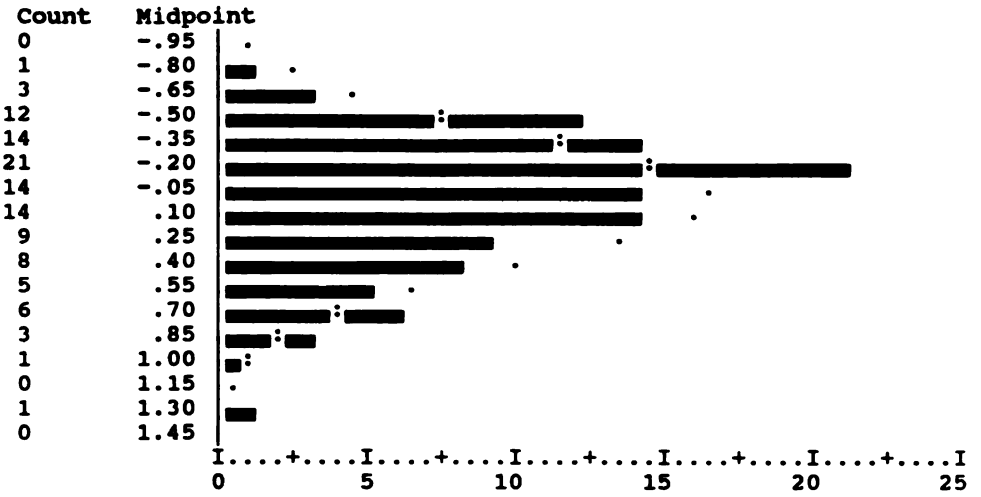


HISTOGRAM FREQUENCY FOR TEXTBOOKS BY THE SPSSX PROGRAM

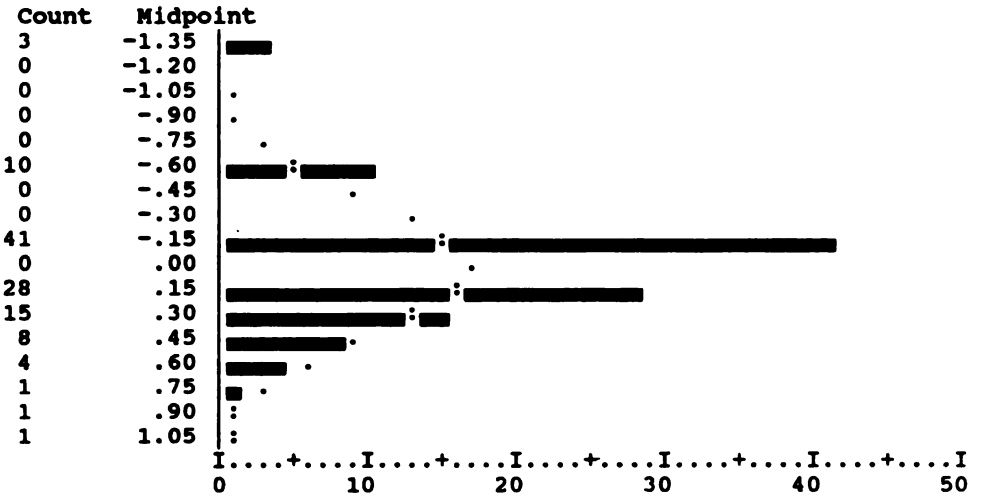


APPENDIX 5-4

HISTOGRAM FREQUENCY FOR EQUIPMENTS AND FACILITIES BY THE SPSSX PROGRAM

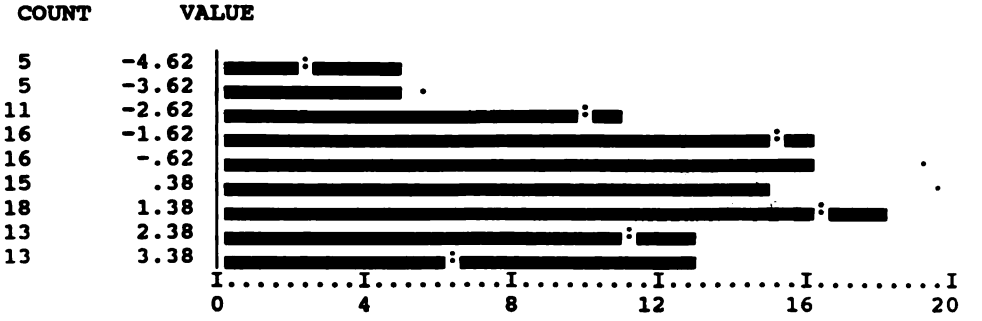


HISTOGRAM FREQUENCY FOR INFRASTRUCTURE BY THE SPSSX PROGRAM

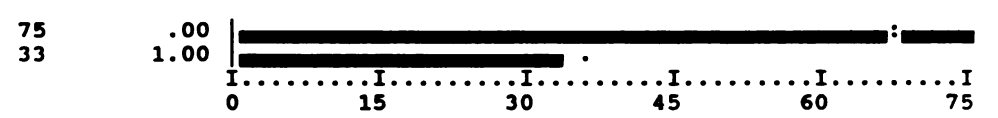


APPENDIX 5-5

HISTOGRAM FREQUENCY FOR REMOTENESS BY SPSSX THE PROGRAM

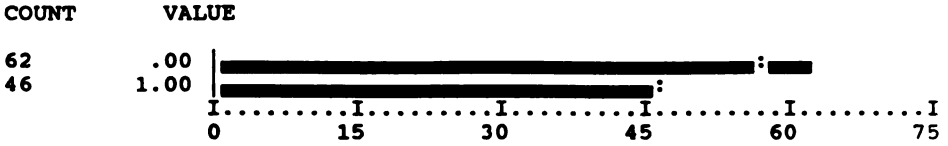


HISTOGRAM FREQUENCY FOR IN-SERVICE TRAINING IN MEASUREMENT AND EVALUATION BY THE SPSSX PROGRAM

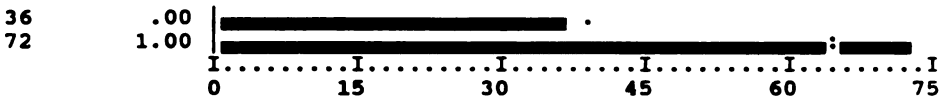


APPENDIX 5-6

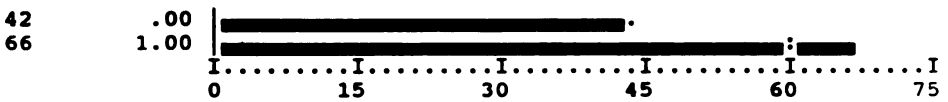
HISTOGRAM FREQUENCY FOR IN-SERVICE TRAINING IN USE OF INSTRUCTIONAL MATERIAL BY THE SPSSX PROGRAM



HISTOGRAM FREQUENCY FOR IN-SERVICE TRAINING TEACHER STUDENT INTERACTION BY THE SPSSX PROGRAM

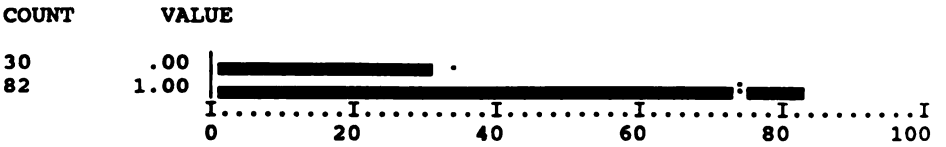


HISTOGRAM FREQUENCY FOR IN-SERVICE TRAINING IN CURRICULUM AND TEACHING CONTENT BY THE SPSSX PROGRAM

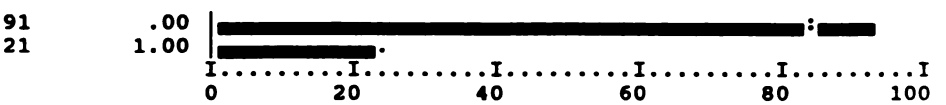


APPENDIX 5-7

HISTOGRAM FREQUENCY FOR PRESERVICE EDUCATION BY THE SPSSX PROGRAM

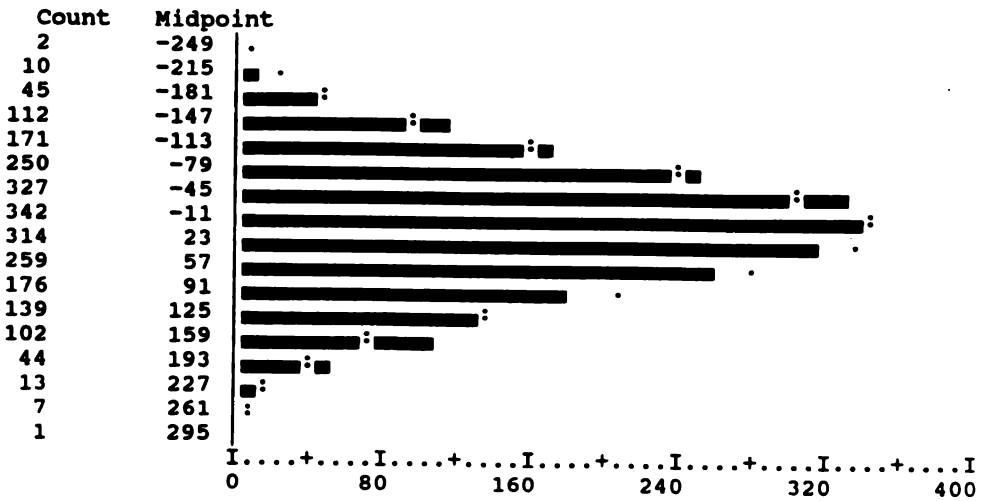


HISTOGRAM FREQUENCY FOR GEOGRAPHICAL REGION NORTH BY THE SPSSX PROGRAM

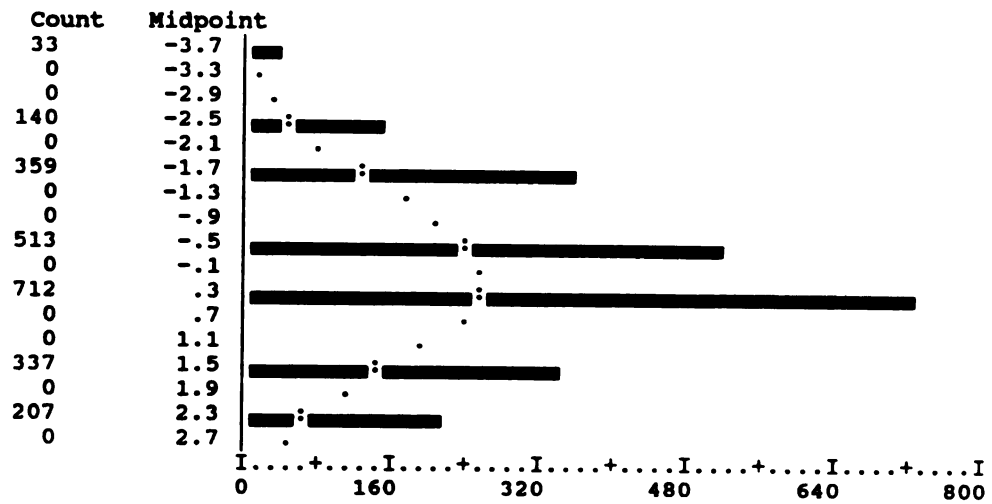


APPENDIX 6-1

HISTOGRAM FREQUENCY FOR TOTAL STUDENT ACHIEVEMENT BY THE SPSSX PROGRAM



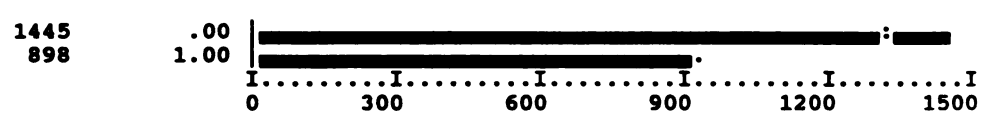
HISTOGRAM FREQUENCY FOR GRADE POINT AVERAGE BY THE SPSSX PROGRAM



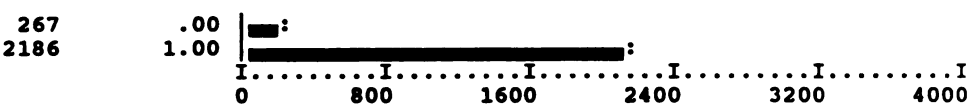
APPENDIX 6-2

HISTOGRAM FREQUENCY FOR STUDENT SOCIOECONOMIC STATUS BY THE SPSSX PROGRAM

HISTOGRAM FREQUENCY FOR STUDENT PREPRIMARY EDUCATION BY THE SPSSX PROGRAM

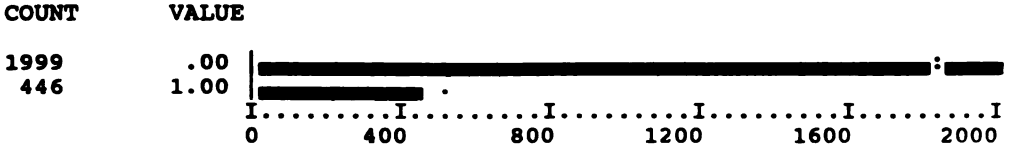


HISTOGRAM FREQUENCY FOR STUDENT BREAKFAST BY THE SPSSX PROGRAM

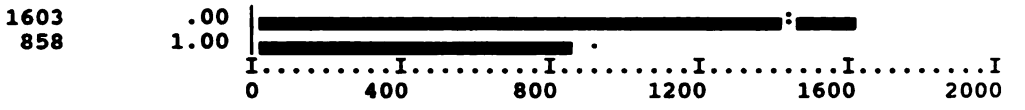


APPENDIX 6-3

HISTOGRAM FREQUENCY FOR STUDENT REPETITION BY THE SPSSX PROGRAM



HISTOGRAM FREQUENCY FOR STUDENT LANGUAGE AT HOME BY THE SPSSX PROGRAM



APPENDIX 7-1

FACTOR ANALYSIS FOR TOTAL STUDENT ACHIEVEMENT BY THE SPSSX PROGRAM

- - - - F A C T O R A N A L Y S I S - - - -

Analysis Number 1 Listwise deletion of cases with missing values

Extraction 1 for Analysis 1, Principal-Components Analysis (PC)

Initial Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
TH11	1.00000	*	1	6.20383	27.0	27.0
TH12	1.00000	*	2	1.71827	7.5	34.4
TH13	1.00000	*	3	1.12649	4.9	39.3
TH2	1.00000	*	4	1.02164	4.4	43.8
MA11	1.00000	*	5	.99552	4.3	48.1
MA12	1.00000	*	6	.89085	3.9	52.0
MA21	1.00000	*	7	.83922	3.6	55.6
MA22	1.00000	*	8	.80534	3.5	59.1
MA3	1.00000	*	9	.79851	3.5	62.6
MA4	1.00000	*	10	.78061	3.4	66.0
WEX	1.00000	*	11	.73368	3.2	69.2
LEX11	1.00000	*	12	.71370	3.1	72.3
LEX12	1.00000	*	13	.70330	3.1	75.4

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
LEX13	1.00000	*	14	.68163	3.0	78.3
LEX14	1.00000	*	15	.65578	2.9	81.2
LEX21	1.00000	*	16	.61869	2.7	83.9
LEX22	1.00000	*	17	.61049	2.7	86.5
CD11	1.00000	*	18	.58150	2.5	89.0
CD12	1.00000	*	19	.56376	2.5	91.5
CD13	1.00000	*	20	.53284	2.3	93.8
CD21	1.00000	*	21	.50597	2.2	96.0
CD22	1.00000	*	22	.49864	2.2	98.2
CD23	1.00000	*	23	.41972	1.8	100.0

PC Extracted 4 factors.

APPENDIX 7-1 (cont'd.)

FACTOR ANALYSIS FOR TOTAL STUDENT ACHIEVEMENT BY THE SPSSX PROGRAM

Factor Matrix:

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
TH11	.42623	.00390	.11425	-.22955
TH12	.54556	-.01848	.05403	-.33084
TH13	.50817	-.05013	.04986	-.04529
TH2	.42918	-.06317	.08845	.62678
MA11	.68379	-.18819	-.11977	-.05702
MA12	.50300	-.15249	-.27493	-.27028
MA21	.66284	-.13600	-.21337	.10953
MA22	.63563	-.18081	-.19697	.03527
MA3	.68165	-.13638	-.15093	.25526
MA4	.59465	-.14482	-.09331	.35696
WEX	.58501	-.04556	.09429	-.11036
LEX11	.60257	.02120	.08853	-.11259
LEX12	.58302	-.04431	.06947	-.16061
LEX13	.55740	-.00401	.06161	.06663
LEX14	.49634	.02168	.13976	.11699
LEX21	.60252	-.04026	.09317	-.26618
LEX22	.62561	-.13008	.01487	-.05551
CD11	.26375	.70815	-.27527	.01202
CD12	.13253	.66299	-.34232	.02535
CD13	.42641	.54858	-.22422	.03186
CD21	.31666	.29033	.39429	-.00173
CD22	.29741	.39451	.40075	-.06987
CD23	.29238	.23782	.53046	.18184

APPENDIX 7-1(cont'd.)

FACTOR ANALYSIS FOR TOTAL STUDENT ACHIEVEMENT BY THE SPSSX PROGRAM

Final Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
TH11	.24743	*	1	6.20383	27.0	27.0
TH12	.41035	*	2	1.71827	7.5	34.4
TH13	.26529	*	3	1.12649	4.9	39.3
TH2	.58885	*	4	1.02164	4.4	43.8
MA11	.52057	*				
MA12	.42490	*				
MA21	.51537	*				
MA22	.47676	*				
MA3	.57119	*				
MA4	.51071	*				
WEX	.36538	*				

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
LEX11	.38406	*				
LEX12	.37249	*				
LEX13	.31895	*				
LEX14	.28004	*				
LEX21	.44419	*				
LEX22	.41161	*				
CD11	.64696	*				
CD12	.57494	*				
CD13	.53405	*				
CD21	.34004	*				
CD22	.40958	*				
CD23	.45650	*				

Varimax Rotation 1, Extraction 1, Analysis 1 - Kaiser Normalization.

Varimax converged in 7 iterations.

Note: that 27 of the variation is explained by factor 1 with an Eigenvalue of 2.2, indicating that Total Student Achievement is made up of only one primary factor.

APPENDIX 7-1 (cont'd.)

FACTOR ANALYSIS FOR TOTAL STUDENT ACHIEVEMENT BY THE SPSSX PROGRAM

Rotated Factor Matrix:

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
TH11	.46085	.01555	.03020	.18411
TH12	.62119	.00617	.06369	.14277
TH13	.43774	.22862	.03515	.14202
TH2	-.00295	.73776	-.00199	.21106
MA11	.62378	.36000	.03713	-.02210
MA12	.60141	.10404	.10437	-.20369
MA21	.50760	.48699	.12534	-.06961
MA22	.53645	.41930	.07286	-.08866
MA3	.43694	.60850	.10001	.00175
MA4	.30889	.64166	.04825	.03533
WEX	.53192	.20812	.03182	.19523
LEX11	.53311	.20205	.09470	.22374
LEX12	.55982	.16872	.04403	.16940
LEX13	.40328	.33472	.07983	.19468
LEX14	.31621	.32929	.05143	.26262
LEX21	.63251	.08819	.03835	.18675
LEX22	.55622	.30194	.00830	.10488
CD11	.06673	.03215	.78846	.14072
CD12	-.03153	-.00644	.75689	.03212
CD13	.21638	.15918	.66139	.15639
CD21	.16726	.05045	.11568	.54418
CD22	.16748	-.03811	.19590	.58455
CD23	.04791	.18239	.00170	.64879

Factor Transformation Matrix:

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
FACTOR 1	.79573	.51490	.20216	.24662
FACTOR 2	-.21541	-.20985	.84654	.43924
FACTOR 3	-.05872	-.12744	-.49226	.85906
FACTOR 4	-.56299	.82134	.01318	.09091

BIBLIOGRAPHY

BIBLIOGRAPHY

Aitkin, M. & Longford, N. (1986). Statistical modeling issues in school effectiveness studies. Journal of the Royal Statistical Society, Series B 149: 1-143.

Anderson, C.S. (1982). The search for school climate: a review of the research. Review of Educational Research, 52,368-420.

Anderson, L. (1987). The classroom environmental study: Teaching for learning. Comparative Education Review. 31, 1, 69-85.

Anderson, L. & Djalil, A. (1989). The impact of research-based teacher training program on Indonesian teachers, classrooms, and students. Teaching and Teacher Education, 5, 3, 165-178.

Anderson, L.W., Ryan, D.W., & Shapiro, B.J. (1989). The IEA classroom environment study. Oxford: Pergamon press.

Armitage, J., Batista, J. Harbison, R.W., Holisinger, D.B., & Helio, R. (1986). School quality and achievement in rural Brazil. Washington DC: The World Bank, Education and Training Department.

Arriagada, A. (1983). Determinants of sixth grade student achievement in Peru. Unpublished Document. Washington DC: The World Bank.

Awolowo, A. (1986). Teacher effectiveness towards the definition of a complex concept. (Eric Document Reproduction Service No. ED 281-848).

Bhumirat, C., Kidchanapanish, P., Arunrungrueng, P., & Shinatrakool. (1987). Research and evaluation on the quality of primary education in Thailand. Bridges Project International Development.

Bridge, R., Judd. C., & Moock, P. (1979). The determinants of educational outcomes: The impact of families, peers, teachers and schools. Cambridge, MA: Ballinger.

Brookover, W., Beady, C., & Moock, P. (1979). The determinants of educational outcomes: The impact of families, peers, teachers, and schools. Cambridge, MA: Ballinger.

Boothroyd, R.A., & Chapman, D.W. (1987). Gender Differences in student achievement in Liberia. International Journal of Educational Development,7,2,99-105.

Brophy, J.E. & Good, T.M. (1986). School effects. Handbook of Research on Teaching. 570-602.

Brophy, J.E. & Good, T.M. (1986). Teacher behavior and student achievement. Handbook of Research on Teaching. 328-375.

Bryk, A.S. & Valerie, L.E. (1989). A multilevel model of the social distribution of high school achievement. Sociology of Education.62, 172-192.

Bryk, A.S. & Raudenbush, S.W. (in press). Methodology for cross level organizational research. To appear in the Research in Organizational Behavior.

Bryk, A.S. & Raudenbush, S.W. (1992). Hierarchical Linear Models for Social and Behavioral Research: Application and Data Analysis Method. (In Press).

Bryk, A.S., Raudenbush, S.W. Seltzer, M., & Congdon, R. (1988). An Introduction to HLM: Computer Program and User's Guide. University of Chicago Department of Education.

Burstein, L. (1983). The analysis of multilevel data in educational research and evaluation. Review of Research in Education, 8,158-233

Chapman, D.W. & Carrier, C.A. (1990). Improving educational quality: A global perspective. New York. Greenwood press.

Chapman, D.W. (1990). Rethinking teacher training in Subsaharan Africa. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Chen, Y.H., Clark, T.B., & Schaffer, E.C. (1989). Teaching variables and Mathematics Achievement in the context of sixth classrooms in Taiwan. Teaching and Teacher Education, 4,4,115-123.

Coladarci, T., & Gage, N.L.(1984). Effects of minimal intervention on teacher behavior and student achievement, American Educational Research Journal,21, 539-555.

Coleman, J.S. et al., (1966). Equality of educational opportunity (Washington D.C., Government Printing Office).

Cook, T.D. & Campbell, D.T. (1979). Quasi-experimentation, design and analysis issues for field settings. Chicago: Rand Mcnally. [3,20,21,458,718,732,752].

Cronbach, L.J. (1976). Research on classrooms and schools: Formulation of questions, Design and Analysis. Stanford, CA: Stanford Evaluation Consortium.

Dempster, A.P., Laird, N.M., & Rubin, D.B. (1979). Maximum likelihood from incomplete data via the EM algorithm. Journal of Royal Statistical Society, 39, (Series B) 1-8.

Dempster, A.P., Rubin, D.B. & Tsutakawa, R.K. (1981). Estimation in covariance components models. Journal of The American Statistical Association, 76,341-353.

Devoe, D.E. (1991). Teacher behavior directed towards individual students in Elementary physical education. Journal of Classroom Interaction, 26,1,9-14.

Doyle, W. (1986). Classroom organization and management. In M.C. Wittrock (Ed.), Handbook of Research on Teaching (3rd., pp.392-431). New York: Macmillan.

Emmer, E.T., Sanford, J., Clements, B., & Martin, J. (1982). Effective classroom management and organization in junior high schools: An experimental investigation. Austin: Research and Development Center for Teacher Education, University of Texas.

Evaluation report submitted to the office of the National Primary Education Commission, Bangkok, Thailand, (in Thai).

Evertson, C.M., Weade, R.J., Emmer, E.T., Crawford, J. (1985). Effective classroom management and instructions: An exploration of models. Final report (NIE grant 83-0063). Washington, DC: National Institute of Education.

Fisher, C., Filby, N., Marliane, R., Cahen, L., Dishaw, M., Moore, J., & Berliner, D. (1978). Teaching behaviors, academic learning time and student achievement: Final report of phase 111-B. Beginning Teacher Evaluation Study: San Francisco: Far West Laboratory.

Fresko, B. & Ben-chaim, D. (1985). The impact of inservice teacher education on mathematics teacher subject-matter confidence, and competency. Teaching and Teacher Education,1,4, 317-324.

Fullan, M. (1982), The Meaning of Educational Change. New York: Teachers College Press.

Fuller, B. (1987). Raising school quality in developing countries: What investments improve school quality? Review of Educational Research, 57, 255-291.

Fuller, B. (1989). Vocal teachers, silent pupils? Life in Botswana classroom. Paper presented at the annual meeting of the comparative and International Education Society, Boston.

Fuller, B. & Heyneman, S. (1989). Third world quality: current collapse, future potential. Educational Researcher, 18,2,12-19.

Fuller, B. & Lockeed, M. (1987). Policy choice and school efficiency in Mexico. Policy, Research, and External Affairs Working Paper, Population and Human resources Department. Washington DC: The World Bank.

Fuller, B. & Gormank, E. (1986). When does education boost economic growth? School expansion and school quality in Mexico. Sociology of Education,59, 168-181.

Gage, N.L. & Grouws, D.A. (1979). The Missouri mathematics effectiveness project: an experiment in fourth grade classrooms. Journal of Educational Psychology, 71,335-362

Gage, N.L. & Needles, M.C. (1989). Process-product research on teaching: A review of criticisms. The Elementary School Journal, 89,3,254-300.

Goldstein, H. (1987). Multilevel models in social and educational research. New York: Oxford University Press.

Griffin, G.A. (1983). Implications of research for staff development programs. The Elementary School Journal, 83,4,414-423

Guthrie, G. (1986). Current research in developing countries: The impact of curriculum reform on teaching. Teaching and Teacher Education,2,1,81-89.

Haddad, W., Carnoy, M., Rinaldi, R., & Regel, O. (1991). Education and development: evidence for new priorities. Policy, Research, and External Affairs Working Paper, Population and Human Resources Department. Washington, DC: The World Bank.

Hawes, H.W. (1976). Universal Primary Education in UNESCO Literature. UPE and World Bank Report. Prospects,6,1,39-56.

Heyneman, S.P. Farrell, J.P., & Sepulveda-Stuarto, M. (1978). Textbooks and achievement: What we know . Washington D.C.: The World Bank.

Heyneman, S.P. & Jamison, D.T. (1980). Student learning in Uganda: Textbook availability and other factors. Comparative Education Review,24,206-220.

Heyneman, S.P. & Loxley, W.A. (1983). The effect of Primary-School quality on academic achievement across twenty-nine High and Low income countries. American Journal of Sociology,88,6,1162-1194

Husen, T., Saha,L., & Noonan, R. (1978). Teacher training and student achievement in less developed countries. Policy, Research, and External Affairs Working Paper, Population and Human Resources Department. Washington, D.C.: The World Bank.

Ingvarson, L. & Mackenzie, D. (1988). Factors affecting the impact of inservice courses for teachers: Implications for policy. Teaching and Teacher Education, 2, 139-155.

Jamison, D.T., Searle, B., Galda, K., & Heyneman, S.P. (1981). Improving elementary mathematics education in Nicaragua: textbooks and radio on achievement. American Psychological Association, 556-567.

Jatoi, H. & Mcginn, N. (1991). Are Women better teachers? Data from Pakistan. Paper presented at the annual meeting of the Comparative and International Education Society, Pittsburgh.

Jencks, C. (1972). Inequality: A reassessment of the effect of family and schooling in America (New York, Basic Books).

Jenkins, F. (1988). A model for multilevel Path Analysis. Unpublished doctoral dissertation. Michigan State University, East Lansing.

Joyce, B. & Showers, B. (1980). Improving inservice training: Messages of the research. Educational Leadership, February, 379-385.

Kelly, E.F. (1984). Evaluation of the Improved Efficiency of Learning Project in Liberia, Africa: Overview and Peculiar Problems, Albany: The Evaluation Consortium, School of Education, State University of New York at Albany.

Kelly, G. & Lassa, P. (1983). The quality of learning in Nigerian Primary education: The UPE program 1976-1982. International Review of Education, 29,2,231-246.

Krieger, M. (1988). African policy linking education and development: standard criticism and new departure. International Review of Education,34,3,293-311.

Lockheed, M. & Bruns, B. (1990). School effects on achievement in secondary mathematics and Portuguese in Brazil. Policy, Research, and External Affairs Working Paper, Population and Human Resources Department. Washington, D.C.: The World Bank.

Lockheed, M.E., Vail, S., & Fuller, B. (1987). How textbooks affect achievement in developing countries: Evidence from Thailand. Educational Evaluation and Policy Analysis, 8, 379-392.

Lockeed, M.E. & Komenan, A. (1989). Teaching quality and student achievement in Africa: The case of Nigeria and Swaziland. Teaching and Teacher Education, 5, 379-392.

Lockheed, M.E. & Longford, N. (1991). School effects in mathematics achievement gain. In S.W. Raudenbush J.D. Willms (eds.) Schools, Pupils, and Classrooms: International Studies of Schooling from a Multilevel Perspective, pp 131-166. New York: Academic Press.

Lockheed, M.E. & Zhao, Q. (1991). The effect of local control on student achievement and attitudes in secondary schools in the Philippines. Paper presented at the annual meeting of the Comparative and International Education Society, Pittsburgh.

Lockheed, M.E. & Longford, N.T. (1990). A Multilevel model of school effectiveness in a developing country. Policy, Research, and External Affairs Working Paper, Population and Human Resources Department. Washington, D.C.: The World Bank.

Lockheed, M., Fuller, B., & Nyirongo S. (1989). Family effects on student achievement. Sociology of Education, 62, 239-256.

Lockheed, M.E. (1987). School and classroom effects on student learning gain: The case of Thailand. Policy, Research, and External Affairs Working Paper, Population and Human Resources Department. Washington, D.C.: The World Bank.

Loxley, W. (1984). Quality of Schooling in the Kalahari. Paper presented at the annual meeting of the Comparative and International Education Society, Houston, Texas.

Mandelville, G.K. & Rivers, J.L. (1991). The South Carolina PET study: teachers perceptions and student achievement. The Elementary School Journal, 91,4,377-407.

Marshall, J. (1988). A general statement on staff development evaluation. Journal of Staff Development,9,1,2-8.

Mason, W.M., Wong, G.M., & Entwistle, B. (1983). Contextual analysis through the multilevel linear model, in Sociological Methodology. San Francisco: Jossey Bass.

Mason, W.M., Wong, G.Y., & Entwistle, B. (1984). The multilevel linear model: A better way to do contextual analysis. Sociological Methodology. London: Jossey-Bass.

Mcdowell, D.W. (1980). The impact of the national policy on education on indigenous education in Nigeria. International Review of Education, 26,1,49-63.

Menlo, A., Marich, M., Collet, L., Evers, T., Fernandez, R., & Ferris, L. (1990). Teaching practices in the United States in the context of comparison with practices in four other countries. Comparative Education, 26,2/3,227-247.

Ministry of health report (1980). Office of the ministry of health. Bangkok, Thailand.

Nitsaisook, M. & Anderson, L. (1989). An experimental investigation of the effectiveness of inservice teacher education in Thailand. Teaching and Teacher Education, 5,4, 287-302.

Nuthal, G. & Alton-lee, A. (1990). Research on teaching and learning: Thirty years of change. Elementary School Journal, 90,5,547-571.

Odaet, C.F. (1991). Implementing educational policies in Uganda. Policy, Research, and External Affairs Working Papers, Population and Human Resources Department. Washington, D.C.: The World Bank.

Olson, M.L. (1987). Identification of criteria for effective teacher inservice education: application to the Kent inter institutional workshop for improving school programs. Unpublished doctoral dissertation. Michigan State University, East Lansing.

Pedhazur, E.J. (1982). Multiple regressions in behavioral research (2nd ed). New York: Holt Rinehart and Winston.

Purkey, S.C. & Smith, M.S. (1983). Effective schools: A review. The Elementary School Journal, 83,4, 426-452.

Psacharopoulos, G. (1991). Why educational policies can fail. Policy, Research, and External Affairs Working Paper, Population and Human Resources Department. Washington, D.C.: The World Bank.

Raudenbush, S.W. (1984). Application of a hierarchical linear model in educational research. Unpublished doctoral dissertation: Harvard University, Massachusetts.

Raudenbush, S.W. (1988). Theory and analysis in multilevel research: convergence or divergence. Paper presented at the conference on Catholic University of Nijmegen, the Netherlands.

Raudenbush, S.W., Bhumirat, C., & Kamali, M. (in press). Predictors and consequences of primary teachers' sense of efficacy and students' perceptions of the quality of instruction in Thailand. To appear in the International Journal of Educational Research.

Raudenbush, S.W. & Bryk, A.S. (1988). Methodological advances in analyzing the effects of classrooms on student learning. Review of Research in Education, 15, 423-474.

Raudenbush, S.W., Kidchanapanish, S., & Kang, S.J. (1991). The effect of pre-primary access and quality on educational achievement in Thailand. Comparative Education Review, 35,2,255-273.

Raudenbush, S.W., Eamsukkawat, S., Di-ibor, I., Kamali, M., & Taoklam, W. (1991). On-the-job improvements in teacher competence: policy options and their effects on teaching and learning in Thailand. Paper presented at the annual meeting of the Comparative and International Education Society, Pittsburgh.

Raudenbush, S.W. & Chinnapat, B. (1989). The distribution of resources for primary education and its consequences for educational achievement in Thailand. Michigan State University, East Lansing.

Raudenbush, S.W. & Bhumirat, C. (1988). Results of a national survey of primary education in Thailand, part 1: school size, sector, facilitities, equipments, textbooks, and teaching materials as a predictor of pupil achievement. Michigan State University, East Lansing.

Raudenbush, S.W. & Bryk, A.S. (1986). A hierarchical model for studying school effects. Sociology of Education, 59, 1-17.

Reynolds, D. (1982a). The search for effective schools. School Organization,2,215-237.

Rosenberg, B. (1973). Linear regression with randomly dispersed parameters. Biometrika, 60,61-75.

Rowan, B., Raudenbush, S.W. & Kang, S.J. (1991). School climate in secondary schools: A multilevel analysis. American Journal of Education, 99,2,238-266.

Riddel, A. (1991). An alternative approach to the study of school effectiveness in Zimbabwe. Paper presented at the annual meeting of the Comparative and International Education Society, Pittsburgh.

Rutter, M. (1983). School effects on pupils progress: Research findings and policy implications. Child Development, 54, 1-29.

Sanborn, M.A. (1987). Beyond Pedagogy: Practical information and suggestions. Journal of Physical Education, Recreation and Dance,58,1,46-54.

Searle, B.W., Mertaugh, M., Read, A., & Cohen, P. (1989). Improving the quality of textbooks in China. Policy, Research, and External Affairs Working Paper, Population and Human Resources Department. Washington, DC:The World Bank.

Stevenson, R.B. (1987). Staff development for effective secondary schools: A synthesis of research. Teaching and Teacher Education,3,3,233-248.

Schmidt, W.H. (1990). School-based, flexible staff specialization. Unpublished manuscript, College of Education, Michigan State University.

Thompson, A.R. & Greenland, J.J. (1981). The implication of UPE for teacher education. International Review of Education,29,2, 199-213.

Vongkeao, B. (1982). The role of the village school teacher in rural development. of the province Mahasarakan Thailand. Unpublished doctoral dissertation, Southern Illinois University.

Vulliamy, G. (1987). School effectiveness research in Papua New Guinea. Comparative Education, 23,2,209-223.

Walberg, H. (1984). Improving the productivity of America's schools. Educational Leadership, 41,8,19-30.

Warwick, D.P. & Reimers, F. (1991). Good schools and poor schools: how they differ. Paper presented at the annual meeting of the Comparative and International Education Society, Pittsburgh.

Wheeler, C.W., Chuaratanaphong, J., Bhumirat, C., Easukkawat, S., Pumsard, S., Sirigirirakal, V., & Sookpokakit, B.(in press). School clusters in Thailand: A management strategy for improving primary school quality. To appear in the International Journal of Educational Research

Weeler, C.W., Raudenbush, S.W., & Pasigna, A. (1988). Policy initiatives to improve primary school quality in Thailand: An essay on implementation, constraints, and opportunities for educational improvement. (BRIDGES Report Series No. 5). Cambridge, MA: Harvard Institute for International Development.

White, R., Conwell, C., & Passe, J. (1991). The effect of equals inservice upon allocated time during mathematics instruction. Journal of Classroom Interaction. 26,1,23-31.

Wood, F.N. & Thompson, S.R. (1980). Guidelines for better staff development. Educational Leadership, 37,374-378.