The Methodological Issues of Variance in Teacher Behavior and Student Achievement: The Relationship of Variance to School and Teacher Effectiveness.

Linda Jane Crone
Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_disstheses/5427

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

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

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

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

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
The methodological issues of variance in teacher behavior and student achievement: The relationship of variance to school and teacher effectiveness

Crone, Linda Jane, Ph.D.

The Louisiana State University and Agricultural and Mechanical Col., 1992
THE METHODOLOGICAL ISSUES OF VARIANCE IN TEACHER BEHAVIOR AND STUDENT ACHIEVEMENT: THE RELATIONSHIP OF VARIANCE TO SCHOOL AND TEACHER EFFECTIVENESS

A Dissertation

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Doctor of Philosophy in

The Department of Administrative and Foundational Services

by

Linda J. Crone
B.Mu.Ed. Indiana University, 1965
M.Mus. University of Arizona, 1978
December 1992
Acknowledgements

I would like to express my sincere gratitude to Dr. Charles Teddlie, who served as my committee chairperson throughout this dissertation. Without his inspiration and guidance, this dissertation would not have been possible. Dr. Abbas Tashakkori, who has not only served as a committee member, but has taught me most of what I know about statistics and data analysis, has also provided invaluable assistance.

I am especially grateful to Dr. Cornelia Yarbrough for inciting my interest in research and helping me to discover "what I wanted to be when I grew up". I also appreciate her service on my committee.

I also appreciate the time and effort that Dr. Eugene Kennedy and Dr. Mic Lang have spent serving on my committee. Their insights on the design and methodology have been especially helpful. Mic, thanks for forcing me to clearly defend my design.

In addition, I wish to express my gratitude to my friend and mentor, Horace Olivier, who has taught me everything I know about data processing, which was essential to my analysis. I also want to thank my good friend, Dr. Janet Rami, who encouraged me and inspired me with her own hard work.
I wish to express my thanks to the Louisiana Department of Education. Some of the issues dealt with in this dissertation were in direct response to work that was done for Betty Kraft's School Effectiveness Study. The research done by my colleagues here at the LDE, has also contributed to my dissertation. A special thanks to Dr. Bobby Franklin for his help.

Finally, I wish to extend a very special thanks to my parents. Their love, support, and belief in me has kept me going.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measures</td>
<td>76</td>
</tr>
<tr>
<td>Issue I: Means and variances of teacher behavior</td>
<td>76</td>
</tr>
<tr>
<td>Issue II: Means and within teacher variance of student achievement</td>
<td>80</td>
</tr>
<tr>
<td>Issue III: Means and within school variance of student achievement</td>
<td>81</td>
</tr>
<tr>
<td>Variance as an Inferential Statistic</td>
<td>81</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>84</td>
</tr>
<tr>
<td>Chapter Four: Results</td>
<td>90</td>
</tr>
<tr>
<td>Overview</td>
<td>90</td>
</tr>
<tr>
<td>Issue I</td>
<td>92</td>
</tr>
<tr>
<td>Limitations of the sample</td>
<td>92</td>
</tr>
<tr>
<td>Review of methods of analysis</td>
<td>93</td>
</tr>
<tr>
<td>Results of quantitative analysis</td>
<td>94</td>
</tr>
<tr>
<td>Results of teacher interviews</td>
<td>97</td>
</tr>
<tr>
<td>Results of principal interviews</td>
<td>102</td>
</tr>
<tr>
<td>Issue II</td>
<td>105</td>
</tr>
<tr>
<td>Review of methods of analysis</td>
<td>105</td>
</tr>
<tr>
<td>Results of analysis on mean student achievement</td>
<td>107</td>
</tr>
<tr>
<td>Results of analysis on within teacher variance</td>
<td>112</td>
</tr>
<tr>
<td>Issue III</td>
<td>115</td>
</tr>
<tr>
<td>Methods of analysis for Issue III</td>
<td>115</td>
</tr>
<tr>
<td>Results of analysis on mean student achievement</td>
<td>117</td>
</tr>
<tr>
<td>Results of analysis on within school variance</td>
<td>124</td>
</tr>
<tr>
<td>Issue IV</td>
<td>129</td>
</tr>
<tr>
<td>Analysis of means of student achievement for high-, middle-, and low-SES Schools</td>
<td>129</td>
</tr>
<tr>
<td>Results of analysis for Issue IV</td>
<td>131</td>
</tr>
<tr>
<td>Chapter Five: Discussion</td>
<td>133</td>
</tr>
<tr>
<td>Summary</td>
<td>133</td>
</tr>
<tr>
<td>Limitations of the Study</td>
<td>139</td>
</tr>
<tr>
<td>Conclusions Regarding Issue I</td>
<td>143</td>
</tr>
<tr>
<td>Conclusions Regarding Issue II</td>
<td>152</td>
</tr>
<tr>
<td>Conclusions Regarding Issue III</td>
<td>154</td>
</tr>
<tr>
<td>Conclusions Regarding Issue IV</td>
<td>157</td>
</tr>
<tr>
<td>Comparison of Findings for Issue II and Issue III</td>
<td>160</td>
</tr>
<tr>
<td>Summary of Variance Results</td>
<td>163</td>
</tr>
<tr>
<td>Implications</td>
<td>166</td>
</tr>
<tr>
<td>Recommendations for Further Research</td>
<td>167</td>
</tr>
</tbody>
</table>
List of Tables

Table 3.1  
Analysis of Research Questions .............................................................. 88

Table 4.1  
Comparison of Means by School Effectiveness Type on Seven Dimensions of Teaching Behavior ........................................................... 96

Table 4.2  
Comparison of Variance in Scores on Teacher Behavior for Effective Versus Ineffective Schools ................................................................. 98

Table 4.3  
Comparison of Mean Student Achievement for Effective, Typical, and Ineffective Teachers ................................................................. 107

Table 4.4  
Results of Analysis on Mean Student Achievement for Teacher Effectiveness Groups and SES ................................................................. 108

Table 4.5  
Comparison of Mean Student Achievement for Middle- and Low-SES Groups ................................................................. 109

Table 4.6  
Comparison of Mean Student Achievement Across Teacher Effectiveness Categories for Middle- and Low-SES ................................................................. 110

Table 4.7a  
Results of Mean Student Achievement Across Teacher Effectiveness Categories for Low-SES Students ................................................................. 111

Table 4.7b  
Results of Mean Student Achievement Across Teacher Effectiveness Categories for Middle-SES Students ................................................................. 111

Table 4.8a  
Results of Student-Newman-Keuls Test for Mean Student Achievement of Low-SES Students by Teacher Effectiveness Group ................................................................. 112

Table 4.8b  
Results of Student-Newman-Keuls Test for Mean Student Achievement of Middle-SES Students by Teacher Effectiveness Group ................................................................. 112
Table 4.9
Comparison of Average Within Teacher Variance of Student
Achievement for Effective, Typical, and Ineffective Teachers . . . . 113

Table 4.10
Comparison of Average Within Teacher Variance of Student
Achievement for Middle- and Low-SES Students ................. 114

Table 4.11
Comparison of Average Within Teacher Variance of Student
Achievement by Middle- and Low-SES Subgroups for Effective,
Typical, and Ineffective Teachers ................................. 115

Table 4.12
Comparison of Mean Student Achievement for Effective,
Typical, and Ineffective Schools ...................................... 118

Table 4.13
Results of Analysis of Mean Student Achievement for School
Effectiveness Groups and SES ........................................ 119

Table 4.14
Comparison of Mean Student Achievement for Middle- and
Low-SES Groups ......................................................... 120

Table 4.15
Comparison of Mean Student Achievement Across School Effectiveness
Categories for Middle- and Low-SES ................................ 121

Table 4.16a
Results of Mean Student Achievement Across School Effectiveness
Categories for Low-SES Students ..................................... 122

Table 4.16b
Results of Mean Student Achievement Across School Effectiveness
Categories for Middle-SES Students ............................... 122

Table 4.17a
Results of Student-Newman-Keuls Test for Mean Student Achievement
of Low-SES Students by School Effectiveness Group ............ 123

Table 4.17b
Results of Student-Newman-Keuls Test for Mean Student Achievement
of Middle-SES Students by School Effectiveness Group ........ 124
Table 4.18  
Comparison of Average Within School Variance of Student  
Achievement for Effective, Typical, and Ineffective Schools . . . . . . . . 125

Table 4.19  
Results of Analysis on Within School Variance of Student  
Achievement for School Effectiveness Groups and SES . . . . . . . . 126

Table 4.20  
Comparison of Average Within School Variance of Student  
Achievement for Middle- and Low-SES Students . . . . . . . . 126

Table 4.21  
Comparison of Average Within School Variance of Student  
Achievement by Middle- and Low-SES Subgroups for  
Effective, Typical, and Ineffective Schools . . . . . . . . 127

Table 4.22a  
Results of Analysis of Within School Variance of Student Achievement  
Across School Effectiveness Groups for Low-SES Students . . . . 128

Table 4.22b  
Results of Analysis of Within School Variance of Student Achievement  
Across School Effectiveness Groups for Middle-SES Students . . . . 129

Table 4.23  
Comparison of Mean Student Achievement by Middle- and Low- 
SES for High-, Middle-, and Low-SES Schools . . . . . . . . 132
Abstract

Three different issues were involved in this study. Issue I examined the means and variances of teacher behavior for teachers in effective schools versus teachers in ineffective schools. Five teachers were evaluated from grade three and grade five of five effective and six ineffective schools (25 teachers from effective schools and 30 teachers from ineffective schools). Observations of seven dimensions of teaching behavior were conducted. In every case, the group means of teacher behavior were higher for the teachers in the effective school category. An examination of the coefficients of variation indicated that the dispersion of scores for the teachers in effective schools was less than for teachers in ineffective schools.

Qualitative interviews with both teachers and principals were conducted to find possible explanations for the differences found in the means and coefficients of variation of teacher behavior. Results indicated that effective schools had better teacher socialization processes, stronger principals, more strictly enforced schoolwide discipline policies, and more thoughtful and thorough means of teacher selection/dismissal.

Issues II and III dealt with the equity concern in education. Issue II was concerned with the means and variances of achievement for different
socioeconomic (SES) levels of students who were taught by effective, typical, and ineffective teachers (60 teachers, 20 in each group). Issue III examined the means and variances of student achievement for students in effective, typical, and ineffective schools (162 elementary schools, 54 schools in each effectiveness category).

The differences in the results of the group means for Issue II compared to Issue III show that the differential effectiveness of the teacher and the school have similar influences on student achievement. There is slightly more differentiation in school effectiveness categories than in teacher effectiveness categories.

When examining variance, the teacher and the school yielded opposite results. For Issue II, the smallest variance was found in the effective teacher group. For Issue III, the effective school group had the largest variance. Effective teachers appear to be realizing the goal of equity, whereas effective schools appear to be widening the dispersion of scores.
Chapter One: Introduction

**Issues of Variance**

The use of variance as the dependent variable has begun to achieve recognition in educational research. Variance has been an integral part of many statistical procedures. Variance, in one of its many forms, provides the conceptual basis for correlation, regression, ANOVA, and many other statistical computations. It is also used as the denominator for both T-tests and Z-tests. Standard deviation and variance are known as descriptive statistics informing us of the dispersion of individual scores around the mean of a group of scores. Recent research has demonstrated that these measures of variability give us more than just the necessary figure to use in the denominator of a T-test.

Industrial literature indicates a growing recognition of the importance of investigating the variability in job performance. A British study on issues of utility analysis states:

A significant advance in industrial psychology over the last decade has been the development of decision theoretical equations for estimating the utility of selection. One of the most important terms in these equations is the standard deviation of performance. (Smith, 1989, p. 189)

A quote from American industry also emphasizes the importance of the standard deviation:
A critical parameter in studies of the economic utility of personnel selection and other personnel programs (such as training) is the standard deviation of employee contributions in dollars. We would like to know how much employees in the same job typically differ in productivity, probably the most important variable in industrial/organizational psychology. (Schmidt & Hunter, 1983, p. 407)

The field of kinesiology also uses variance as the dependent variable. This field is concerned with the consistency (or inconsistency) of an individual’s performance over several trials. Consequently, the variable error, which is the term used for the standard deviation of a subject’s trials around the mean of those trials, is an essential outcome in this field (Schmidt, 1988).

A few early studies were conducted in education where inferences were made concerning variance. Birch and Lefford (1967) compared variance in visual discrimination for children of different age groups. Johnson and Baker (1973) researched the effect of computerized administration versus human administration of a problem solving task. One of the inferences made was that there would be differences in the variance of performance. Hushak (1977) hypothesized that "schools reduce the variance of cognitive skills over time." He made comparisons of variance in word identification, reading vocabulary, and reading comprehension for different grade levels.
Recent studies in educational research are also beginning to look at variance as a significant factor in teacher performance. In their comparison of effective versus ineffective schools, Teddlie, Kirby, and Stringfield (1989) pointed out the small variability in teacher behavior within effective schools. Virgilio, Teddlie, and Oescher (1991) further explored the differences in teaching behaviors of teachers from schools with different levels of effectiveness, and found that teachers in effective schools behaved more similarly than those in typical schools, and those in typical schools behaved more similarly than those in ineffective schools.

Variance in student achievement is another growing issue in education. Bloom (1984) reviewed different methods of teaching in which the result was to both increase mean achievement and decrease variance. Coleman (1985) also conducted a study where variance was an essential factor. This was a cross-national study in which he reported that, although the Japanese were found to have the greatest mean gain in science, they also had the greatest dispersion of scores.

Recognizing the importance of examining variability in achievement, Raudenbush and Bryk (1987) explore the viability of using different hierarchical models to find those school level variables that contribute to differences in dispersion of scores. One premise of their study was that
variance reduction is a goal of education. Lee and Bryk (1989) used HLM modeling techniques to identify specific school characteristics that contribute to "an equitable distribution of achievement across the diverse social class, racial/ethnic, and academic backgrounds of students". (p. 172)

American education is becoming more and more concerned with equity in education (which would be reflected by a comparatively small variance of achievement scores). It is the philosophy of most of the leading educators that all children can and should learn. Levine and Lezotte (1990) recommend that the effectiveness of a school should not even be considered until the data is disaggregated by SES to examine the gap in scores between subgroups within a school. Equity strives for little or no gap of achievement scores, especially between the low- and middle-SES groups of students within a school (Levine & Lezotte, 1990).

The purpose of this study is to examine three different issues regarding means and variances. Issue I examines the means and variances of teacher behavior for effective versus ineffective schools and seeks possible explanations for any differences found. The unit of analysis for this issue is the teacher. Issue II is concerned with the means and variances of different SES levels of student achievement at the teacher level, and examines differences in means and within class variances of student achievement by middle- and low-SES subgroups for effective, typical, and
ineffective teachers. The teacher is the unit of analysis for Issue II. Issue III looks at the means and variances of different SES levels of student achievement at the school level, and compares means and within school variances of student achievement by middle- and low-SES subgroups in effective, typical, and ineffective schools. Issue III uses the school as the unit of analysis.

**Background and Definition of Variables**

(Operational definitions found in Appendix A)

**School effectiveness.**

Until the Coleman report (Coleman et al., 1966) very little research had been done on school effectiveness. Students were required to go to a specific school, generally to the school that was within their own neighborhood. Some teachers had reputations as being "good" or "bad", but the "effectiveness" of an entire school was not often discussed, except perhaps by the very rich who went off to the "best" private schools. Coleman et al. (1966) examined the degree of effect that schools had on student achievement, and concluded that schools made very little difference. They reported that the main factor influencing student achievement was the socioeconomic status and home environment of the students.
Studies then began to emerge which reported schools that did "make a difference" (Edmonds & Frederiksen, 1979; Good & Weinstein, 1986; Weber, 1971). The early studies (Edmonds & Frederiksen, 1979; Weber, 1971) sought out inner-city schools with low-SES and conducted case studies comparing schools which they considered to be "effective" with those considered to be "ineffective". Schools that were achieving at a much higher academic level than other schools with similar attributes were considered effective.

The measure of student achievement scores as an operational definition of school effectiveness has been carried on through most subsequent research. Other variables are occasionally considered, such as dropout rate, student attendance, or pupil teacher ratio; but student achievement (specifically on some type of standardized test) is the most measurable and comparable of variables.

Different algorithms have been used to classify the "effectiveness" of a school based on their achievement test scores. The most commonly used algorithm utilizes regression analysis predicting achievement from socioeconomic status. The student’s actual scores are then compared to the predicted scores. That difference between the actual scores and the predicted scores is called the residual. This residual reflects the amount of
"effect" a school has made on the students within that school. A positive residual indicates that the school has helped the students to attain higher achievement than was expected for students at that particular socioeconomic status. If the residual is negative, it is assumed that the school has had an adverse effect on the students. Cut off points are then established, generally +1 and -1 standard deviations above/below the mean of the residuals, to categorize schools as effective and ineffective (Klitgaard & Hall, 1974; Lang, 1991).

**Teacher effectiveness.**

Teacher effectiveness can also encompass a broad range of definitions. Even though student achievement is sometimes considered, teacher effectiveness is more often defined by specific teacher behaviors which have been identified as attributes that are necessary for a teacher to be "effective". These attributes fall into four different domains: classroom management, classroom instruction, classroom climate, and time-on-task (Virgilio et al., 1991). Research in teacher effectiveness (Brophy & Good, 1986) has found that the classrooms of effective teachers are much better organized than other classrooms. Classroom procedures are conducted in an efficient manner. This includes such activities as handing out and
collection of materials, the movement of students throughout the class, and
the overall management of student behavior.

Brophy and Good (1986) also discussed the findings from literature
concerning teacher instruction. This is not only concerned with knowledge
of the subject matter, but includes instructional techniques of focusing on
the lesson, summarizing the lesson, and proper question and answer
procedures (i.e. a sufficient amount of questions, the appropriate type of
questions, and proper wait time).

The climate of a classroom has been found to play an essential part in
the student's ability to learn. Climate deals with the actual physical
attributes of the room (a cheerful, pleasant looking room with sufficient
lighting, appropriate temperature, etc.), and also includes the personality of
the teacher and the emotional atmosphere that is created by his/her
presence (Gage, 1965).

Time-on-task is another important variable for teacher effectiveness.
In their Follow Through Evaluation Study, Stallings and Kaskowitz (1974)
examined correlations between academic achievement and different process
variables. They found that time spent on academics provided the clearest
pattern of positive relationships to achievement. Time spent on
nonacademic activities yielded negative correlations with academic
achievement. In the California Early Childhood Education Study, Stallings, Fairweather, and Needels (1977) reported that classes made greater academic gains when more time was spent on academic activities.

While examining teacher behavior in effective versus ineffective schools, Teddlie et al. (1989) used qualitative methods to find teaching behaviors that could effectively discriminate between teachers in the two types of schools. Using guidelines from the literature, they came up with ten indicators, nine of which demonstrated significantly more effective behavior for teachers in effective schools. Those indicators included time-on-task, presentation of new material, independent practice, high expectations, positive reinforcement, minimal interruptions, discipline, friendly ambience, and appearance of the room.

Socioeconomic status.

Even though there is now sufficient evidence to conclude that schools do "make a difference" (Good & Weinstein, 1986), it is still true that, to a great extent, student achievement does depend on the home environment or socioeconomic status of the students. When regression analysis is used to classify schools as effective or ineffective, socioeconomic factors are found to be the strongest indicator of academic achievement (Teddlie et al., 1989).
Most of the early school effectiveness studies were conducted in urban, inner-city schools (Edmonds & Frederiksen, 1979; Rutter, Maughan, Mortimore, Ouston, & Smith, 1979; Weber, 1971). Since there was no evidence to suggest that the findings of these studies would generalize to other SES levels, socioeconomic status became a contextual concern (Good & Brophy, 1986). Studies began to emerge which compared both school and teacher behavior for different SES levels. Although some of the indicators have been found to be consistent over different SES levels, evidence demonstrates that there are many different techniques which are needed to effectively teach students of different SES levels.

Hallinger and Murphy (1986), and Teddlie, Virgilio, and Oescher (1990) both conducted studies which compared schools with different SES levels, and reported similar results. The major contextual differences found in both studies were in regards to expectations, principal style, rewards, and curricular offerings.

In the classroom, some teacher behaviors that were found to need adjustments for different SES levels were pacing, question and answer techniques, the type and amount of work provided, and the amount of praise (Brophy & Evertson, 1976).
Measures of school effectiveness.

As a general rule, school effectiveness studies have mostly used some form of student academic achievement as the indicator of school effectiveness. State level "school report cards" are beginning to consider other variables besides achievement scores as indicators of a school's worth. Variables such as dropout and attendance data, and percent of certified teachers are being included in state reports (Louisiana Department of Education, 1991).

The achievement scores, though, are the outcomes which are of primary interest to most policy makers. Many of the states are now using achievement scores as the criterion for incentive award programs; Louisiana, South Carolina, and California are examples of the states which are now providing awards for school academic achievement. According to the National Education Goals Panel (1991), education in the US is going to become much more outcome oriented, with nation-wide testing for student achievement.

Achievement scores can be either national norm-referenced test or criterion-referenced test. There are pros and cons for using either one. The NRT provides a better indication of how schools are performing according to national standards. Criterion-referenced tests are often
constructed to measure the attainment of state or local curriculum guides and are considered to be a better indicator of the school’s ability to teach what it purports to teach.

**Variance.**

Measures of variation, as defined by Hinkle, Wiersma, and Jurs (1988, p.56) "are lengths of intervals on the scale of measurement that indicate the variation, or spread, of scores in a distribution." The most commonly used measures of variation are the variance and standard deviation. "The variance is defined as the average of the sum of squared deviations from the mean" (Hinkle, et al., 1988, p.58). "The standard deviation is the square root of the variance. It is expressed in the same units as the original measurement of the variable" (Hinkle et al., 1988, p.61).

In addition to the actual variance, a standardized measure of variation is sometimes used when different variances are being compared. This standardized variance is called the coefficient of variation and is computed by dividing the standard deviation by the mean (Kazmier, 1978) or dividing the standard deviation by the mean and multiplying by 100 (SAS Institute Inc., 1985). This creates a unitless measure of variation which makes comparison of variances more accurate.
An examination of these different types of variance can provide valuable information. Variance in teacher "effectiveness" is just as important an issue to education as the variance of employee performance is to industry. Geske & Teddlie (1990) point out that, despite the fact that our raw materials (the students) are beyond our control, and the output is not directly measured in dollars and cents, we still need to be concerned with the productivity of our schools. There are some educational studies that have examined variance of teacher productivity in differentially effective schools.

Two studies on teacher behavior in effective/ineffective schools observed the smaller variance of teacher behavior in effective schools (all teachers being grouped towards the "effective" end of the scale). The question was then posed as to what factors within the school caused this truncation of variance (Teddlie et al., 1989; Virgilio et al., 1991).

Virgilio et al. (1991) suggested that the cause of the absence of ineffective teachers in effective schools could be explained by either teacher selection or teacher socialization. Teacher socialization was believed to be the more influential factor as Teddlie, Falkowski, Stringfield, Desselle, and Garvue (1984) had found that only 15% of the principals in effective schools felt that they played the primary role in
teacher selection (in the areas researched, the central office made the hiring decisions).

Variance of student achievement is also an important concern in education. Indirectly, variance of student achievement has been the focus of one of the major issues in American education: that of equity versus efficiency. Where those who advocate efficiency are concerned with teaching the majority of students in the most efficient manner, the proponents of equity feel that it is the responsibility of the schools to bring up the scores of the low-SES students to more closely reflect the scores of the entire population.

This concern for equity in education began with Ron Edmonds in 1979. Equity deals with the opinion introduced by Edmonds: "that all children, excepting those of certifiable handicap, are educable, and the behavior of the school is critical in determining the quality of that education" (Edmonds & Frederiksen, 1979). They believed that to call a school effective, not only should the mean performance level be high, but there should be little or no gap between the achievement of the working class and middle class of students. If this were the case, the variance of student achievement within a school should be comparatively small. Conversely, if a school is not reaching the different subgroups within the
school, one would find larger gaps between subgroups, hence, a larger within school variance.

The only study that has directly examined variance of achievement scores for effective and ineffective schools reported finding just the opposite effect. Lark, Bluust, and Coldiron (1984) conducted a study on variance of achievement scores in the Pennsylvania schools. They classified schools as effective/ineffective using a regression model, and found that those schools which were rated as effective actually had greater variance than the ineffective schools. They also found that those schools classified as effective using the regression algorithm had larger gaps between the subgroups than those classified as ineffective.

The findings of a study by Crone & Tashakkori (1992) may clarify this issue. In examining the variance of student achievement for a large sample of eighth graders using the NELS (National Education Longitudinal Study) Data tapes, it was found that variance issues could only be properly addressed when examined in conjunction with the socioeconomic status of the students. Little difference was found in within school variance of schools with different effectiveness levels until the sample was divided into upper and lower SES levels. The schools that were in the top 30% of SES and those in the bottom 30% were analyzed separately. It was found
that the upper SES schools had the smallest variance in the effective schools. Conversely, the lower SES schools had the smallest variance in the ineffective schools.

It is difficult to conjecture as to whether or not these results would apply to the elementary school students. Motivational factors for the different ages may yield quite different results. These findings, though, do point towards a need to examine variance of student achievement in the context of SES.

**Research Questions**

The purpose of Issue I was to compare the means and variances of teacher behavior of teachers in effective schools versus teachers in ineffective schools. If differences were found, then possible causes for these differences were looked for by using qualitative methods of research (teacher and principal interviews). The independent variable for Issue I is school effectiveness, and the dependent variable is teacher behavior.

The research questions which emerged from Issue I are:

1. Is there a difference in teacher effectiveness in effective elementary schools versus ineffective elementary schools?

2. Is there a difference in variance in teacher effectiveness in effective elementary schools versus ineffective elementary schools?
3. If a difference does exist in teacher effectiveness in the two school effectiveness categories, what contributed to this difference?

Issue II involves an examination of the student achievement of the students taught by the teachers for Issue I. These teachers were rated as effective, typical, or ineffective based on the observed behavior, and comparisons made of the different SES levels of student achievement for the teacher effectiveness groups. Of interest in this study was the comparison of the means of student achievement for the SES subgroups taught by the differentially effective teachers. Also, it was desired to examine the dispersion of student achievement scores for these teachers. The main intent here was to determine whether effective teachers are able to lessen the gap in achievement between SES subgroups within their class. The independent variables are the three teacher effectiveness categories, and the two SES levels of students within each classroom. The mean of student achievement is the first dependent variable for Issue II. The second dependent variable for this issue is the within class variance of student achievement. The reason for the examination of the interaction is to determine whether the different subgroups are getting differential treatment by the teacher effectiveness groups.
The research questions for Issue II are:

4a. Is there a difference in student achievement for effective, typical, or ineffective teachers?

4b. Is there a difference in student achievement for different SES subgroups of students?

4c. Is there an interaction between teacher effectiveness and SES subgroups?

5a. Is there a difference in variance of student achievement for effective, typical, or ineffective teachers?

5b. Is there a difference in variance of student achievement for different SES levels of students?

5c. Is there an interaction effect on variance of student achievement?

Issue III examines student achievement as it is affected by the effectiveness of the school. The means and variances of student achievement for middle- and low-SES students were compared for three different school effectiveness categories. As in Issue II, the main interest was to determine how equitably the different subgroups within each school are being taught. Again, the examination of the interaction effect will address this concern. School effectiveness categories and SES subgroups are the independent variables, with the means of student achievement being
the dependent variable. A second dependent variable for this issue is the within school variance of student achievement.

The research questions for Issue III are:

6a. Is there a difference in student achievement for effective, typical, or ineffective schools?
6b. Is there a difference in student achievement for different SES subgroups of students?
6c. Is there any interaction effect between school effectiveness and SES levels?
7a. Is there a difference in variance of student achievement for effective, typical, or ineffective schools?
7b. Is there a difference in variance of student achievement for different SES levels of students?
7c. Is there any interaction effect on variance of student achievement?

The Hypotheses

Questions 1, 2, 4a, 4b, 4c, 5a, 5b, 5c, 6a, 6b, 6c, 7a, 7b, and 7c form the basis for the hypotheses for this study. Question 3, being qualitative in nature, is not subject to hypothesis testing.
Issue I, question 1 states that teachers in effective elementary schools demonstrate higher levels of teacher behavior than teachers in ineffective schools.

Issue I, question 2 contends that teachers in effective elementary schools have less variance of teacher behavior than teachers in ineffective elementary schools.

Issue II, question 4 is: (a) students taught by effective teachers have higher student achievement than those taught by typical teachers, and students taught by typical teachers have higher achievement than those taught by ineffective teachers (b) middle-SES students have higher means of student achievement than low-SES students, and (c) an interaction between teacher effectiveness and SES levels exists.

Issue II, question 5 states that: (a) differences occur in within class variance of students being taught by effective, typical, or ineffective teachers, (b) differences exist in within class variance of student achievement for different SES subgroups of students, and (c) an interaction between school effectiveness and SES levels does occur.

Issue III, question 6 is that: (a) students in effective schools have higher means of student achievement than students in typical schools, and students in typical schools have higher means of student achievement than
students in ineffective schools, (b) mean student achievement for middle-SES students is higher than mean student achievement for low-SES students, and (c) there is an interaction effect.

Issue III, question 7 states: (a) there is a difference in within school variance of student achievement for students in effective, typical, or ineffective schools, (b) there is a difference in within school variance of student achievement for different SES levels of students, and (c) there is an interaction between the two independent variables.

Significance of the Study

The issue of means and variances of teacher behavior in effective versus ineffective schools is, to some extent, a replication of Teddlie et al. (1989) and Virgilio et al. (1991). Both studies reported finding less variance of teacher behavior in effective schools. When this is coupled with higher mean behavior scores, as found in both studies, this would imply that there are less ineffective teacher behaviors in effective schools. This study sought to carry these findings a step further, and investigate the possible causes for the reduction of ineffective teaching behaviors.

As mentioned before, Virgilio et al. (1991) believed that this truncation of variance in effective schools could be caused by either teacher selection or teacher socialization. The process of
selection/dismissal certainly needs to be compared for effective/ineffective schools. Do the principals do a better job of hiring in effective schools? Do the good schools, because of their reputation, attract good teachers? Are ineffective teachers dismissed from effective schools? Or does something happen once they enter the school to make the teachers more effective, and if so, what happens?

There are many school effectiveness variables that could have an effect on teacher behavior. Geske and Teddlie (1990) pointed out that the productivity of a school is, to a large extent, dependent on the school leadership (generally by the principal). They gave three areas in which school leadership has an impact on school effectiveness: (a) goal orientation, (b) instructional leadership, and (c) assignment of students and teachers to classrooms.

A teacher's knowledge and ownership of the school goals could provide the teacher with more direction and individual goals. This could be especially advantageous for improving teacher behavior if the school goals were academically oriented.

The principals in effective schools are found to make more frequent visits to the classrooms, give more feedback to the teachers regarding their behavior, and take a more active part in curriculum development and
implementation (Virgilio et al., 1991). In a recent PBS documentary highlighting four exceptionally good schools in the US (MacNeil/Lehrer Productions, 1990) Roger Mudd said of one principal: "He is not the kind of leader who spends a lot of time in his office." The principal himself stated: "I do a lot of managing by walking around. I like the opportunities to interact with the kids."

A principal can also have an influence on the variability of productivity of teacher behavior by the thoughtful assignment of students and teachers to classrooms (Virgilio et al., 1991; Geske & Teddlie, 1990). For example, separating students who together may create disruptive behavior would have an effect on both time-on-task and classroom management.

One other school effectiveness variable which could possibly have an effect on teacher behavior is that of school climate. Does the school have a clean, pleasant, cheerful atmosphere? Is it a friendly, enjoyable place to be? The school climate can certainly affect the teachers’ emotional state, and positively or negatively affect their attitude toward teaching and toward their students.

Qualitative methods were employed to examine the possible causes for these differences in variance of teacher behavior. Individual teacher
interviews were conducted using open ended questions. The interview form included sections on teacher recruitment/dismissal, teacher socialization, principal’s instructional leadership, goal orientation and assignment of students to classrooms. This adds to the growing literature on the process that makes teachers more homogeneously effective in effective schools.

The issue of variance of student achievement for effective and ineffective schools has produced inconclusive and conflicting data. Those educators who subscribe to the equity issue consider that a school is not actually effective, regardless of its mean achievement, unless the subgroups within the school are being effectively taught (Edmonds & Frederiksen, 1979; Levine & Lezotte, 1990).

In discussing the goals for the year 2000, the National Education Goals Panel (1991) states:

The academic performance of elementary and secondary students will increase significantly in every quartile, and the distribution of minority students in each level will more closely reflect the student population as a whole. (p. 44)

Hence, to be effective, a school would need to have a comparatively small variance in addition to high achievement. The most commonly used method of classifying schools for research purposes is the regression method, which generally uses aggregated school level achievement scores.
These could possibly mask large differences between subgroups. An investigation of the variance of those schools which are classified as effective/ineffective by regression methods would give a better indication of whether all students are being reached.

The only known published study that deals directly with variance of student achievement supports the efficiency philosophy (opposed to the equity philosophy) of education. In explaining the findings of greater variance in schools classified as effective by the regression model, Lark et al. (1984) express the belief that the purpose of the schools is to maximize each child's potential, and in so doing, the spread in student achievement would become greater. It is possible that Lark et al. (1984) looked at the issue in too simplistic a way. Variance of student achievement must be considered as it relates to other variables, such as variance of SES, mean achievement scores, and residuals.

Crone & Tashakkori (1992) found that for eighth grade students the variance patterns differ for different SES levels. The high-SES schools had the smallest variance of student achievement in the effective schools, this dispersion of scores gradually increased as the school effectiveness decreased. The low-SES schools had the largest within school variance of student achievement in the effective schools, this gradually decreased as
the schools became less effective. This may or may not generalize to elementary school. Elementary students are not as capable of independent work and learning as eighth graders. It is possible that by eighth grade, unless the students are extremely uplifted by an effective school, or extremely suppressed by an ineffective school, they may be learning whatever they can on their own. This may not be as easy for elementary students to do.

One could also logically argue that the same pattern in variance would be expected to occur for student achievement as was found in teacher effectiveness. If the ineffective schools have larger variances, that must indicate that there are at least some effective teachers within those schools. The students of those few effective teachers would be expected to attain higher achievement scores than the rest of the students, therefore widening the variance. If this were the case, then the ineffective schools would be expected to have larger variance than the effective schools. Using schools that contain a representative sample of both middle- and low-SES, this study seeks to provide evidence to determine whether or not effective schools are more equitably reaching all students.

Variance in student achievement for effective, typical, and ineffective teachers poses all of the same questions as the issue of variance in student
achievement for effective, typical, and ineffective schools. Are the "good" teachers able to teach in a manner so as to reach students of differing SES within the same classroom? Does effective teaching maximize each student's potential and broaden the variation? Or are the effective teachers able to minimize the gap in student achievement and create equity in the classroom? An examination of the means and variances in student achievement for teachers with different levels of effectiveness will address this issue.
Chapter Two: Review of Literature

Overview

The review of literature will begin with a section on school effectiveness, and will then discuss the literature on teacher effectiveness. This will be followed by a review of studies involving variance of teacher behavior. Next, the chapter will review measures of effectiveness. The final section will include a discussion of context issues.

School Effectiveness

There is much controversy as to the operational definition of school effectiveness. The most commonly used measure is an average school score for student achievement on standardized tests. Some argue that this is an extremely limited view (Purkey & Smith, 1983) which gives no indication of a school’s ability to teach other important topics such as computer skills, writing, or problem solving skills. Other indices which are often used in school incentive programs are student attendance, teacher attendance, and dropout rates, for grades 7-12 (Louisiana Department of Education, 1991).

There are also many concerns with the equity of using an overall school average (Levine & Lezotte, 1990). It is feared that an average may mask the performance of minorities, low-SES students, or the low ability students within a school.
Once an average has been obtained, there is even further confusion concerning the many different procedures used to label a school as "effective". For the purpose of this study, "school effectiveness" will refer to the academic achievement scores of the students within a particular school. The most common means of measuring achievement for effectiveness classification is with the use of regression residual scores, but other measurement and equity issues will be discussed in this chapter.

Until the mid-1960's, virtually nothing substantial had been done in respect to research in school effectiveness. The well known Coleman report (Coleman et al., 1966) stated that the major factor contributing to student achievement was that of family background and that schools make little difference in student achievement. It appears that many researchers disagreed with this conclusion, as studies began to emerge which sought to provide evidence of schools which did make a difference. The studies during the late sixties, though, concentrated on trying to link student achievement to easily attained variables such as class size, teacher salaries, the number of library books, the book series used in class, or new school buildings (Purkey & Smith, 1983). None of these variables were found to have much relationship to achievement.
The literature in the 1970's was basically involved with the identification of effective schools. Weber (1971) was one of the first to attempt to refute Coleman's findings, and also to examine the processes involved in effective schools. He concentrated on finding successful reading instruction taking place in inner-city schools. Four schools were identified which became the focus of a case study. In looking at the process involved, Weber determined common characteristics for these schools. These included strong leadership; high expectations; orderly, pleasant atmosphere; emphasis on basic skills; and evaluation of progress. Even though the sample was small and there was no comparison group, these traits have been confirmed by subsequent studies.

There are two studies (Brookover, Beady, Flood, Schweitzer, & Wisenbaker, 1979; Klitgaard & Hall, 1974) that utilized large samplings of schools to determine whether effective schools do exist. Klitgaard and Hall (1974) examined three large data sets: one was a sample of Michigan state schools, another from New York City, and the third was from Project Talent high school data. Using background factors as predictors, they analyzed the distribution of residuals, and came up with their definition of an "effective school": a school that achieved at least one standard deviation above the mean of the residuals. In the Michigan study about
9% of the schools were found to be effective. In New York the unit of analysis was the school district and 30 out of 627 districts were found to be effective. Even though the percentage was very small, the findings did indicate that effective schools do exist.

Another large study of a sample of Michigan schools was conducted by Brookover et al. (1979) which also concluded that schools do make a difference. They found that schools with the same input variables had very different outcomes. In comparing schools with the same SES levels, there was as much as one standard deviation difference in mean achievement scores.

Rutter et al. (1979) also found significant school effects in the London secondary schools. They investigated the process involved in effective schools and found that factors associated with student achievement were an academic focus, teachers' expectations, time actually spent teaching, and the degree of teacher interaction with the entire class.

In the elementary schools of Detroit, Edmonds and Frederiksen (1979) determined the process variables involved in effective schools to be strong leadership, high expectations, orderly atmosphere, emphasis on basic skills, and frequent monitoring of student progress (this confirmed Weber's findings). The authors made no claim to these results being generalizable
to different geographic areas, SES levels, or grade levels. By this time, though, the schools were under a great deal of pressure, by both school administrators and state departments, to improve achievement scores. With all involved eager to find a recipe for improving achievement scores, the findings of Edmonds and Frederiksen (1979) were embraced as the "five correlates of effective schools".

Other studies produced other "correlates" of effective schools. Purkey and Smith (1983) reviewed some of the most salient school effectiveness literature of their time. They reported that the most common characteristics of school effectiveness which were found in outlier studies were better control or discipline, and high staff expectations of student achievement. In case studies (Brookover & Lezotte, 1979; California State Department of Education, 1980; Glenn, 1981; Levine & Stark, 1981) researchers reported that findings centered around strong leadership, high expectations, clear goals, strong academic emphasis, and monitoring of student progress.

Program evaluation studies (Armor et al., 1976; Doss & Holley, 1982; Hunter, 1979; Trisman, Waller, & Wilder, 1976) yielded common characteristics of effective schools as being school site-management, instructional leadership, staff stability, curriculum articulation and
organization, school-wide staff development, parental involvement and support, school-wide recognition of academic achievement, maximized learning time, and district support.

Levine (1982) concluded that the successful inner-city schools in Chicago were characterized by well aligned curricular objectives, staff that paid attention to the pacing of instruction, teaching of higher-level cognitive skills, increased quality of homework, emphasis on parental involvement, and supportive, skilled administrators. In a study of London elementary schools, Mortimore and Sammons (1987) found 12 traits that they felt made important contributions to effective schools. They included such characteristics as purposeful leadership of staff by principal, involvement of teachers in curriculum planning and development, consistency among teachers, intellectually challenging teachers, and work centered environment.

Levine and Lezotte (1990) provide an update of the correlates in their review and analysis of effective schools research. The characteristics which they found were the most often correlated with unusually effective schools were: (a) productive school climate and culture, (b) focus on student acquisition of central learning skills, (c) appropriate monitoring of student progress, (d) practice-oriented staff development at the school site,
(e) outstanding leadership, (f) salient parent involvement, (g) effective instructional arrangements and implementation, and (h) high operationalized expectations and requirements for students. They point out that there are many other correlates that have been cited by researchers. It is not to be assumed that these characteristics are unimportant, but simply that there has not been sufficient replication to provide empirical evidence of their importance.

It is evident that there are some similarities (need for leadership, an environment conducive to learning, high expectations), but far from total agreement on the recipe for creating effective schools. It is no wonder that Cuban (1983) issued "A friendly but cautionary note" on effective schools. He feels that schools that may try to implement these findings need to consider four problems with the research: (a) all the descriptions of effective schools didn't tell how to create one, (b) the language is fuzzy—what is "climate" or "leadership"? (c) most effectiveness studies concentrate on nothing but academic achievement, and (d) most research was context specific. Levine (1982) also expressed concern about the implementation of research findings, but felt that the problem was not in what changes should be made, but in the school's desire and ability to carry out these changes.
There are reports, though, of projects that made quite noticeable improvements using guidelines based on findings from the literature. Milwaukee’s Project Rise is one example. McCormack-Larkin and Kritek (1983) report the results of this project, which began as an attempt to improve 18 elementary and two middle schools which had the lowest achievement scores of all the schools throughout the city. Based on the literature, they developed Project Rise upon the beliefs that all students could learn basic skills, that the underachievement of low-SES and minority students was caused by inappropriate school expectations, and that the literature did provide concrete practices that could be used to improve achievement. They worked on improving school climate, curriculum instruction, supportive services, evaluation, and parent support. A great deal of emphasis was put on working with the teachers to change their attitudes about the abilities of low-SES students. The two year results showed a substantial increase in math scores (almost reaching the city average) and some increase in reading scores. The project was deemed successful enough that the procedures were adopted for the entire city.

New York City’s School Improvement Project was based on Edmond’s correlates of effective schools. Clark and MacCarthy (1983) report
encouraging results, especially for those schools where the project had strong principal supervision and coordination.

Sirotnik (1985), and Good and Brophy (1986) stress the fact that school improvement programs can only be successful if they are tailor-made to the particular schools and their specific circumstances. They also emphasize that teachers and staff must have ownership in the project by being involved in the innovations and change. In their review of literature, Good and Brophy (1986) also express the concern that most school effectiveness research to that time had been conducted in urban schools, was based largely on student performance on standardized achievement test, and that more research needed to be done on school-level processes.

Lezotte and Taylor (1990) conclude that there is now a framework for the process of creating effective schools which is generalizable to all schools, regardless of the context. They published and summarized eleven case studies of school improvement processes (the word "process" is used rather than "program" as all studies emphasize the fact that school improvement is an ongoing process rather than a short term program). The most commonly used correlates for improving schools were a clearly stated mission, strong instructional leadership, a climate conducive to learning, high expectations for students and staff, student performance outcome
measures, and data-driven improvement planning. Mentioned in almost every case study was the commitment that all students can and will learn. Other characteristics which were evident in many of the case studies were parental and community involvement; commitment of staff to the educational program; good communication between staff, administrators, and parents; shared decision making; and disaggregation of data.

The disaggregation of data reflects the commitment that these school districts placed on the importance of both quality and equity in their schools. This is in harmony with the National Center for Effective Schools’ (Levine & Lezotte, 1990) definition of school effectiveness, which states that in order to be effective a school should be successful in teaching the intended curriculum of basic skills to all students. They believe that a school should not only have a high level of achievement, but the achievement level should not vary significantly for different subsets of students. Some states are now requiring school districts to disaggregate data to see if all subgroups are being effectively taught.

Other criticisms of school effectiveness research are that none has examined the behavior of individual teachers from one school to another, and that little longitudinal work has been done to examine the stability of school ratings. The Louisiana School Effectiveness Study (LSES) sought
to address these issues (Stringfield & Teddlie, 1988). To date, four phases of this study are complete. Phase I was a pilot study using a modified version of Brookover's 1979 school climate scales. Phase II was based on the premise that different SES levels may require different teaching methods and school climates. This study looked at 76 randomly selected schools (effective, typical, and ineffective) within two SES levels (middle and low). They used both regression and ANOVA techniques to classify schools into different effectiveness status levels. Phase III (Teddlie et al., 1989) controlled for SES and added teacher effectiveness as a variable. This 16 school case study used a sample representing urban, suburban, and rural schools. To be classified as effective a school must have exhibited achievement significantly above that which would be expected for two years in a row. Phase IV looks at the stability of school effects on academic achievement over a five year time period for the 16 schools from Phase III. The contextual findings from LSES Phase II and III will be discussed below.

Most recent school effectiveness research has emphasized considerations of the human aspects of schooling (as opposed to the physical aspects). The importance of a strong principal in initiating improvements in a school is stressed by Hall, Rutherford, Hord, and
Huling (1984). Stringfield and Teddlie (1988) found more active, involved principals in lower SES effective schools. Levine (1986) believes that the most important trait of school reform is that of improving the quality of teachers, which must be done by changing the school culture and environment. Witte and Walsh (1990) found in the Milwaukee public schools that, although the variables having the highest correlation to achievement scores were SES and school location, school environment, parental involvement, and teacher control were significantly related to both language arts and math scores at all grade levels.

Good and Brophy (1986) called for studies that help us to understand the qualitative aspects of schooling. Researching variables such as school culture and environment call for more qualitative methods of analysis. Open ended questions can elicit more information than a simple response on a Likert scale. If worded properly, open ended question can also avoid the socially desirable answers that people are likely to give on a questionnaire. The qualitative methods of observation are also advantageous in researching such factors as principal’s leadership style, school culture or environment. If we go into schools with simple closed-ended questions to be answered we may miss some very important factors. The case study approach yields very valuable information when we are
searching for what makes schools effective. The best way of approaching this issue using the case study approach is to compare effective schools to ineffective schools and determine what qualities are inherent in one group or the other.

Teacher Effectiveness

Although few researchers have linked school effectiveness and teacher behavior, teacher effectiveness research has yielded some common results which are consistent with the school effectiveness findings. The importance of high academic expectations has been mentioned in many teacher effectiveness studies (Powell, 1980; Hathaway, 1983; Brophy & Good, 1986). There is very little question that, at least in elementary schools, time spent on academic activities has a direct correlation with academic achievement (Stallings & Kaskowitz, 1974; Good & Grouws, 1975; Brophy & Evertson, 1976; Powell, 1980; Levine & Ornstein, 1989). Virgilio et al. (1991) found, though, that high time-on-task was not an essential factor for high academic achievement in junior high schools.

Another oft mentioned correlate of effective teaching which coincides with the school effectiveness research is that of an environment conducive to learning, which involves both positive classroom management and

There are some traits found in teacher effectiveness studies that have not been frequently examined in school effectiveness studies. One of these is that of interactive teaching, rather than students being engaged in seat work (Brophy, 1982; Brophy & Good, 1986; Stallings & Kaskowitz, 1974). Another commonly found characteristic of effective teaching is that of faster pacing (Brophy, 1982; Brophy & Good, 1986; Good & Grouws, 1975). The pacing issue, though, seems to be context specific. Brophy and Evertson (1976) discuss the importance of slower pacing for low-SES students. It is also suggested (Tobin, 1980; Tobin & Capie, 1982) that fast pacing is best for basic skills in the lower grade levels, but when dealing with higher level skills in the upper grades, some wait time is needed after teacher's questions and before their responses.

The importance of appropriate academic feedback to high achievement gains is stressed in many studies, including the Canterbury Studies (Wright
Brophy and Evertson (1976) conducted a study in which they identified teachers who were stable in their effects on student achievement. They then compared those teachers who produced high achievement to those who produced low achievement. They found that teachers who produced high achievement were businesslike, task oriented, spent the most time on academic activities, had a positive attitude assuming everyone can learn, were persistent in working with the slow students, monitored the entire class, and moved around the room during seat work assignments. Those teachers with the lowest student achievement scores were more concerned with personal relationships and affective objectives than with academics. They appeared to be disillusioned teachers who disliked students and concentrated on exercising their authority. Other findings from this study are context specific and will be discussed in the section on context issues.

In addition to the previously mentioned findings by Good and Grouws (1975), they also reported that teachers with high achieving students had more teacher-student interaction on academics, and gave clearer instructions than their counterparts with low achieving students. This
study by Good and Grouws was conducted in fourth grade mathematics classes.

As part of the California Beginning Teacher Evaluation Study, an ethnographic study was conducted of 20 second grade and 20 fifth grade students. At each grade level 10 teachers were considered to be "more effective", and 10 were considered to be "less effective". Tikunoff, Berliner, and Rist (1975) reported the findings from this ethnography. They found some characteristics which were consistent over the two grade levels and over subject matter (reading and mathematics). The effective teachers were characterized by an enjoyment for teaching, a polite and pleasant attitude, high expectations, frequent monitoring of progress, knowledge of the subject matter, commitment to instruction of the subject matter, pacing appropriate to the group, and respectful treatment of the students.

There are some characteristics of teacher effectiveness which tend to demonstrate a curvilinear relationship. Although appropriate feedback, teacher student interaction, and respect for students are all characteristics which are found to be positively correlated with student achievement, Coker et al. (1980) found negative correlations to student achievement for overemphasis on praise and rewards, overemphasis on eliciting and
responding to student questions, and overemphasis on student input in
decision making.

Rosenshine (1983) concluded from the teacher effectiveness research
that students taught with structured curricula do better than those taught
with a discovery method. This structure is found to be especially
important with younger, slower, less experienced students. Students also
benefit from having the instruction come directly from the teacher.
Rosenshine also emphasized the importance of reviews--daily, weekly, and
monthly--with reteaching when necessary.

Brophy and Good (1986) reviewed and summarized the literature on
teacher effectiveness and found that those characteristics of teaching which
consistently correlated with student achievement included such traits as the
opportunity to learn, time allocated to academics, expectations, classroom
management, work appropriate to the level of the student, active teaching,
clarity of teaching, appropriate questioning techniques, and appropriate
feedback.

As early as 1979, it was demonstrated by Good and Grouws that the
findings from research could be successfully implemented in the
classroom. They trained teachers on key instructional behaviors for
teaching mathematics, and compared the trained teachers to a control
group. The teachers who received the training exhibited significantly higher behavior on conducting reviews, checking homework, actively engaging students in seatwork and making homework assignments. The students of the teachers who received the training scored significantly higher on their mathematics test scores.

**Variance in Teacher Effectiveness**

Variance issues have recently been receiving the attention of educational researchers. The results of LSES-III (Teddlie et al., 1989) indicated that teachers in effective schools not only displayed different behavior from teachers in ineffective schools but also had less variation of behavior. The teaching behaviors that showed the biggest differences in variance were time-on-task, expectations, discipline, and display of student work.

Virgilio et al. (1991) further explored the differences in teaching behaviors of teachers from effective versus ineffective schools and found that teachers in effective schools behaved more similarly than those in typical schools, and those in typical schools behaved more similarly than those in ineffective schools. This trend was most evident in elementary schools, but was also found in junior high school. The difference in range was caused by the ineffective schools having more low scores at the
bottom end of the scale. The authors theorized that this truncation of range is caused by a combination of teacher selection and teacher socialization. In the areas researched, the district office was reported to make most of the teacher selection decisions, so it is believed that the major difference in variation of teacher effectiveness is due to teacher socialization within the schools.

One area of socialization, teacher induction, was investigated by Kirby (in press). Results indicated that effective schools were more supportive of beginning teachers on all but one out of 14 items on a beginning teacher questionnaire. This questionnaire looked at three areas of socialization: assistance, monitoring, and team-building. This study reported less variance in effective schools for every item on the beginning teacher questionnaire.

Besides teacher induction, other possible causes for less variation in teaching behavior in effective schools may be attributable to the principal’s impact on instructional leadership, goal orientation, and assignment of students and teachers to classrooms (Geske & Teddlie, 1990). The implementation of new curricular programs was found to be highly correlated with principals’ change facilitator style (Hall et al, 1984): the more effective the principals were as change facilitators, the more new curricular programs were found. This indicates that the more effective
principals are more likely to get teachers to try new practices, which would replace older, possibly ineffective practices.

As discussed in the above section on school effectiveness, clearly defined academic goals are an often stated characteristic of effective schools. An effective principal plays a major part in setting the academic goals for the teachers and students (Virgilio, et al., 1991).

The careful and thoughtful assignment of students to classes is discussed by Good and Brophy (1986) as an important factor in school effectiveness. Assigning the teacher to a specific class, taking into consideration the characteristics of the class and the strengths and weaknesses of the teacher, is equally important.

Another possible explanation for the variance differences in effective and ineffective schools is what, if anything, the principal does with ineffective teachers. Bridges (1986) presents the scenario of how schools get "stuck" with incompetent teachers. This usually begins with incompetent principals who fail to properly evaluate their teachers. These teachers become tenured, and then it is very difficult to legally dismiss them. A new conscientious principal is assigned to the school, and refuses to keep these teachers; consequently, the simplest alternative is to transfer
them to another school. Bridges stresses the extreme importance of frequent evaluation of teachers, especially new teachers who may most benefit by intervention and remedial assistance. One study (Stringfield, Teddlie, & Suarez, 1985) reported an effective principal as saying "I try to work with them. There are many things we can teach teachers." In two cases, though, this principal had to get rid of unsatisfactory teachers. One teacher transferred, the other was not recommended for tenure. The ineffective school which was paired with this school had no forced transfers of teachers out of the school.

Recent studies confirmed that there is a difference in variance of teacher behavior for effective versus ineffective schools. The variance in behavior is less for the effective schools, with all teachers grouped toward a higher mean of behavior. The ineffective schools yield a greater variety of both effective and ineffective teaching behaviors. One study sought to explain this difference in variance by examining beginning teacher socialization process (Kirby, in press), and found that beginning teachers were carefully and thoughtfully inducted into effective schools. Beginning teachers in effective schools perceived that they were provided with more help, and were more frequently monitored than beginning teachers in
ineffective schools. This is the only known research that has sought to explain this difference in variance of teacher behavior.

Measures of Effectiveness

Most educational research at the school level has used student academic achievement as the criterion for effectiveness. This is a limited measure of effectiveness, but this variable is of the greatest concern to the educational audience (i.e. parents, politicians, school board members). This variable is also the most measurable of the school outcomes (Good & Brophy, 1986), which contributes to its frequent and often exclusive use as the dependent variable.

Even though various techniques have been suggested for measuring academic achievement, the most commonly used methods utilize regression residual scores and determine effectiveness by performance above or below prediction based on students' background. Other methods use mean achievement scores, and may be a simple comparison of one school's standing with another, looking at trends of improvement, rates of gains, or disaggregation of data for different subgroups within a school (Levine & Lezotte, 1990). Many educational researchers express concern about the use of aggregate scores masking differences in group effects within the
schools (Cuban, 1983; Edmonds & Frederiksen, 1979; Frederick & Clauset 1985; Good & Brophy, 1986; Good & Weinstein, 1986). Although it is not a commonly used method, Levine and Lezotte (1990) advocate disaggregation of data by student SES and race/ethnicity.

Ron Edmonds was one of the first to bring up the question of equity in education. Edmonds and Frederiksen (1979) suggested that schools may not be equally effective for different groups within the school. They stratified students according to race and home background and compared mean achievement scores for the pupil subgroups. They found that schools that were effective for one group were not always effective for others. Economic factors appeared to account for the biggest differences in student achievement (more than racial differences). Therefore, the characteristics of effective teaching for low-SES students may be different from those of middle- or high-SES students.

Frederick and Clauset (1985) addressed the equity concern by using different algorithms to classify a school as effective or ineffective. The six different algorithms used produced inconsistent results for school effectiveness. They concluded that the use of aggregate scores often masks critical differences among various cohort groups within a school.
Very little attention has been given to the actual variance of overall scores within a school. Following the line of reasoning of those who believe in equity in education, one would conclude that the more effective schools would have less variance in achievement scores. The only study found that actually compares variance in achievement scores for effective versus ineffective schools reports just the opposite.

Lark et al. (1984) conducted a study of the Pennsylvania schools, in which they looked at the variance of schools which had been classified as effective or ineffective according to the regression method, and found that those schools labelled as effective actually had greater variance than those rated as ineffective. They also classified the schools as effective/ineffective according to the differences in mean achievement between the high- and low-SES students in each school. Schools with the smallest differences were labelled as effective to be consistent with the theoretical position that the effective schools would lessen the gap between SES groups. They found that this classification was totally inconsistent with the regression algorithm. In fact, they report that the larger differences between mean achievement of subgroups were found in the schools labelled as effective by the regression method. They contend that the larger differences in the effective schools is as it should be, since the
purpose of schools is to maximize each student's potential—which would widen the dispersion of student outcomes.

In a four year study of the London schools, Mortimore and Sammons (1987) reported that effective schools are effective for all groups, and less effective schools are less effective for all groups. The overall differences in SES were not completely removed, but they reported that "...on average a student from a blue-collar worker's family attending an effective school achieved more highly than one from a white-collar family background attending one of the least effective schools." (Mortimore & Sammons, 1987, p. 6)

A study by Crone and Tashakkori (1991) further examines this variance issue using the student component of the National Education Longitudinal Study (NELS) data file. NELS consists of a stratified random sampling of eighth grade students from over 1000 schools throughout the US. Data provided for each student includes scores on a four section standardized test, demographic information, plus student responses to questions on locus of control and self-concept. Using a standardized test composite score for reading and math, a mean achievement score and standard deviation for each school was computed. Mean SES indicators were also calculated for each school. A forward
regression indicated the best predictors of student achievement to be parent education, percent minority, and family income. With the use of this regression model, the schools were classified into six groups according to their residuals: (a) those residuals which fell below the 10th percentile were labelled ineffective outliers, (b) residuals between the 10th and 30th percentile were called ineffective, (c) residuals between the 30th and 50th percentile were referred to as ineffective typical, (d) residuals between the 50th and 70th percentile were labelled effective typical, (e) residuals between the 70th and 90th percentile were considered effective, and (f) residuals above the 90th percentile were called effective outliers.

An examination of the correlations between mean achievement scores and variance of achievement scores yielded interesting results. In the effective outliers and effective schools there was a strong negative correlation between these two variables: as the achievement scores increased the variance decreased. In the ineffective outliers and ineffective schools the reverse occurred: the strong positive correlation indicated that as the achievement scores decreased the variance also decreased. The ineffective typical and effective typical schools had no significant correlation between achievement scores and variance. This same pattern of
correlations occurred between variance of achievement scores and average SES.

A consideration of these correlations led to a separate examination of the variances of student achievement for upper and lower SES schools. It was found that the upper SES schools had the smallest variance in the effective schools. Conversely, the lower SES schools had the smallest variance in the ineffective schools. This was true not only in the extreme outliers, but also in the effective and ineffective groups. It was concluded that school SES and effectiveness have additive effects. When both are high, the students have homogeneously high performance; when both are low, the performance is homogeneous and low. Variance becomes larger when one of the two variables (SES or effectiveness) is high and the other is low. In low-SES effective schools the dispersion of scores is larger as some students are influenced by their poor background rather than the effective school. In ineffective high-SES schools, some students perform well as a function of family background while others are more affected by the poor quality of the school.

An analysis of the relationship between the variance in student achievement and the variance in SES resulted in no relationship between the two variables in the ineffective outlier schools, but as the schools
increased in effectiveness the relationship between the two variances increased also. These findings of Crone and Tashakkori indicate that the patterns of variance in student achievement cannot provide adequate information without considering the SES of the students involved.

It would seem logical to assume that the same pattern found in variance of teacher behavior might be found in variance of student achievement (that the more effective schools would have the smaller variance since all students would be brought up to a higher minimum level of achievement). This may not be a logical assumption though, when one considers that schools can transfer teachers, but are unable to transfer the weaker students. The findings of Crone and Tashakkori lead one to believe that (for eighth grade students) even the most effective schools are not able to bring up the scores of all of the lower SES students. If these findings generalize to the elementary school, one would not expect to find smaller variance in effective schools with low-SES student enrollment.

Whether or not the results found for eighth grade students will generalize to elementary students is another consideration. Elementary students, with the proper encouragement, may be much more willing to try to achieve than middle school students. Hence, the pattern in elementary
schools may be quite different. These results, though, do point out the necessity of considering SES when examining variance of achievement.

**Context Issues**

When researching the issue of equity (whether subgroups are being effectively reached within a school), context issues need to be considered. Research indicates that different strategies may be needed to effectively teach students from different SES levels. With desegregation and bussing, more schools are composed of a mixture of SES levels. Both school effectiveness and teacher effectiveness literature are revealing that what works for one SES level of students does not always work for another SES level. Hence, the challenge to teachers and schools to effectively teach subgroups within the schools becomes even greater.

Teddlie, Stringfield, Wimpleberg, and Kirby (1990) enumerated the contextual findings from Phase Two and Phase Three of the Louisiana School Effectiveness study (LSES). They found four characteristics of effective schools which applied to both low and middle-SES schools. Those included a clear academic mission and focus, orderly environment, high academic engaged time-on-task, and frequent monitoring of student progress.
The contextual differences found in LSES were in the areas of expectations, principal style, rewards, curricular offerings and teacher experience. The staff in effective middle-SES schools expressed both high present and future expectations for their students. The low-SES schools placed more importance on high present expectations.

The principal style in effective middle-SES schools was more often found to be that of the manager, with the teachers having more autonomy, where the low-SES schools seemed to be more effective with an initiator as principal. (One characteristic of the initiator is that he/she takes a more active role in the instructional leadership of the school.)

There were also more external rewards in the low-SES schools, while the middle-SES schools de-emphasized overt rewards. The curriculum was expanded beyond the basic skills for the effective middle-SES schools, while the effective lower SES focused on mastery of basic skills. The effective middle-SES schools had a high degree of contact with the community, where the lower SES schools often purposely kept the community out if they deemed the community effect to be negative. Teddlie, et al. (1990) also found more experienced teachers in the effective middle-SES schools and less experienced teachers in the effective low-SES
schools. Low-SES principals in effective schools also reported having more authority in selecting their staff.

In a study conducted in another part of the country Hallinger and Murphy (1986) reported very similar results when they compared low and high-SES effective schools. As found in LSES there were some traits that were common between high- and low-SES schools: these included strong goals, safe, orderly environment, clear mission, instructional leadership, high expectations, and well-coordinated curriculum. They also found quite similar results to LSES in regard to contextual differences. They reported that the effective low-SES schools focused more on mastery of basic skills.

The Hallinger and Murphy (1986) findings concerning principal styles for high-/low-SES schools were consistent with the LSES findings for middle-/low-SES schools: the principals in high-SES schools were found to exercise less direct control over the classroom instruction, giving the teachers more autonomy than was found in the low-SES schools. The system of rewards was found to be much more frequent and elaborate in low-SES schools than in high-SES schools. Hallinger and Murphy also found that the high expectations for low-SES students were focused on the present, where the high expectations for high-SES students were both present and future. The issue of community involvement was also
consistent with Teddlie et al. (1990). They found that the high-SES schools encouraged community involvement and the low-SES schools often discouraged that involvement. The only disparity in the contextual differences found in this study as compared to the LSES study was that Hallinger and Murphy did not mention any differences in teacher experience.

It seems reasonable to assume that some of these contextual differences would also apply at the teacher level. The teacher with low-SES students would need to provide more focus on basic skills, more overt rewards, and concentrate on present expectations. Brophy and Evertson (1976) reported differences in teaching style for high- and low-SES students. High-SES students needed to be stimulated with interesting, extra work; low-SES students simply needed assignments that they could handle and a teacher that would see that the assignments were done. High-SES students needed to move at a faster pace with larger steps than the low-SES students. Teachers of high-SES students get the best results when they do not allow students to call out answers to questions (not letting the most assertive dominate the discussions). Conversely, the calling out of answers correlated positively with student achievement for low-SES students. Low-SES students showed positive results from more
academic praise, where high-SES students’ achievement showed no correlation to praise.

Good, Ebmeier, and Beckerman (1978) found that low-SES students needed more monitoring and supervision. They also benefitted from a more relaxed climate and increased academic praise. Low-SES students showed more achievement gain with factual, product questions rather than open ended questions.

Rosenshine (1983) reported that younger, slower, less experienced students needed small steps, detailed and redundant instructions and explanations, many examples, many questions and opportunities for practice, a high success rate, seatwork broken down into small segments, and overlearning of material.

Considering all of these differential results, it is probable that different groups within a particular school may not be getting equal opportunities for education. Many studies have singled out schools which are basically at one SES level or the other (high or low, or middle or low). In reality many of our schools are a combination of SES levels. This brings us back to Edmonds’ concern as to whether different groups within a school are being effectively taught. If the school has a larger proportion of high- or middle-SES students and is classified as effective, are the low-SES students also getting the type of instruction that they need? Likewise, if
the majority of the students are low-SES, are the middle-SES students being properly challenged to produce their maximum effectiveness?
Chapter Three: Method

Overview

This chapter will begin with an overview of the research designs for the three issues to be examined in this study. After a discussion of the design, the sampling procedures will be described. A section delineating the different measures used to answer the research questions will then follow. The last section of this chapter will begin with a discussion of the methodological issues involved in making inferences regarding variances, and then describe the methods which will be employed to analyze the research questions.

Research Design

Issue I: The mean and variance of teacher behavior.

The purpose of Issue I was to add empirical evidence to the findings of Teddlie et al. (1989), and Virgilio et al. (1991) in regard to the differences found in the means and variances of teacher behavior in effective and ineffective schools. The research design for this issue is causal comparative, and compares two groups of outlier schools (effective versus ineffective). Each group consisted of six schools, with five teachers in each school. One effective school had to be dropped from the sample,
so the final sample included five effective schools and six ineffective schools. As discussed in the section on Sampling, this school was not found to be typical of an effective school.

There is some criticism of the outlier approach. Purkey and Smith (1983) contend that the groups demonstrating extremes in behavior should be compared to a group of average (or typical) subjects. In this study, though, it was decided that the outlier approach was the most advantageous. Two considerations went into making this decision. The first consideration was that this approach best distinguishes unique behavioral characteristics of different groups. As qualitative analysis was used to answer one of the questions in Issue I, these comparisons of extremes was desirable. The second consideration was practical. As Issue I required extensive time for observations and interviews, 12 schools were the maximum that could be included. It was deemed more appropriate to have two groups with 6 schools each, rather than three groups of four schools each. The amount of information to be gained from addition of the typical level was not deemed worth the additional time required to gather it.

As was done in both the Teddlie and Virgilio studies, teacher behaviors were rated on a Likert-type scale by observers. The
observations were done by three different observers, each filling out the Virgilio Teacher Behavior Inventory (Virgilio & Teddlie, 1989) and the Stallings Classroom Snapshot (Stallings & Kaskowitz, 1974) for every teacher in the sample. Only two of the three observers were used in the final analyses, as the scores of the third observer lacked the necessary variability for comparison of teacher behavior. An average of the two different observations were computed for each item on the two measures.

Means and variances of teacher behavior were computed for each school. Scores for each school included means and variance scores for the different domains of teacher behavior (classroom management, classroom climate1, classroom climate2, time-on-task, instruction1, and instruction2). These means and variances were averaged and compared for teachers in effective schools and teachers in ineffective schools.

The different variance scores of teacher behavior for each school were transformed to coefficients of variation. The coefficient of variation is a standardized measure of variability which is calculated by dividing the standard deviation by the mean and multiplying by 100 (SAS Institute Inc., 1985). Going back to the world of business, Kazmier (1978) explains the need for standardization of variances, even on the same scale of measurement. He asks which would have more variability—a stock issue
with a mean price of $150 with a standard deviation of $12, or a stock
issue with a mean price of $5 with a standard deviation of $1? Obviously,
without a consideration of the difference in means, the standard deviation
of $12 would appear to be greater. When changed to the coefficient of
variation, the standard deviation of $12 becomes $8, and the standard
deviation of $1 becomes $20. Consequently, as the means of teacher
behavior were found to be higher for the teachers in effective schools, it
was felt that the coefficient of variation would give a more accurate
comparison of variance than the actual variance or standard deviation.

Once coefficients of variation were determined for each school, a
comparison was made between the coefficients of variation for teacher
behavior in effective schools, versus the coefficients of variation for
teacher behavior in ineffective schools. The comparisons were of a
descriptive nature, as no statistical test of significance was found for the
coefficient of variation.

Analysis was conducted on the mean and variance of teacher behavior
for each separate domain. These scores were averaged and compared for
all teachers in effective schools versus all teachers in ineffective schools.

As differences were found, an examination of possible causes for this
disparity was conducted. Qualitative methods were employed to
investigate possible reasons for differences in means and variances. Individual teacher interviews were administered, using open ended questions. After one school was eliminated from the sample, 55 teacher interviews were used. These consisted of five teachers in each of five effective schools (25 teachers) and five teachers in each of six ineffective schools (30 teachers).

These interviews sought to establish whether differences in means and variances demonstrate any relationship to teacher selection, teacher socialization, principal’s instructional leadership, goal orientation, assignment of students to classroom, or any combination of these variables. The responses to the interviews were compared for teachers in effective schools to teachers in ineffective schools.

In addition to the teacher interviews, the principals of the 11 schools that remained in the sample for the final analyses were also interviewed. Altogether interviews of five principals of effective schools and six principals of ineffective schools were used for analysis. Questions of similar nature to the teacher interviews were asked of the principals. The principal interviews were compared between the two school effectiveness groups.
**Issue II: The mean and within class variance of student achievement.**

Issue II dealt with a comparison of the means and variances of student achievement for teachers with varying levels of teacher behavior. The teacher behavior scores from the observations done for Issue I were used to categorize the teachers as effective, typical, or ineffective. As the effectiveness of the school was not a factor in Issue II, the teachers from all 12 schools were used. This included 60 teachers.

An overall teacher behavior index was needed for teacher categorization. Since the two instruments used to evaluate the teachers were on two different scales of measurement, the scores were converted to \( Z \)-scores in order to create this index. The score for each domain of teacher behavior was converted to a \( Z \)-score using the domain mean and standard deviation of all teachers in the sample. Once the procedure was done for the seven domains, the \( Z \)-scores were averaged to get a composite mean for each teacher.

The teachers were then rank ordered according to their composite score, and divided into three equal groups. The highest third was labelled effective, the middle third typical, and lowest third ineffective. Each group consisted of 20 teachers.
As these teachers are either third or fifth grade teachers, the achievement scores used for Issue II were the grade three and grade five criterion-referenced test scores administered by the Louisiana Educational Assessment Program (LEAP). This LEAP test includes sections on Language Arts and Mathematics which were combined so that an average composite achievement score was used for each teacher (see section on Measures for a detailed explanation of the procedures used to combine these scores).

For each teacher there were two different means and two different variances calculated; one mean and variance for the low-SES students taught by each teacher and one mean and variance for the middle-SES students taught by each teacher. In order to dichotomize the students according to SES, free lunch status was used for disaggregation. Other indicators of SES, such as parent's education and occupation are also important but the data available was not as complete. In a forward stepwise regression which was done on the data, 45% of the variation in student achievement was explained by free lunch status. The data on parent's occupation or education added very little to the explained variance. Consequently, the students who received free/reduced lunch were considered low-SES. Those not receiving free/reduced lunch were...
considered middle-SES. As the free lunch status is determined by family income and family size, the SES status will hereafter be referred to as 'family income and family size.'

**Issue III: The mean and within school variance of student achievement.**

The purpose of Issue III was to compare the means and within school variances of middle- and low-SES levels of student achievement scores for effective, typical, and ineffective schools. The achievement scores selected for use in Issue III were a combination of the grade three and grade five criterion-referenced test scores administered by the Louisiana Educational Assessment Program (LEAP). For this issue, it was desired that schoolwide means and variances be provided for each SES level of students. Hence, the student level scores were transformed to Z-scores and both language arts and math in grades three and five were combined for each SES level. The two school level means and variances became the dependent variable.

The design is causal comparative, with the comparison of means and variances in student achievement being done for the main effect of the three school effectiveness groups (effective, typical, and ineffective), for the subgroups of students disaggregated by SES, and for the interaction effect. The same methods of disaggregation were used as were discussed
for Issue II, dividing the students into two different SES levels; middle and low. Again, family income and family size were deemed the most appropriate available variables to use for disaggregation.

**Sampling**

Twelve elementary schools were chosen for the sample which was used for both Issue I and Issue II. As Issue I entailed classroom observations, as well as qualitative interviews, the sample had to be limited to that which was feasible. One of the twelve schools was eliminated from Issue I as being atypical of an effective school. Although the regression clearly identified this school as effective, it was not known until visiting the school that it had a gifted and talented program which accounted for around 30% of the population. The school did not demonstrate any of the characteristics of an effective school. It was decided that the gifted and talented students, not the teacher behavior, must have raised the test scores.

For the purpose of this study it was necessary to choose schools which had at least some mixture of both low- and middle-SES (in order to have representative samples of both subgroups). Frequency counts of the percent of low-SES students in each school were run on all elementary schools within two school districts: one being the school district of a large southeastern city, and the other being a rural area near that city. The use
of both urban and rural schools was felt to make the study more
generalizable to the entire state and perhaps even the entire region. Only
those schools which had a 30%-70% ratio of low-SES students were used
to select the sample.

For Issue I this sample of schools was then stratified into two groups;
effective and ineffective schools. This classification was done by utilizing
a multiple regression analysis, predicting student achievement from indices
of socioeconomic status (SES). The student achievement scores were
transformed to Z-scores, using the state means and standard deviations for
each subject area and grade level. An overall school mean of all student
Z-scores was computed. The regression procedure was then done on the
school level data to establish the effectiveness status classification.

Different SES indicators (family income, family size, parent’s
education, parent’s employment) were used in a forward stepwise
regression to determine the best predictors of student achievement. The
forward stepwise regression identifies the smallest set of variables that are
needed to maximize the explained variance. The computer does this
process on a mechanical basis, taking the single best predictor first, then
adding the next best predictor, etc. This process is done until the change
in explained variance is no longer significant. Since there was no
preconceived theory as to the best predictors of student achievement for this data, the stepwise regression was considered appropriate. The forward stepwise was preferred over the backward stepwise regression "since it attempts to build on the explained variance as opposed to reducing the amount of explained variance". (Keppel & Zedeck, 1989, p. 408)

The best regression model was found to be with the use of family income, family size, and father’s education as the predictors. This yielded an $R^2$ of .47. The residual for each school was used to label the schools as effective or ineffective. The residual is an indicator of the effect of the school after the influence of socioeconomic status has been partialed out. Consequently, a positive residual indicates that the students’ achievement scores are higher than was predicted for that school, based on the SES of that school. A negative residual shows that the students have performed below what was predicted of them.

For this study, +/- .75 standard deviations were used as the cutoff point. Lang (1991) found that +/-1 standard deviations is not the most desirable in controlling for the possibility of chance agreement. He reported finding that, when using +/-1 standard deviations as the cutoff point for school effectiveness categories, there were large differences between the Kappa
coefficient and the agreement ratio. This suggested that substantial chance agreement was occurring. This chance agreement was due to the clustering of schools in the central cell. As the cut off points were brought in closer to the mean, the chance agreement was minimized. Lang suggests +/- .674 as being less subject to chance agreement. The distribution of the data for this study makes the use of +/-1 too limiting in the amount of schools available, so it was deemed more appropriate to use a decision point closer to that recommended by Lang.

Consequently, those schools which had residuals of over .75 standard deviations above the mean of the residuals were considered effective. Those schools with residuals below -.75 standard deviations below the mean of the residuals were considered ineffective. From each of these two strata, six schools were randomly selected.

Five teachers were randomly selected from each of the twelve schools. Only teachers from grades three and five were considered since those are the ones for which the CRT test data is available. As mentioned above, only 11 schools were in the final sample for Issue I, this included 25 teachers in effective schools, and 30 teachers in ineffective schools.

For Issue II all twelve schools that were in the original sample were used. This was deemed appropriate because the one school was eliminated
for Issue I because of questions as to its effectiveness status. Since Issue II dealt with teacher effectiveness, not school effectiveness, the teachers from this school were also used. All students being taught by these 60 teachers were included in the sample for Issue II.

Issue III used all regular, public elementary schools throughout the state. This was limited to those schools which contain both grades three and five. Approximately 750 schools were included. Using the school level CRT data for grades three and five a regression was run. By the time the analysis was conducted for Issue III, the data had undergone considerable clean up. This was done because the data was used for the state’s school incentive program (Crone, Franklin, Caldas, Ducote, & Killebrew, in press). The clean up involved eliminating schools that had little or no test data (alternative and special education schools), eliminating scores that were incorrectly coded as regular education when they should have been special education, and making corrections for incorrect or nonexistent school codes. Also, extensive analysis had been conducted on the data (Franklin, Caldas, Crone, Killebrew, & Ducote, in press) to find the best predictors of these achievement scores. Hence, the amount of variance in achievement scores explained by the regression model was much larger than the prediction model for Issue I. The final regression
model used for Issue III included family income, family size, percent
minority, the interaction between income and community type, and percent
attendance. This yielded an $R^2$ of .69. [In their examination of effective
schools models for the Milwaukee area, Witte and Walsh, 1990, also found
an interaction effect for SES and community type.]

Employing the residuals in the same manner as described above, the
schools were placed into three categories: (a) effective, (b) typical, and
(c) ineffective. The distribution of schools throughout the state was
analyzed before designating the decision points for classification. The
distribution was found to be normal, and the decision points were again
felt to be most appropriately placed at +/- .75

As in Issue I, the sample was then limited to those schools that had a
30-70% ratio of low-SES students. As one of the main concerns was the
within school variance and how it related to school effectiveness and SES
levels, it was important to have schools that had a mixture of middle- and
low-SES students. The final sample included 54 schools in each school
effectiveness category; 162 schools in all.
Measures

Issue I: Means and variances of teacher behavior.

The independent variable for Issue I was school effectiveness which was determined by the regression procedure described above. The SES measures used as the predictor variables in the regression are collected as part of the statewide testing program. The classroom teachers report the parent education and parent employment level data, and the students report the family income and family size.

The student achievement was measured using the criterion-referenced test results for grades three and five. These tests are administered as part of the Louisiana Educational Assessment Program (LEAP). This program gives CRT tests to grades three, five, seven, ten, and eleven in all public schools throughout the state. This is a test produced by the State Department of Education, which is developed to measure the attainment of the curricular guidelines specified by the state. They are not minimum skills tests, but are designed to measure grade level skills. The state’s curriculum guides are constructed with specific standards for each grade level and subject area. The items in the CRT are then designed and validated in order to reflect those standards. For grades three and five this
test includes two subject areas: language arts and math. (Louisiana Department of Education, 1989).

Reliability for the CRT test for grade three was reported as .94 for the language arts portion, and .93 for the mathematics test. For grade five, reliability coefficients were .93 for language arts and .93 for mathematics. These results were from the KR-20 test of reliability, using the actual 1989 LEAP administration (Louisiana Department of Education, 1989).

Content validity was also established for the test. Initially an item bank was developed which was designed to provide items which were matched on both content and item difficulty.

In order to get one aggregate test score for each school, the raw scores for both grade levels and subject areas were converted to Z-scores with the use of the state population means and standard deviations for each separate grade level and subject area. Purkey and Smith (1983) felt that the use of only one subject area and grade level gives a very limited view of a school's effectiveness. Witte and Walsh (1990) found that different variables were needed to predict the different subject areas of reading and mathematics. Consequently, multiple subject areas and grade levels were used. An overall school average was then computed. This composite
mean of the $Z$-scores for each school is what was used as the dependent variable in the regression to determine school effectiveness.

Teacher effectiveness, the dependent variable, was measured using the Virgilio Teacher Behavior Instrument (Appendix B). This instrument was developed by Virgilio to provide an easily coded, research oriented instrument to determine teacher effectiveness in the classroom. The reliability and validity of the VTBI were reported by Teddlie, Virgilio, and Oescher (1990). They reported high Cronbach’s alpha reliability coefficients: total inventory (.96), classroom management (.88), instruction (.96), and classroom climate (.85). A panel of judges determined content validity. A correlation of .64 between the VTBI and time-on-task was shown for concurrent validity.

Construct validity was established by a five factor solution consistent with the theoretical structure of the instrument. These included: (a) classroom management factor, (b) instructional strategies factor, (c) instructional presentation/questioning factor, (d) classroom social/psychological environment factor, and (e) classroom physical environment factor. Inter-rater reliability was computed for the observers on the instrument prior to the study. The Pearson correlation yielded an $r$ of .93.
In addition to the VTBI, time-on-task was measured by use of the Classroom Snapshot portion of the Stallings Observation System (Stallings & Kaskowitz, 1974). Inter-rater reliability was established for the three observers who were using the instrument for this study. The correlation was $r = .93$. The instrument itself has been used in studies of teacher effectiveness (Stallings & Kaskowitz, 1974; Stallings, 1980; Teddlie et al., 1989). Reliability studies were conducted by Stallings and Kaskowitz (1974). They reported inter-rater reliability of .70 on most of the variables. They also found that data seemed to be reliable across days. Stallings (1980) also concluded that the instrument was an effective predictor of student achievement.

The teacher interview form was designed specifically for this study. Based on school effectiveness literature, those indicators which had been found to make a difference (Geske & Teddlie, 1990; Virgilio et al., 1991) in the effectiveness of a school were considered for this interview form. Those indicators were chosen that would most obviously have an effect on teacher behavior, such as teacher selection, evaluation, etc. An instrument by Rosenholtz, which is designed to measure teacher socialization, was also used as a model in the development of interview questions (Rosenholtz, 1989). The teacher interview form includes open ended
questions on teacher selection and socialization, evaluation procedures, school goals and rules, and assignment of students to classrooms (Appendix C).

The principal interview form was also designed for the study, and again includes open ended questions (Appendix D). Basically the same questions were asked of the principal as were asked of the teachers. Many of the questions on the teacher interview were regarding the activities and conduct of the principal. It was deemed appropriate to look into the school's policies and procedures from the principal's viewpoint as well as the teacher's viewpoint.

Issue II: Means and within teacher variance of student achievement.

The independent variable for Issue II was teacher effectiveness. The ratings which each teacher received on the Virgilio Teacher Behavior Inventory and Stallings Classroom Snapshot were combined into one composite teacher behavior index. This composite teacher behavior score was used to classify the teachers as effective, typical, and ineffective.

The mean and variance of individual $Z$-scores for the two SES levels of students being taught by each teacher provided the dependent variable for Issue II. The means and within teacher variances were averaged and
compared for each overall category of teacher effectiveness (effective, typical, and ineffective).

**Issue III: Means and within school variance of student achievement.**

The independent variable for Issue III was school effectiveness. The mean and variance of the individual student composite CRT scores became the dependent variable for Issue III. These means and variances were averaged for each SES subgroup within each school. The SES subgroups were compared within each school category (effective, typical, and ineffective).

**Variance as an Inferential Statistic**

As discussed in Chapter One, there is some literature, both in industrial research and educational research, that makes inferences about variance. This is, though, a rather rare procedure in educational research. Games (1972), and Martin and Games (1977) suggest that one reason why there is so little research testing hypothesis about equality of variances is that the available tests for sample variances are extremely sensitive to the normality assumption. Martin and Games (1977) also point out that the existing tests for homogeneity of variance are restricted to single factor designs. Both studies recommend that the solution to these problems is a
modified Box test, which performs an ANOVA on logarithmic transformations of variance estimates. The two procedures recommended, though, use forms of the jackknife and are rather complicated, as even they admit.

Keppel (1973) discusses the need to examine the within-group variance in certain circumstances:

Usually, we have little sustained interest in the within-group variances ... This does not always have to be the case. Important changes in behavior caused by the different treatments may not be revealed in average performance. Or, if we look for them, systematic differences among the treatment conditions might be reflected in average performance and in the variability of the subjects within these conditions. (Keppel, 1973, p.80)

Keppel believes that the test for homogeneity of variance is not sensitive enough to differences in within-group variance to use when hypothesis testing. He suggests a procedure developed empirically by Levene (1960) with the use of several Monte Carlo experiments, which consists of an ANOVA conducted on absolute deviation scores.

Hays (1973, chap.11) also discusses the importance of testing hypothesis of variances. He, too, emphasizes the need to be cautious about the overall distribution when using the F-test on variances. As long as the distribution is normal or the sample size is large, he does recommend using the F-test for hypothesis testing on variances.
The field of kinesiology has a great deal of interest in the within subject variation of performance. They use this "error variance" as the dependent variable in studies which search for techniques that may improve the consistency of motor skills. This within subject standard deviation of performance is used as the data points in ANOVA procedures. There is no mention of transformations or special techniques for this measure of variability (Schmidt, Young, Swinnen, & Shapiro, 1989; Wulf & Schmidt, 1989).

This study plans to examine many variances simultaneously. In addition to descriptive analysis, the intention is to use the variances in a mixed factorial design for Issues II and III. Before proceeding with this plan, a careful examination of the distribution was conducted; the overall distribution of student scaled scores was found to be normal. It is recommended by Bryk and Raudenbush (1988) that once sample variances have been transformed to approximate a normal distribution, hypothesis testing techniques from normal distribution theory can then be applied. Consequently, it was deemed appropriate to do logarithmic transformations of the variances, and then use them as data points in a regular mixed factorial ANOVA. Issue II and Issue III utilize large enough sample sizes to justify proceeding cautiously with the planned analysis.
Data Analysis

Hypothesis 1 states that there is a difference in teacher behavior in effective/ineffective schools. The means of teacher behavior for each separate dimension were compared for teachers in effective schools versus teachers in ineffective schools. This analysis was done with the use of a multivariate ANOVA, and the univariate test for each dimension of teacher behavior. (See Table 3.1 for a summary of all research questions and analysis procedures.)

Research question 2 states that a difference does exist between the variance of teacher behavior in effective elementary schools versus ineffective elementary schools. A descriptive analysis was done on the coefficients of variation which were computed for each domain of teacher behavior.

Research question 3 asks what factors may contribute to any possible differences in means and variances of teacher behavior for effective versus ineffective schools. Qualitative analyses were done on the teacher and principal interviews. The responses to the interview questions were analyzed using the constant comparative method. Once the responses to each item were placed into categories, they were compared for teachers
and principals in effective and ineffective schools. The responses of teachers and principals were compared for each school effectiveness type.

Research question 4a is that there is a difference in student achievement for students taught by effective, typical, or ineffective teachers. The teacher is the unit of analysis. A comparison of mean achievement scores for the effective, typical, and ineffective teachers was done by examining the between subject effect in a 3 X 2 mixed factorial ANOVA.

Research question 4b is that there is a difference in student achievement for different SES subgroups of students. A comparison of mean achievement scores for the two subgroups of students was done by examining the within subject effect of a 3 X 2 mixed factorial ANOVA.

Research question 4c states that there is an interaction of teacher behavior and SES levels. This was answered by looking at the interaction of the 3 X 2 mixed factorial ANOVA.

Research question 5a states that there is a difference in within class variance of students being taught by effective, typical, or ineffective teachers. The within class variance of student achievement for each of the 60 teachers was computed. Along with descriptive analysis, the between subject effect of a 3 X 2 mixed factorial was used to answer this question.
An additional test of variance was conducted using the $E_{\text{max}}$ test for homogeneity of variance. This was done using an average of the variances for each teacher effectiveness group.

Question 5b looked at the within class variance of student achievement for different SES subgroups of students. Two different within class variances were computed for each teacher: one for low-SES students and one for middle-SES students. The within subject effect of the 3 X 2 mixed factorial was used to analyze these variances. The interaction was examined for question 5c. Descriptive analysis was also conducted.

Question 6a maintained that there is a difference in student achievement for effective, typical, or ineffective schools. The mean achievement scores for students in these three school categories was compared by looking at the between subject factor of a 3 X 2 mixed factorial ANOVA.

Question 6b states that there is a difference in mean achievement for different SES subgroups of students. The within subject factor of a 3 X 2 mixed factorial ANOVA was used to analyze the means for the two SES subgroups. The interaction effect was examined to answer 6c.

Question 7a states that there is a difference in within school variance of student achievement for students in effective, typical, or ineffective
schools. The within school variance was determined for each school and used as a data point in a 3 X 2 mixed factorial ANOVA. The main effect for school effectiveness groups was analyzed to answer question 7a. The $F_{\text{max}}$ test for homogeneity of variance was also used. The average variance for each school effectiveness group was used for comparison.

Research question 7b is that there is a difference in within school variances of student achievement for different SES levels of students. Two different within school variances were calculated, one for the low-SES students and one for the middle-SES students. The within subject factor of a 3 X 2 mixed factorial ANOVA was used to analyze these variances. Descriptive analyses were again done for 7a, 7b, and 7c.
Table 3.1

**Analysis of Research Questions**

<table>
<thead>
<tr>
<th>Issue</th>
<th>Research question 1 (mean teacher behavior)</th>
<th>Research question 2 (variance of teacher behavior)</th>
<th>Test of Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>teacher</td>
<td>teacher</td>
<td>MANOVA/UNIVARIATE</td>
</tr>
</tbody>
</table>

**Issue II**

<table>
<thead>
<tr>
<th>Research question 4a (mean student achievement, group effect)</th>
<th>Research question 4b (mean student achievement, SES effect)</th>
<th>Research question 4c (mean student achievement, interaction)</th>
<th>Research question 5a (variance of student achievement, group effect)</th>
<th>Research question 5b (variance of student achievement, SES effect)</th>
<th>Research question 5c (variance of student achievement, interaction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>teacher</td>
<td>teacher</td>
<td>teacher</td>
<td>teacher</td>
<td>teacher</td>
<td>teacher</td>
</tr>
<tr>
<td>3 X 2 Mixed Factorial*</td>
<td>3 X 2 Mixed Factorial*</td>
<td>3 X 2 Mixed Factorial*</td>
<td>3 X 2 Mixed Factorial*</td>
<td>3 X 2 Mixed Factorial*</td>
<td>3 X 2 Mixed Factorial*</td>
</tr>
</tbody>
</table>
### Table 3.1 (continued)

<table>
<thead>
<tr>
<th>Unit of Analysis</th>
<th>Test of Significance</th>
</tr>
</thead>
</table>

#### Issue III

<table>
<thead>
<tr>
<th>Research question 6a</th>
<th>school</th>
<th>3 X 2 Mixed Factorial*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(mean student achievement, group effect)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Research question 6b</th>
<th>school</th>
<th>3 X 2 Mixed Factorial*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(mean student achievement, SES effect)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Research question 6c</th>
<th>school</th>
<th>3 X 2 Mixed Factorial*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(mean student achievement, interaction)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Research question 7a</th>
<th>school</th>
<th>3 X 2 Mixed Factorial*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(variance of student achievement, group effect)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Research question 7b</th>
<th>school</th>
<th>3 X 2 Mixed Factorial*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(variance of student achievement, SES effect)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Research question 7c</th>
<th>school</th>
<th>3 X 2 Mixed Factorial*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(variance of student achievement, interaction)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Mixed Factorial consist of one between teacher factor and one within teacher factor.
Chapter Four: Results

Overview

This chapter will be divided into four main sections. The first three sections will be based on the statistical analysis and results for each of the three main issues addressed in this study. As the different research questions employ various techniques for analysis, the statistical procedures utilized in this study will be reviewed at the beginning of each section reporting the findings for the issues. Section four will discuss one additional finding that was not related to the initial hypotheses.

Hence, the first section will review the statistical techniques used to analyze the issue regarding the means and variances of teacher behavior for effective versus ineffective schools. This will be followed by the results of these analyses.

In addition, this section will include the qualitative analysis, which is also part of Issue I. The purpose of the qualitative analysis was to find reasons for the differences in means and variances of teacher behavior for teachers in effective versus ineffective schools.

Section two will discuss the statistical techniques and results for Issue II. This issue explores the means and variances of student achievement for different subgroups of students taught by effective, typical, and ineffective teachers.
Section three will review the statistical procedures and discuss the results for Issue III. Issue III deals with the means and variances of student achievement for different subgroups of students in effective, typical, and ineffective schools.

Issue IV will relate the findings of an additional analysis, which was conducted on the means and variances of student achievement for different subgroups of students in schools with high-, medium-, and low-SES levels. This analysis was done after it was determined that the average of the means of student achievement for the final sample used in this study was above 50, although the transformation was done to a mean of 50. This sample for the main study only included schools that had more than 30% low-SES students and less than 70% low-SES students. The reason for this sampling procedure was so that a representative sample of both low and middle-SES students would be available in each school.

The transformation, however, to $T$-scores was based on the entire sample of schools regardless of SES. This was deemed appropriate so that the scores for this sample would be representative of the entire population. After determining that the sample used for the study had a mean achievement score that was above average, I decided to conduct an
analysis of the means and variances of the schools which were left out of the final sample.

**Issue I**

**Limitations of the sample.**

As discussed in Chapter 3, teacher behaviors were observed by three different observers, each completing one observation of all 60 teachers. The three observers first spent a day in the same school, and conducted observations in the same classroom at the same time, to establish interrater reliability. The interrater reliability using the Pearson correlation coefficient was .93 for the VTBI and .93 for the Stallings. The actual observations were done on different days and at different times of day.

Further examination of the results of the three separate observers indicated, however, that one observer had very little variation in scores, thus limiting the overall variability of the scores. The scores of this one observer were eliminated. Hence, the results only reflect the observations of two different observers.

One of the twelve schools was also eliminated from the analysis. When selecting the schools, it was not known that one of the schools had around 30% of the student population classified as gifted and talented. This could obviously elevate the mean and residuals of student scores for
that school. Once inside the school, the observers saw very little evidence of the characteristics of an effective school; therefore, it was decided that this school would not be included in the analysis. Consequently, the final sample for Issue I includes five effective schools and six ineffective schools.

**Review of methods of analysis.**

The first research question for Issue I stated that there is a difference in teacher behavior in effective/ineffective schools. The means for seven different dimensions of teacher behavior were analyzed using the MANOVA procedure. The overall multivariate test was first examined, and the univariate tests were only considered if the MANOVA value was significant.

Research question 2 stated that a difference exists between the variance of teacher behavior in effective elementary schools versus ineffective elementary schools. The coefficient of variation was used to examine the dispersion of scores for the seven different dimensions of teacher behavior. They were analyzed solely with the use of descriptive statistics, as no procedures were found for testing the significance of this data.
The third part of Issue I concerned the qualitative analysis of teacher and principal interviews, which were conducted to determine the causes of any differences found in the means and variances of teacher behavior in effective versus ineffective schools. The responses to the open ended interview questions were analyzed using the constant comparative method (Lincoln & Guba, 1985). Each individual's response for each question was placed on an index card. If one respondent gave more than one answer, each answer was treated separately (placed on a different card). The cards were then sorted and compared within the effective and ineffective school groups.

Results of Quantitative Analysis.

The results concerning Issue I, question 1 indicated that there was a difference in five of the seven dimensions of teacher behavior for teachers in effective schools versus teachers in ineffective schools. Table 4.1 contains the means, standard deviations, and $F$-values for the univariate tests. The multivariate results indicated a significant difference between the two groups of teachers, $F(7,47) = 5.75$, $p < .0001$. In every case the mean was higher for the teachers in effective schools than for the teachers in ineffective schools.
The means for the first two variables (time-on-task) are expressed as percentages. The interactive time-on-task reflects the percent of time that the students are actively working with the teacher, rather than the students working on their own. The overall time-on-task indicates the percent of time the students were actually observed as being on task. Although the means were higher for the teachers in effective schools, there was no significant difference found for either dimension of time-on-task for the two different groups of teachers.

The remaining five dimensions of teacher behavior were scored on a 6-point Likert-type rating scale (5 being considered the most effective, and 6 indicating that the behavior was not observed). According to the factor analysis from the validation of the Virgilio Teacher Behavior Instrument, instructional items were divided into two dimensions. The first dimension (instruction1) deals with instructional strategies (evaluating student progress, demonstrating a variety of teaching methods, providing students with review and practice), while instruction2 pertains to presentation/questioning (accuracy of material, clarity of directions and explanations, questioning techniques, etc.). Climate1 deals with the social/psychological environment of the room, where climate2 is an
indicator of the actual physical attributes of the room (Teddlie, Virgilio, & Oescher, 1990).

Table 4.1

Comparison of Means by School Effectiveness Type on Seven Dimensions of Teaching Behavior

<table>
<thead>
<tr>
<th>Variable</th>
<th>Teachers in Effective Schools (n=25)</th>
<th>Teachers in Ineffective Schools (n=30)</th>
<th>F-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time-on-task (interactive)</td>
<td>.55 ± .18</td>
<td>.46 ± .16</td>
<td>3.47</td>
</tr>
<tr>
<td>Time-on-task (overall)</td>
<td>.73 ± .11</td>
<td>.70 ± .13</td>
<td>.99</td>
</tr>
<tr>
<td>Management</td>
<td>4.21 ± .61</td>
<td>3.35 ± .68</td>
<td>23.74**</td>
</tr>
<tr>
<td>Instruction1</td>
<td>3.70 ± .51</td>
<td>2.81 ± .73</td>
<td>26.85**</td>
</tr>
<tr>
<td>Instruction2</td>
<td>3.92 ± .56</td>
<td>2.94 ± .70</td>
<td>32.65**</td>
</tr>
<tr>
<td>Climate1</td>
<td>4.23 ± .59</td>
<td>3.46 ± .65</td>
<td>20.83**</td>
</tr>
<tr>
<td>Climate2</td>
<td>3.72 ± .76</td>
<td>3.06 ± .73</td>
<td>10.60*</td>
</tr>
</tbody>
</table>

Note. Time-on-task scores range from .00 (0%) to 1.00 (100%). The other skills range from 1 (low) to 5 (high).

*p<.002 **p<.0001

Each of these five dimensions had significantly higher means for the teachers in the effective schools. Except for interactive time-on-task, and
climate2, the standard deviations for the dimensions were smaller in the effective schools.

As seen in Table 4.2, the standardized measures of variation (coefficients of variation) were smaller for the effective schools on all seven dimensions of teacher behavior. The largest differences in variation of scores was evident between the two school effectiveness categories on the dimensions of overall time-on-task and instruction1 and instruction2. These differences (especially in the two dimensions of instruction) were caused by the much lower scores found in the ineffective schools. This indicates that the effective schools are somehow bringing up the lower end of the range of teacher behavior. There are some "good" teachers found in ineffective schools, but few "bad" teachers found in effective schools. These results are direct replications of the findings of Teddle, Kirby, and Stringfield (1989), and Virgilio, Teddlie, and Oescher (1991).

Results of teacher interviews.

The results of the qualitative analysis utilized to answer question #3 indicated that there were characteristics found in effective schools that were not found in ineffective schools which could possibly explain the quantitative differences in the means and variances of teacher behavior.
Table 4.2

Comparison of Variance in Scores on Teacher Behavior for Effective Versus Ineffective Schools

<table>
<thead>
<tr>
<th>Variable</th>
<th>Lowest Score</th>
<th>Highest Score</th>
<th>Range</th>
<th>Coefficient of Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEACHERS IN EFFECTIVE SCHOOLS (n=25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time-on-task (interactive)</td>
<td>.15</td>
<td>.85</td>
<td>.71</td>
<td>31.22</td>
</tr>
<tr>
<td>Time-on-task (overall)</td>
<td>.55</td>
<td>.90</td>
<td>.36</td>
<td>10.39</td>
</tr>
<tr>
<td>Management</td>
<td>2.37</td>
<td>5.00</td>
<td>2.64</td>
<td>11.78</td>
</tr>
<tr>
<td>Instruction1</td>
<td>2.74</td>
<td>4.61</td>
<td>1.88</td>
<td>12.28</td>
</tr>
<tr>
<td>Instruction2</td>
<td>2.75</td>
<td>4.88</td>
<td>2.14</td>
<td>12.42</td>
</tr>
<tr>
<td>Climate1</td>
<td>2.60</td>
<td>4.90</td>
<td>2.31</td>
<td>12.71</td>
</tr>
<tr>
<td>Climate2</td>
<td>2.33</td>
<td>5.00</td>
<td>2.68</td>
<td>17.88</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Lowest Score</th>
<th>Highest Score</th>
<th>Range</th>
<th>Coefficient of Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEACHERS IN INEFFECTIVE SCHOOLS (n=30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time-on-task (interactive)</td>
<td>.08</td>
<td>.75</td>
<td>.68</td>
<td>34.45</td>
</tr>
<tr>
<td>Time-on-task (overall)</td>
<td>.48</td>
<td>.96</td>
<td>.49</td>
<td>19.01</td>
</tr>
<tr>
<td>Management</td>
<td>2.10</td>
<td>4.60</td>
<td>2.51</td>
<td>18.86</td>
</tr>
<tr>
<td>Instruction1</td>
<td>1.28</td>
<td>4.30</td>
<td>3.03</td>
<td>23.74</td>
</tr>
<tr>
<td>Instruction2</td>
<td>1.38</td>
<td>4.38</td>
<td>3.01</td>
<td>22.04</td>
</tr>
<tr>
<td>Climate1</td>
<td>2.30</td>
<td>4.80</td>
<td>2.51</td>
<td>16.97</td>
</tr>
<tr>
<td>Climate2</td>
<td>1.50</td>
<td>4.67</td>
<td>3.18</td>
<td>22.07</td>
</tr>
</tbody>
</table>

Note. The coefficient of variation is computed by dividing the standard deviation by the mean and multiplying by 100.
The analysis of the teacher interviews (Appendix C) showed that there was more cohesiveness among the teachers in the effective schools. The responses to three different questions revealed that teachers in effective schools both work and socialize more together.

Of the teachers who came to the school after or at the same time as the tenure of the current principal began, we asked how they found out about the school policies and procedures when they first arrived. This included 13 teachers in effective schools and fourteen teachers in ineffective schools. Eight of the teachers in effective schools mentioned that they found out from other teachers, whereas only three of the teachers in ineffective schools mentioned other teachers as a source of information. In the ineffective schools the most common response to this question was that they found out from the policy handbook (11 responses), compared to only five teachers in effective schools who mentioned the policy handbook.

When asked how they were made to fit in, the teachers who came at the same time or after the present principal indicated a better socialization process in the effective schools. Eleven teachers in effective schools named other teachers as helping. Only seven teachers in ineffective schools said that the other teachers helped them to fit in.

All teachers were asked what kind of input they have regarding school goals. Twelve of the teachers in effective schools responded that the
teachers work together. Only two of the teachers in ineffective schools answered that teachers worked together on the goals.

Another question that yielded noticeably different responses for the differentially effective schools concerned changes made when a new principal took over. This item was only answered by teachers who were in the school before the present principal (12 teachers in effective schools and 16 teachers in ineffective schools). The question asked of the teachers was "When the new principal came, how did you find out the policies and procedures of the new principal?"

Over half of the responses pertained to the actual changes, or lack of changes made by the principal. Seven responses in ineffective schools were that the policies and procedures basically remained the same, whereas only two teachers in effective schools stated this. In the effective school group there were 15 concrete examples cited of changes made, whereas the ineffective group mentioned 10 differences. These 10 differences referred to by the ineffective group were more nebulous than the changes made in effective schools. Examples of the changes mentioned in effective schools are: four mentions of instating discipline policies, plus changes in academics, homework policies, rebuilding of the library, improvement of grounds, etc. Examples of ineffective school changes made by principals...
are 'less pressure', 'more organized', 'more committees', and 'he meets with grade level teachers'.

When all teachers in the schools were asked how they know what their principal thinks of their teaching capabilities, the answer for teachers in both effective and ineffective schools most commonly alluded to some sort of formal observation or evaluation. There was equal mention of this for both groups. A difference did occur, though, in reports of the principal being in and out of the classroom on an informal basis. Eleven teachers in effective schools commented on this, where only six teachers in ineffective schools made any mention of the principal conducting informal classroom observations.

A major difference was also seen in the responses to the question asked of all teachers concerning the establishment of conduct rules for their classes. In effective schools, thirteen teachers mentioned schoolwide rules, nine mentioned assertive discipline, and only four said they decided on the rules by discussing them with the class. In the ineffective schools, schoolwide rules were referred to six times, assertive discipline was mentioned twice, and eleven teachers established their classroom rules by discussing them with their students.
The only other outstanding difference concerned how new teachers were selected for the school. In this sample, all teachers stated that the principal had a major role in the hiring. The main difference was that in the ineffective schools four teachers were hired after being student teachers there at that particular school. There was no mention of new teachers being selected from student teachers in the effective schools. This confirms recent results from TerHaar (1992).

Results of principal interviews.

The responses of the principal interviews (Appendix D) yielded similar results to those of the teacher interviews. These interviews were conducted with five principals of effective schools and six principals of ineffective schools. All questions reflect responses from all eleven principals. When asked what kind of changes they made when first taking over as principal, 20 changes were mentioned by principals in effective schools compared to 11 changes in ineffective schools.

In response to the question concerning the input of teachers regarding school goals and policies, four of the principals in effective schools stated that the teachers were on the committees in which the goals and policies were developed, only one principal in the ineffective school group mentioned that a teacher committee developed the goals and policies.
Other principals in ineffective schools alluded to discussing the goals and policies with the teachers in meetings.

When asked how the conduct rules were established for their school, five of the principals in effective schools talked of working with the teachers to develop schoolwide policies. An example of a response from a principal of an effective school demonstrates the cohesiveness of the school:

We use Assertive Discipline. We sat down as a faculty and looked at a variety of different discipline programs. We chose Assertive Discipline. I believe in getting faculty input. That way if there is a problem we are all in on it. I took the training and then came back and conducted my own inservices. Then the teachers came up with their classroom programs. All the teachers turned in their rules to me. We did it schoolwide.

On the other hand, the ineffective principals mostly talked of following parish guidelines, or gave very noncommittal answers such as the following.

We follow the parish discipline policies. I tried several things. The policy and direction has to come from the principal. I tried Assertive Discipline. It didn’t work though, the little kids don’t understand. I expect them to behave.

Some differences emerged in the principal interviews that were not seen in the teacher interviews. These had to do with the way the principals reported dealing with teachers who did not subscribe to their
philosophies, what they looked for in hiring teachers, future goals for the school, and the most important role of the principal.

When discussing action taken with teachers who did not subscribe to their philosophy of education, three effective school principals talked of working with the teachers; where only one ineffective school principal discussed "personally working out a few small problems." The most common response from principals of ineffective schools was that there were no problems, or that they just got rid of the teachers. Quotes from effective/ineffective groups, respectively, demonstrate the differences:

I observed 90% of the time the first year. When I saw problems I would suggest changes and ways to improve. I sent some to workshops. After a time, if there was no improvement, then I informed them that it would go on their evaluation. I then began remediation. A few did retire/resign/get transferred.

I won’t hesitate to remove ineffective teachers. I’ve had some that just didn’t fit into this setting.

When asked how they select new teachers, all felt that they had input. The differences were in what they looked for. The principals of effective schools seemed to look more for personality traits of their prospective teachers-creativity, flexibility, concern and caring for children. Two ineffective school principals did mention looking for teachers that like children, but the most common response of these principals was that they asked about the philosophy of education or discipline. No effective school
principal mentioned asking about a teacher's philosophy or discipline policy.

In discussing future goals for the school, four effective school principals talked of goals that pertained directly to student learning (i.e., bringing up science and social studies test scores, more time for learning). Only one ineffective school principal referred to learning-related goals. Other responses from principals in ineffective schools concerned more indirect methods of influencing student learning, such as computer labs, parent education, or full time librarian.

In regard to what they consider the most important role of a principal, most principals in both groups expressed the belief that instructional leadership was the most important. A noticeable difference was that three of the principals in ineffective schools went on to add that, due to administrative work, they do not have time to fulfill this role.

**Issue II**

**Review of methods of analysis.**

The first research question (question 4) for Issue II looked at means of student achievement for middle- and low-SES subgroups of students taught by effective, typical, and ineffective teachers. The main effects for both variables (teacher effectiveness, and SES groups) were examined, as well
as the interaction, and post hoc analysis for the differential teacher
effectiveness groups.

The analysis was conducted using a 3 X 2 mixed factorial design. The
three different teacher effectiveness groups were the between subject
factor. As the teacher was the unit of analysis, the two different means of
student achievement for the different SES levels of students taught by each
teacher was the within subject factor. The error term for the within
teacher factor and interaction consisted of the interaction of SES levels and
subjects pooled over the teacher effectiveness levels (Keppel & Zedeck,

Question 5 concerned differences in the within teacher variance of
student achievement for different SES levels of students taught by
effective, typical, and ineffective teachers. Two methods of analyses were
utilized on the variances. One method averaged the variance for each
teacher in the different groups and compared using the $E_{\text{max}}$ Test for
homogeneity of variance. The other procedure used the variance for each
teacher as the data point and tested for significance with the use of a
3 X 2 mixed factorial, as discussed above. As tests of significance of
variance differences are very sensitive to distribution form, logarithmic
transformations were done on the variances in order to stabilize the
distribution (Games, 1977).
Results of analysis on mean student achievement.

Results for the main effect of teacher effectiveness groups indicated that there is a difference in the mean of student achievement for teachers with varying degrees of teacher effectiveness. Table 4.3 shows the average means for the three groups of teachers, with Table 4.4 giving the results of the ANOVA. The means are in the expected direction, with the effective teacher group having the highest scores on student achievement.

The overall standard deviation was the smallest in the effective teacher group, followed by the typical teacher group, with the largest standard deviation being in the ineffective teacher group.

Table 4.3

Comparison of Mean Student Achievement for Effective, Typical, and Ineffective Teachers

<table>
<thead>
<tr>
<th>TEACHER EFFECTIVENESS CATEGORY</th>
<th>M</th>
<th>SD</th>
<th>LOWEST SCORE</th>
<th>HIGHEST SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective Teachers (20 Teachers)</td>
<td>52.22</td>
<td>3.06</td>
<td>44.51</td>
<td>58.24</td>
</tr>
<tr>
<td>Typical Teachers (20 Teachers)</td>
<td>49.76</td>
<td>3.14</td>
<td>42.24</td>
<td>55.39</td>
</tr>
<tr>
<td>Ineffective Teachers (20 Teachers)</td>
<td>47.47</td>
<td>3.83</td>
<td>39.38</td>
<td>53.52</td>
</tr>
</tbody>
</table>
Table 4.5 shows the differences in the average of the means of student achievement for the two different SES subgroups of students. There was almost as much difference between the means of student achievement for the two SES levels (4.26) as there was between the means for the effective teachers versus the ineffective teachers (4.75). Upon examination of the highest and lowest scores, the lowest mean scores demonstrated very little difference. The difference between the means of the two SES groups appeared to be caused by the upward extension of the range in the middle-SES group. There was no interaction effect between teacher effectiveness and SES levels, indicating that effective teachers were effective for both SES groups and ineffective teachers were ineffective for both SES groups.

Table 4.4

Results of Analysis on Mean Student Achievement for Teacher Effectiveness Groups and SES

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F-Value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>2</td>
<td>462.02</td>
<td>231.01</td>
<td>11.72</td>
<td>.001</td>
</tr>
<tr>
<td>Error</td>
<td>57</td>
<td>1123.12</td>
<td>19.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES</td>
<td>1</td>
<td>544.45</td>
<td>544.45</td>
<td>67.79</td>
<td>.0001</td>
</tr>
<tr>
<td>Group * SES</td>
<td>2</td>
<td>2.43</td>
<td>1.21</td>
<td>.15</td>
<td>ns†</td>
</tr>
<tr>
<td>Error</td>
<td>57</td>
<td>457.76</td>
<td>8.03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† not significant.
Table 4.5

Comparison of Mean Student Achievement for Middle- and Low-SES Groups

<table>
<thead>
<tr>
<th>SES LEVELS</th>
<th>M</th>
<th>SD</th>
<th>LOWEST SCORE</th>
<th>HIGHEST SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle-SES</td>
<td>51.68</td>
<td>4.37</td>
<td>38.39</td>
<td>60.34</td>
</tr>
<tr>
<td>(60 teachers)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-SES</td>
<td>47.42</td>
<td>3.95</td>
<td>38.23</td>
<td>58.51</td>
</tr>
<tr>
<td>(60 teachers)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The means separated by group and SES are shown in Table 4.6. Again they were all in the expected direction, with the highest means found in the middle-SES group taught by effective teachers. The low-SES group scored right around four points lower (slightly more than one standard deviation) than the middle-SES group when taught by effective and typical teachers. There was a larger gap between the SES subgroups when taught by ineffective teachers (4.66 points). The standard deviation of scores for both subgroups was also somewhat larger for the ineffective teachers.
Table 4.6

Comparison of Mean Student Achievement Across Teacher Effectiveness Categories for Middle- and Low-SES

<table>
<thead>
<tr>
<th>TEACHER EFFECTIVENESS CATEGORY</th>
<th>SES LEVEL</th>
<th>M</th>
<th>SD</th>
<th>LOWEST SCORE</th>
<th>HIGHEST SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective Teachers (20 teachers)</td>
<td>Middle-SES</td>
<td>54.07</td>
<td>3.98</td>
<td>44.82</td>
<td>60.33</td>
</tr>
<tr>
<td></td>
<td>Low-SES</td>
<td>49.96</td>
<td>3.23</td>
<td>43.89</td>
<td>58.51</td>
</tr>
<tr>
<td>Typical Teachers (20 teachers)</td>
<td>Middle-SES</td>
<td>51.42</td>
<td>3.61</td>
<td>42.13</td>
<td>57.07</td>
</tr>
<tr>
<td></td>
<td>Low-SES</td>
<td>47.40</td>
<td>3.25</td>
<td>38.32</td>
<td>51.37</td>
</tr>
<tr>
<td>Ineffective Teachers (20 teachers)</td>
<td>Middle-SES</td>
<td>49.54</td>
<td>4.41</td>
<td>38.39</td>
<td>55.02</td>
</tr>
<tr>
<td></td>
<td>Low-SES</td>
<td>44.89</td>
<td>3.72</td>
<td>38.23</td>
<td>55.52</td>
</tr>
</tbody>
</table>

The separate F-test for the effect of the teacher effectiveness groups on the different SES groups indicated that differences do occur for both of these groups (Table 4.7a and 4.7b). Consequently, the post hoc analysis was appropriate. The results of the Student-Newman-Keuls Test for the low-SES group showed that the means were significantly different for all three teacher effectiveness categories (Table 4.8a). When looking at the means for the middle-SES group, one can see that there was a difference between the means for the effective teachers compared to the typical and
ineffective teachers, but no difference occurred between the typical teachers compared to the ineffective teachers (Table 4.8b). Also, it was interesting that low-SES students with effective teachers perform at the same level as middle-SES students with ineffective teachers. This partially replicates results from LSES-II (Teddlie et al., 1984) which looked at schools rather than teachers.

Table 4.7a

**Results of Mean Student Achievement Across Teacher Effectiveness Categories for Low-SES Students**

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F-Ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>257.58</td>
<td>128.79</td>
<td>11.09</td>
<td>.0001</td>
</tr>
<tr>
<td>Error</td>
<td>57</td>
<td>661.69</td>
<td>11.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>59</td>
<td>919.26</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.7b

**Results of Mean Student Achievement Across Teacher Effectiveness Categories for Middle-SES Students**

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F-Value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>206.87</td>
<td>103.43</td>
<td>6.41</td>
<td>.003</td>
</tr>
<tr>
<td>Error</td>
<td>57</td>
<td>919.20</td>
<td>16.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>59</td>
<td>1126.07</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4.8a

Results of Student-Newman-Keuls Test for Mean Student Achievement of Low-SES Students by Teacher Effectiveness Group

<table>
<thead>
<tr>
<th>TEACHER EFFECTIVENESS GROUP</th>
<th>N</th>
<th>M</th>
<th>GROUPING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective Teachers</td>
<td>20</td>
<td>49.96</td>
<td>A</td>
</tr>
<tr>
<td>Typical Teachers</td>
<td>20</td>
<td>47.41</td>
<td>B</td>
</tr>
<tr>
<td>Ineffective Teachers</td>
<td>20</td>
<td>44.89</td>
<td>C</td>
</tr>
</tbody>
</table>

*Note.* Means with the same letter are not significantly different.

Table 4.8b

Results of Student-Newman-Keuls Test for Mean Student Achievement of Middle-SES Students by Teacher Effectiveness Group

<table>
<thead>
<tr>
<th>TEACHER EFFECTIVENESS GROUP</th>
<th>N</th>
<th>M</th>
<th>GROUPING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective Teachers</td>
<td>20</td>
<td>54.07</td>
<td>A</td>
</tr>
<tr>
<td>Typical Teachers</td>
<td>20</td>
<td>51.42</td>
<td>B</td>
</tr>
<tr>
<td>Ineffective Teachers</td>
<td>20</td>
<td>49.54</td>
<td>B</td>
</tr>
</tbody>
</table>

*Note.* Means with the same letter are not significantly different.

Results of analysis on within teacher variance.

Research question 5 stated that there is a difference in within teacher variance of student achievement for different SES subgroups of students.
taught by effective, typical, and ineffective teachers. Although no significant differences were found with either the ANOVA or $F_{\text{max}}$, some interesting trends were evident. Table 4.9 shows the average of within teacher variance of student achievement for the three different teacher effectiveness groups. There was very little difference between the variance of student achievement for the typical and ineffective teachers. The effective teachers had a smaller average variance.

Table 4.9

Comparison of Average Within Teacher Variance of Student Achievement for Effective, Typical, and Ineffective Teachers

<table>
<thead>
<tr>
<th>TEACHER EFFECTIVENESS CATEGORY</th>
<th>MINIMUM VARIANCE</th>
<th>MAXIMUM VARIANCE</th>
<th>AVERAGE VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective Teachers (20 Teachers)</td>
<td>26.50</td>
<td>102.49</td>
<td>61.35</td>
</tr>
<tr>
<td>Typical Teachers (20 Teachers)</td>
<td>44.13</td>
<td>118.06</td>
<td>69.93</td>
</tr>
<tr>
<td>Ineffective Teachers (20 Teachers)</td>
<td>12.09</td>
<td>126.12</td>
<td>70.96</td>
</tr>
</tbody>
</table>

Note. Average variance is calculated by dividing the sum of the within teacher variance for all teachers in that group by the n for that group.
When examining the overall average within teacher variance for the two SES subgroups, very little difference was evident. With the extremely wide range of variances, a difference of 3.22 points between the two averages is negligible (Table 4.10).

Table 4.11 shows the average within teacher variation of student achievement broken down by both teacher effectiveness and SES level. For both SES levels, the smallest variation did occur in the effective teacher group. There was approximately nine and a half points difference in variance between effective and ineffective teachers for both SES subgroups. The effective and typical groups showed similar differences, with the variances being smaller in the effective teacher groups.

Table 4.10

Comparison of Average Within Teacher Variance of Student Achievement for Middle- and Low-SES Students

<table>
<thead>
<tr>
<th>SES LEVEL</th>
<th>MINIMUM VARIANCE</th>
<th>MAXIMUM VARIANCE</th>
<th>AVERAGE VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle-SES</td>
<td>9.42</td>
<td>159.90</td>
<td>65.80</td>
</tr>
<tr>
<td>(60 Teachers)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-SES</td>
<td>11.21</td>
<td>182.77</td>
<td>69.02</td>
</tr>
<tr>
<td>(60 Teachers)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Average Variance is calculated by dividing the sum of the within teacher variance for all teachers in that group by the n for that group.
### Table 4.11

**Comparison of Average Within Teacher Variance of Student Achievement by Middle- and Low-SES Subgroups for Effective, Typical, and Ineffective Teachers**

<table>
<thead>
<tr>
<th>TEACHER EFFECTIVENESS CATEGORY</th>
<th>SES LEVEL</th>
<th>AVERAGE VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective Teachers (20 Teachers)</td>
<td>Middle-SES</td>
<td>58.96</td>
</tr>
<tr>
<td></td>
<td>Low-SES</td>
<td>63.74</td>
</tr>
<tr>
<td>Typical Teachers (20 Teachers)</td>
<td>Middle-SES</td>
<td>69.68</td>
</tr>
<tr>
<td></td>
<td>Low-SES</td>
<td>70.17</td>
</tr>
<tr>
<td>Ineffective Teachers (20 Teachers)</td>
<td>Middle-SES</td>
<td>68.75</td>
</tr>
<tr>
<td></td>
<td>Low-SES</td>
<td>73.16</td>
</tr>
</tbody>
</table>

**Note.** Average Variance is calculated by dividing the sum of the within teacher variance for all teachers in that group by the n for that group.

### Issue III

**Methods of analysis for Issue III.**

The analysis used for Issue III was the same as that used for Issue II. Issue III was also concerned with the means and variances of student achievement. In this case, though, the independent variables were school effectiveness groups and SES levels. The entire population of schools in the state which contained grades three and five (see Chapter Three) were categorized into three levels of effectiveness (effective, typical, and
ineffective) based on the residuals from regression analysis. Schools with residuals .75 standard deviations above the mean of the residuals were classified as effective, those .75 standard deviations below the mean of the residuals were classified as ineffective. From those two groups, only schools with more than 30% of the student population comprised of low-SES, and less than 70% low-SES, were chosen. This left 60 schools in the ineffective group and 54 in the effective group. In order to create an equal number of schools in each group, six schools were randomly eliminated from the ineffective group. It was felt that 54 schools were sufficient, and a balanced design was preferred. Fifty-four schools with a 30%-70% low-SES ratio were then randomly selected out of the remaining 'typical' schools.

The means of student achievement for the two SES subgroups of students in effective, typical, and ineffective schools were tested by using the 3 X 2 mixed factorial ANOVA as discussed in Issue II. As the school was the unit of analysis, the SES with two levels (middle and low) are non-independent (within subject factor). The three school effectiveness categories are independent (between subject factor). This analysis provides the answers for research question 6 regarding differences in student achievement for low- and middle-SES students in effective, typical, and
ineffective schools. Both the two main effects and their interaction were examined to answer this question.

Question 7, concerning within school variance of student achievement for low- and middle-SES levels of students in effective, typical, and ineffective schools, is also addressed using a 3 X 2 mixed factorial ANOVA. As in Issue II the variance in student achievement for each school was used as the dependent variable. Consistent with previous analyses, logarithmic transformations were performed on the variances before using them as data points in the ANOVA. An average of these variances for each group was also obtained and compared with the use of the $F_{\text{max}}$ Test of homogeneity of variance.

Results of analysis on means of student achievement.

The results of the analysis for question 6 indicated that there was a difference in student achievement for low- and middle-SES students in effective, typical, and ineffective schools. Question 6a dealt with the main effect of school effectiveness. The results indicated that there was a significant difference in the means of student achievement for effective, typical, and ineffective schools. The means were in the expected direction, with the effective schools having the highest scores. Table 4.12 shows the
means for the three school effectiveness categories, with Table 4.13 presenting the results of the ANOVA.

Unlike the trend seen in Issue II, the overall standard deviation of scores was larger in the effective schools. The range in scores was also greater for the effective schools, with the highest score widening the gap. There was about 6 points difference in the lowest to highest score in the ineffective schools, and almost 9 points difference in the effective schools.

Table 4.12

Comparison of Mean Student Achievement for Effective, typical, and Ineffective Schools

<table>
<thead>
<tr>
<th>SCHOOL EFFECTIVENESS CATEGORY</th>
<th>M</th>
<th>SD</th>
<th>LOWEST SCORE</th>
<th>HIGHEST SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective Schools (54 schools)</td>
<td>54.49</td>
<td>2.00</td>
<td>49.88</td>
<td>58.68</td>
</tr>
<tr>
<td>Typical Schools (54 schools)</td>
<td>51.86</td>
<td>1.47</td>
<td>47.72</td>
<td>54.88</td>
</tr>
<tr>
<td>Ineffective Schools (54 schools)</td>
<td>48.63</td>
<td>1.62</td>
<td>45.59</td>
<td>51.46</td>
</tr>
</tbody>
</table>
Table 4.13

Results of Analysis on Mean Student Achievement for School Effectiveness Groups and SES

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F-Value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>2</td>
<td>1924.64</td>
<td>962.32</td>
<td>233.78</td>
<td>.0001</td>
</tr>
<tr>
<td>Error</td>
<td>159</td>
<td>654.49</td>
<td>4.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES</td>
<td>1</td>
<td>1553.45</td>
<td>1553.45</td>
<td>537.81</td>
<td>.0001</td>
</tr>
<tr>
<td>Group * SES</td>
<td>2</td>
<td>15.43</td>
<td>7.72</td>
<td>2.37</td>
<td>ns†</td>
</tr>
<tr>
<td>Error</td>
<td>159</td>
<td>459.27</td>
<td>2.89</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† not significant

An examination of the means (Table 4.14) shows the differences between the two SES subgroups of students. As seen in Table 4.13, these differences are significant. Although the mean of student achievement for this sample was, on the average, about two points higher than for the sample used in Issue II, the gap between the two SES levels remained the same. There was a 4.26 point difference between middle- and low-SES groups in Issue II, and 4.38 points difference between the two groups in Issue III. There was no significant interaction effect between school effectiveness categories and SES levels. There was, though, some evidence of differential treatment for the different subgroups within the
Where the middle-SES student achievement scores increase in nearly equal increments for each school effectiveness category, the low SES student achievement is not equally increased by the effective schools. In other words, the effective schools appear to be more effective for the middle-SES students than for the low-SES students.

Table 4.14

Comparison of Mean Student Achievement for Middle- and Low-SES Groups

<table>
<thead>
<tr>
<th>SES LEVEL</th>
<th>M</th>
<th>SD</th>
<th>LOWEST SCORE</th>
<th>HIGHEST SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle-SES</td>
<td>53.84</td>
<td>3.23</td>
<td>46.33</td>
<td>62.31</td>
</tr>
<tr>
<td>(62 schools)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-SES</td>
<td>49.46</td>
<td>2.92</td>
<td>43.02</td>
<td>56.83</td>
</tr>
<tr>
<td>(62 schools)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The means for the two SES subgroups within each school effectiveness group are presented in Table 4.15. The widest gaps between middle- and low-SES were found in the effective school group. In the ineffective and typical schools there was approximately 4 points difference between the means for middle- and low-SES. In the effective schools there was a 5 point difference. Likewise, the larger standard deviations were in the effective schools. In looking at the highest score for the middle-SES group, again it appeared that this is the group which is widening the
dispersion by scoring proportionally higher than any other group, while the low-SES were falling behind in effective schools. Where the highest score for the low-SES group was no higher in the effective schools than in the typical schools, the highest score for the middle-SES students was almost 6 points higher than in the typical schools.

Table 4.15

Comparison of Mean Student Achievement Across School Effectiveness Categories for Middle- and Low-SES

<table>
<thead>
<tr>
<th>SCHOOL EFFECTIVENESS CATEGORY</th>
<th>SES LEVEL</th>
<th>M</th>
<th>SD</th>
<th>LOWEST SCORE</th>
<th>HIGHEST SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective Schools (54 Schools)</td>
<td>Middle-SES</td>
<td>57.05</td>
<td>2.48</td>
<td>52.58</td>
<td>62.31</td>
</tr>
<tr>
<td></td>
<td>Low-SES</td>
<td>52.08</td>
<td>2.13</td>
<td>46.21</td>
<td>56.83</td>
</tr>
<tr>
<td>Typical Schools (54 Schools)</td>
<td>Middle-SES</td>
<td>53.89</td>
<td>1.37</td>
<td>49.19</td>
<td>56.44</td>
</tr>
<tr>
<td></td>
<td>Low-SES</td>
<td>49.68</td>
<td>1.63</td>
<td>46.61</td>
<td>55.90</td>
</tr>
<tr>
<td>Ineffective Schools (54 Schools)</td>
<td>Middle-SES</td>
<td>50.57</td>
<td>1.52</td>
<td>46.33</td>
<td>53.90</td>
</tr>
<tr>
<td></td>
<td>Low-SES</td>
<td>46.63</td>
<td>1.86</td>
<td>43.02</td>
<td>51.03</td>
</tr>
</tbody>
</table>
The influence of the three school effectiveness groups on the mean of student achievement for low-SES students (Table 4.16a) and for middle-SES students (4.16b) yielded significantly different means in both cases. Consequently, the Student-Newman-Keuls Test was conducted on these means (Table 4.17a and 4.17b).

Table 4.16a

Results of Mean Student Achievement Across School Effectiveness Categories for Low-SES Students

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F-Ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>806.04</td>
<td>403.02</td>
<td>113.39</td>
<td>.0001</td>
</tr>
<tr>
<td>Error</td>
<td>159</td>
<td>565.15</td>
<td>3.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>161</td>
<td>1371.18</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.16b

Results of Mean Student Achievement Across School Effectiveness Categories for Middle-SES Students

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F-Value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>1134.04</td>
<td>567.02</td>
<td>164.33</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>159</td>
<td>548.61</td>
<td>3.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>161</td>
<td>1382.65</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The post hoc analysis indicated that the means of student achievement for low-SES students were different for each school effectiveness category. The middle-SES students also had significantly different means for the differentially effective schools. It was also interesting to note that the low-SES students in effective schools outscored the middle-SES students in ineffective schools, as was found in Teddlie et al., (1984).

Table 4.17a

Results of Student-Newman-Keuls Test for Mean Student Achievement of Low-SES Students by School Effectiveness Group

<table>
<thead>
<tr>
<th>SCHOOL EFFECTIVENESS GROUP</th>
<th>N</th>
<th>M</th>
<th>GROUPING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective Schools</td>
<td>54</td>
<td>52.08</td>
<td>A</td>
</tr>
<tr>
<td>Typical Schools</td>
<td>54</td>
<td>49.68</td>
<td>B</td>
</tr>
<tr>
<td>Ineffective Schools</td>
<td>54</td>
<td>46.63</td>
<td>C</td>
</tr>
</tbody>
</table>

*Note. Means with the same letter are not significantly different.*
Table 4.17b

Results of Student-Newman-Keuls Test for Mean Student Achievement of Middle-SES Students by School Effectiveness Group

<table>
<thead>
<tr>
<th>SCHOOL EFFECTIVENESS GROUP</th>
<th>N</th>
<th>M</th>
<th>GROUPING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective Schools</td>
<td>54</td>
<td>57.05</td>
<td>A</td>
</tr>
<tr>
<td>Typical Schools</td>
<td>54</td>
<td>53.89</td>
<td>B</td>
</tr>
<tr>
<td>Ineffective Schools</td>
<td>54</td>
<td>50.57</td>
<td>C</td>
</tr>
</tbody>
</table>

Note. Means with the same letter are not significantly different.

Results of analysis on within school variance of achievement.

The comparison of average within school variance for the three school effectiveness groups showed very little overall difference (Table 4.18). The group effect for the ANOVA confirms that there was no difference in the variation of scores for effective, typical, and ineffective schools (Table 4.19).
Table 4.18

Comparison of Average Within School Variance of Student Achievement for Effective, Typical, and Ineffective Schools

<table>
<thead>
<tr>
<th>SCHOOL EFFECTIVENESS CATEGORY</th>
<th>MINIMUM VARIANCE</th>
<th>MAXIMUM VARIANCE</th>
<th>AVERAGE VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective Schools (54 schools)</td>
<td>47.58</td>
<td>119.85</td>
<td>78.21</td>
</tr>
<tr>
<td>Typical Schools (54 schools)</td>
<td>49.45</td>
<td>110.42</td>
<td>75.63</td>
</tr>
<tr>
<td>Ineffective Schools (54 schools)</td>
<td>49.84</td>
<td>113.27</td>
<td>75.05</td>
</tr>
</tbody>
</table>

Note. Average variance is calculated by dividing the sum of the within school variance for all schools in that group by the n for that group.

However, a difference was found when comparing the within school variance of student achievement for the two SES subgroups of students.

The average variance for the middle-SES students was larger than that of the low-SES students (Table 4.20). The SES effect for the mixed factorial ANOVA substantiated this difference (Table 4.19). There was no significant interaction effect.
Table 4.19

Results of Analysis of Within School Variance of Student Achievement for School Effectiveness Groups and SES

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F-Value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>2</td>
<td>0.01</td>
<td>0.007</td>
<td>0.12</td>
<td>ns†</td>
</tr>
<tr>
<td>Error</td>
<td>159</td>
<td>10.16</td>
<td>0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES</td>
<td>1</td>
<td>0.55</td>
<td>0.55</td>
<td>18.06</td>
<td>.0001</td>
</tr>
<tr>
<td>Group * SES</td>
<td>2</td>
<td>0.09</td>
<td>0.05</td>
<td>1.53</td>
<td>ns†</td>
</tr>
<tr>
<td>Error</td>
<td>159</td>
<td>4.87</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† not significant

Table 4.20

Comparison of Average Within School Variance of Student Achievement for Middle- and Low-SES Students

<table>
<thead>
<tr>
<th>SES LEVEL</th>
<th>MINIMUM VARIANCE</th>
<th>MAXIMUM VARIANCE</th>
<th>AVERAGE VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle-SES</td>
<td>38.90</td>
<td>138.33</td>
<td>80.81</td>
</tr>
<tr>
<td>(162 schools)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-SES</td>
<td>32.78</td>
<td>112.32</td>
<td>71.78</td>
</tr>
<tr>
<td>(162 schools)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Average variance is calculated by dividing the sum of the within school variance for all schools in that group by the n for that group.

The average of the within school variances of the different SES levels for the three different school effectiveness groups showed the largest
variance to be in the middle-SES effective school group. The smallest variance was found in the low-SES typical school group. As shown in Table 4.21, the average variance was larger for the middle-SES group in each school effectiveness category. The most substantial difference was found in the effective school group. In that group there was a gap of over 9 points between the middle- and low-SES groups. The typical schools produced very little difference between the variances of the two SES groups.

Table 4.21

Comparison of Average Within School Variance of Student Achievement by Middle- and Low-SES Subgroups for Effective, Typical, and Ineffective Schools

<table>
<thead>
<tr>
<th>SCHOOL EFFECTIVENESS CATEGORY</th>
<th>SES LEVEL</th>
<th>AVERAGE VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective Schools</td>
<td>Middle-SES (54 schools)</td>
<td>83.02</td>
</tr>
<tr>
<td></td>
<td>Low-SES (54 schools)</td>
<td>73.78</td>
</tr>
<tr>
<td>Typical Schools</td>
<td>Middle-SES (54 schools)</td>
<td>76.77</td>
</tr>
<tr>
<td></td>
<td>Low-SES (54 schools)</td>
<td>72.39</td>
</tr>
<tr>
<td>Ineffective Schools</td>
<td>Middle-SES (54 schools)</td>
<td>77.48</td>
</tr>
<tr>
<td></td>
<td>Low-SES (54 schools)</td>
<td>73.13</td>
</tr>
</tbody>
</table>

Note. Average Variance is calculated by dividing the sum of the within schools variances for all schools in that group by the n for that group.
Although there was no significant group effect on either middle- or low-SES subgroups, (Table 4.22a and 4.22b), there was an observable, nonsignificant difference in the effective school influence on middle-SES students compared to the typical and ineffective school variance. The effective school group appeared to have the influence of widening the dispersion of scores for the middle-SES students. There was no significant difference found in the variances when tested by the $F_{max}$.

Table 4.22a

Results of Analysis of Within School Variance of Student Achievement Across School Effectiveness Groups for Low-SES Students

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F-Ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>0.02</td>
<td>0.008</td>
<td>0.19</td>
<td>ns†</td>
</tr>
<tr>
<td>Error</td>
<td>159</td>
<td>7.28</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>161</td>
<td>7.30</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† not significant.
Table 4.22b

Results of Analysis of Within School Variance of Student Achievement Across School Effectiveness Groups for Middle-SES Students

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F-Value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>0.09</td>
<td>0.05</td>
<td>0.93</td>
<td>ns†</td>
</tr>
<tr>
<td>Error</td>
<td>159</td>
<td>7.75</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>161</td>
<td>7.84</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† not significant.

Issue IV

Analysis of means of student achievement for high-, middle-, and low-SES schools.

As mentioned in the first section of Chapter Four, an additional analysis was conducted, comparing the means of student achievement for middle- and low-SES students in the sample of schools used for Issue III, to the means of student achievement in the two groups of schools that were not used in this sample. Those two groups consisted of schools that had over 70% low-SES, and those schools that were less than 30% low-SES. These schools will be referred to as low- and high-SES schools, respectively. The sample used for Issue III will be referred to as middle-SES schools.
Knowing the strong relationship between SES and student achievement, it was expected that the low-SES schools would have lower means, and the high-SES schools would have higher means. What was of interest was how much lower the low-SES schools were than the middle and high-SES schools. Also of interest was whether the means were lower for both SES subgroups within the low-SES schools. Differences in variance were not relevant here because the comparison of variances was done by effectiveness categories. This issue is relevant to the rest of the study as it pertains to the equity issue. Hence, the examination of means is of primary importance in Issue IV.

As the number of schools in each group was different, descriptive analysis was used to compare these means. There were 26 schools in the high-SES group, 162 schools in the middle-SES group, and 53 schools in the low-SES group. A mixed factorial was run on the data, using Proc GLM, a statistical analysis that adjusts for unbalanced designs (Proc GLM in SAS Institute, Inc., 1985). Caution in interpreting the results was still felt necessary due to the extreme divergence of numbers. The results from the Proc GLM procedure did indicate significant differences between all means.
Results of analysis for Issue IV.

Table 4.23 shows the means for the different SES subgroups in high-, middle-, and low-SES schools. There was approximately two points difference in the mean student achievement for both middle- and low-SES subgroups of students in high versus middle-SES schools. The difference between the means of the low-SES schools as compared to the middle-SES schools was around three points.

Of major concern here was the fact that the middle-SES students attending low-SES schools did not score as high as the low-SES students in the high-SES schools. When looking at the scores for just the low-SES students or just the middle-SES students by the different school SES levels, there was a five point difference in the means of achievement between the high-SES and low-SES schools. This difference was greater than the difference between the SES subgroups within a school. In other words the overall socioeconomic make up of the school had even more effect on student achievement than the individual student socioeconomic status.
Table 4.23

Comparison of Mean Student Achievement by Middle- and Low-SES for High-, Middle-, and Low-SES Schools

<table>
<thead>
<tr>
<th>SCHOOL SES LEVEL</th>
<th>STUDENT SES LEVEL</th>
<th>M</th>
<th>SD</th>
<th>LOWEST SCORE</th>
<th>HIGHEST SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-SES Schools (26 Schools)</td>
<td>Middle-SES</td>
<td>55.48</td>
<td>2.50</td>
<td>50.19</td>
<td>59.36</td>
</tr>
<tr>
<td></td>
<td>Low-SES</td>
<td>51.15</td>
<td>2.49</td>
<td>44.42</td>
<td>55.02</td>
</tr>
<tr>
<td>Middle-SES Schools (162 Schools)</td>
<td>Middle-SES</td>
<td>53.74</td>
<td>3.26</td>
<td>46.33</td>
<td>62.31</td>
</tr>
<tr>
<td></td>
<td>Low-SES</td>
<td>49.40</td>
<td>2.95</td>
<td>43.02</td>
<td>56.83</td>
</tr>
<tr>
<td>Low-SES Schools (53 Schools)</td>
<td>Middle-SES</td>
<td>50.86</td>
<td>4.10</td>
<td>43.06</td>
<td>59.71</td>
</tr>
<tr>
<td></td>
<td>Low-SES</td>
<td>46.36</td>
<td>3.38</td>
<td>40.64</td>
<td>55.61</td>
</tr>
</tbody>
</table>
Chapter Five: Discussion

Summary

Three different issues were involved in this study. The first issue dealt with the means and variances of teacher behavior for teachers in effective schools versus teachers in ineffective schools. Until the studies were done by Teddlie et al., (1989), and Virgilio et al., (1991), very little research had been done linking teacher behavior and school effectiveness. These two studies found that, not only did teachers in effective schools have more positive means for teacher behavior; they also had smaller variance of teacher behavior. These studies did not include any further qualitative data to establish explanations for these differences. Consequently, Issue I sought to replicate the findings of the other two studies, and go a step further in investigating possible causes for the differences.

Schools were classified as effective/ineffective by using residual scores from a regression analysis. Those schools with CRT scores that were .75 standard deviations above what was predicted by the socioeconomic status of the school were considered effective. Those schools .75 standard deviations below prediction were categorized as ineffective. Six effective and six ineffective schools were then chosen for the sample for Issue I.
Five teachers from each school were selected from grade three and grade five.

Observations were made of the teachers during which seven dimensions of teaching behavior were evaluated: interactive time-on-task, overall time-on-task, classroom management, instructional strategies, presentation/questioning, social/psychological environment of the room, and physical attributes of the room. In every case, the mean teacher behavior was higher for the teachers in the effective school group. An examination of the coefficients of variation indicated that the dispersion of scores for the teachers in effective schools was less than for teachers in ineffective schools. These results constitute a strong replication of those studies mentioned earlier.

Interviews with both the teachers and principals were also conducted to find possible explanations for the differences found in the means and variation of teacher behavior. The results of the teacher interviews indicated that: (a) there was a better process of teacher socialization in effective schools (thus replicating a recent study by TerHaar, 1992), (b) teachers work together more in effective schools, (c) principals in effective schools made more changes when taking over as principal,
(d) principals are seen more frequently in the classrooms of effective schools, and (e) effective schools are more likely to have enforced schoolwide discipline policies (quite often with assertive discipline methods). In ineffective schools there was also mention of teachers being selected from those who were student teachers at the school. This was never mentioned in effective schools.

The principal interviews confirmed the findings from the teacher interviews regarding changes made in the schools, teachers working together, and schoolwide discipline. Other differences found in the responses of principals from effective schools versus responses of principals from ineffective schools indicated that principals in effective schools (a) do more to work with teachers who have problems, (b) select teachers based on personality traits as opposed to philosophies of education and discipline, and (c) have future goals for the school that more directly pertain to student learning. Principals from ineffective schools more often complained of not finding time for instructional leadership, although all principals felt that this was the most important role of a principal.

Issues II and III both dealt with the equity concern in education. Equity is an issue that appears to be talked about more than it is researched. Levine & Lezotte (1990) devote several pages of their review
to the issue of equity and disaggregation of data. They point out that little systematic research has been conducted on this issue but that "...calling attention to achievement discrepancies by social class and/or race/ethnicity can both point the way toward needed changes and begin to build in accountability for initiating such changes" (p.7).

Those who advocate equity feel that more should be done to ensure that low socioeconomic and minority students are getting an education equal in quality to that which is received by the high-SES students. Proponents of equity advocate extra efforts being exerted to reach disadvantaged students and improve their achievement scores. Along with this line of reasoning is the belief that effective teachers and schools should not only be able to raise the overall achievement scores of students, but bring up the scores of the minority students to more closely reflect the performance of the entire group (National Education Goals Panel, 1991).

If this were actually happening, one would expect to find smaller variation in scores for effective teachers and effective schools. The means and variances of student achievement were examined to explore this matter.

Issue II was concerned with the means and variances of student achievement for different SES levels of students who were taught by
effective, typical, and ineffective teachers. The teacher observations from Issue I were used to classify the teachers as effective, typical, and ineffective. The means and variances of the criterion-referenced test scores for both low- and middle-SES students were compared for these three teacher effectiveness groups.

The results of the 3 X 2 mixed factorial ANOVA indicated that there was a difference in mean student achievement for students taught by effective/typical/ineffective teachers. There was also a difference found between the means of student achievement for the middle-SES students versus the low-SES students. When the mean scores were separated by teacher effectiveness group and SES level, there were differences found between all teacher effectiveness groups for low-SES students, and differences found between effective teachers versus typical and ineffective teachers in the middle-SES group. For the middle-SES students there was no difference in the means of student achievement for the typical versus the ineffective teachers.

No significant differences were found in the within teacher variance of student achievement. The trend was in the direction that the proponents of equity would expect, with smaller average variance for effective teachers than for typical or ineffective teachers. When taught by effective
and ineffective teachers, there was less dispersion of scores for middle-SES students than for low-SES students.

Issue III examined the means and variances of student achievement for students in effective, typical, and ineffective schools. For this sample, all elementary schools in the state that contain both grades three and five were used in a regression analysis predicting the CRT scores from socioeconomic and demographic variables. The schools were then classified as effective, typical, or ineffective based on their residual scores.

As the main concern was the possibility of differential treatment for different SES subgroups within a school, only schools with a mixture of both low- and middle-SES students were selected. The final sample included 54 schools in each school effectiveness category. The results indicated differences in means of student achievement for all groups when divided by both SES and school effectiveness categories.

When examining the variance of student achievement for the differentially effective schools, a slightly larger average variance was found in the effective schools. This difference was not significant. When using the variances as data points in a 3 X 2 mixed factorial ANOVA, a significant difference was found between the variances of student
achievement for the two different SES levels of students, with the middle-SES having the larger dispersion of scores.

**Limitations of the Study**

The sample of schools for Issue I may be considered incomplete in that it only compares outlier groups (effective versus ineffective). Purkey and Smith (1983) feel that it is important to compare the outlier groups to the schools that are considered average or typical. Two considerations influenced the choice of simply using effective and ineffective school groups. The first consideration had to do with the nature of the analysis for Issue I. As part of Issue I involved qualitative analysis, the outlier approach was preferable in order to distinguish different qualities found in the differentially effective schools. The second consideration was totally practical. As Issue I involved extensive time in the schools conducting observations and interviews, twelve schools were felt to be the maximum that could possibly be included. It was deemed more appropriate to have six schools in each group, rather than to divide the sample into three groups with only four schools per school effectiveness category.

One concern that came to mind when conducting the teacher and principal interviews was that the answers seemed to too closely reflect the results of research. The thought occurred, but was in no way confirmed,
that these teachers and principals have now been educated enough
concerning the correlates of effective schools to provide the pat answers
that research dictates.

A second limitation of this study is that the sample of schools used in
Issue I and Issue II was selected from an area to which it was convenient
to travel. This sample did, though, include both metropolitan schools
(from a city of around 300,000) and rural schools. The means for this
sample are slightly lower than the means for Issue III, which uses the
entire state. This indicates that the sample is not entirely representative of
the entire state, and may or may not generalize to the nation.

Another problem when considering generalizability for Issues II and III
is the fact that the achievement test scores used in this study were from a
state developed grade level test. Issue III did combine the results of
grades three and five, but if other grade levels were used these results may
not be generalizable. Also, if a basic skills or norm-referenced test were
used as a measure of achievement, different results may be expected.

There is some question as to the reliability of the socioeconomic and
demographic data that is provided along with the test results. This data is
provided by either the teacher or the student at the time of the test.
Understandably, the major focus at that time is the test itself, and the
additional information may not have been afforded the attention that it would have received at another time.

The test scores themselves may involve some inaccuracies. The test scores have recently taken on a new importance, as the state has begun a school accountability program, and an incentive award program. Both programs have test results as a major focus. This, along with newspaper publication of test results, may have led to unethical test procedures (e.g., Stringfield & Teddlie, 1991). After finding some cases of excessive erasures, erasure analysis was conducted on schools for the 1989-90 and 1990-91 school years. Only the most extreme of cases were eliminated, though, leaving many that may have been questionable.

Another debatable practice that may have influenced the test scores was that of teaching to the test, which includes drilling the students with parallel items (Mehrens, Phillips, & Schram, 1992). As the observations done for Issue I were in some cases conducted a short time before the administration of this test, these procedures were observed.

The results of the qualitative analysis for Issue I would have been stronger had the question been addressed with quantitative methods as well. Patton (1990) recommends triangulation of data (use of both
qualitative and quantitative analysis) in order to further substantiate the results.

Another limitation of this study is due to the fact that one of the main focuses of this study was on making inferences about variances. The use of variance as the dependent variable is not a common practice in educational research (although it does appear to be more common in kinesiology). One limitation to methods of statistical inference regarding variances is that they are generally restricted to single factor designs. The other problem concerning statistical methods for testing variance is that the methods are extremely sensitive to distribution form. Martin & Games (1977) offer some possible alternatives to the existing methods. Both alternatives involve ANOVA techniques on logarithmic transformations. They admit to the lack of efficiency of the methods proposed, but predict that as researchers become more interested in testing for variation of scores, that "more enlightening designs for investigating variances will be seen in the research literature". (p. 204)

Unfortunately that has not occurred. There are studies that have demonstrated interest in differences in variance (i.e., Lark, et al., 1984; Virgilio, et al., 1991), but have limited their analysis to the purely descriptive. This points to a real need for investigation of possible
statistical methods to use in analyzing variance differences. This was not, however, the purpose of this study. There is no certainty that the most robust and powerful tests were used to analyze the data in Issue II and III. Consequently, even where statistical significance was not found, descriptive analysis was done when differences were apparent.

It must also be taken into consideration that this is a causal-comparative design. The analysis was done on already existing groups. The comparison can imply correlational relationships between school effectiveness and teacher behavior, teacher effectiveness and student achievement, and school effectiveness and student achievement. No cause and effect relationship can be established with this type of design.

**Conclusions Regarding Issue I**

The findings concerning the means of teacher behavior for teachers in effective schools versus teachers in ineffective schools in most cases confirm the findings of both the Teddlie et al., (1989) and Virgilio et al., (1991) studies. The only differences were that the two previous studies found more differences in time-on-task than were found in this study. This lack of significant difference, especially for overall time-on-task, was unexpected. There is no apparent explanation for this outcome. In the Virgilio et al., (1991) the mean overall time-on-task for teachers in
effective and typical schools, respectively, was .79 and .74. On the other hand, the teachers in ineffective schools only had a mean score of .58 for overall time-on-task. Both means in this study (.73 and .70) are closer to the effective and typical school groups in the Virgilio.

One possible explanation may be that the students were on their best behavior in the presence of the observers. These observations were conducted around the same time that a much publicized statewide teacher evaluation program was taking place. In some classes it was quite apparent that the students had been warned ahead of time to be on their best behavior in the presence of observers. Although it was stressed that this was not an evaluation, it was not uncommon to be asked by the students if they had done well.

Differences found in the coefficients of variation of teacher behavior for the effective schools versus the ineffective schools also confirms the findings of the other two previous studies. Not only were the coefficients of variation smaller in the effective schools, the range of scores was truncated at the lower end, a phenomenon discussed by Virgilio, et al., (1991). Something was obviously happening in the effective schools to eliminate the most undesirable teacher behaviors. This was especially true
on the two dimensions of classroom instruction. The lowest scores for Instruction1 and Instruction2 in the effective schools were 2.74 and 2.75, whereas the lowest scores for these dimensions in the ineffective schools were 1.28 and 1.38.

Another fact that confirms these findings is that, of the twenty teachers that were rated as "ineffective", only five of those were in schools that were rated as "effective". In other words, 75% of the ineffective teachers were teaching in ineffective schools. Hence, there is less inadequate teaching behavior in effective schools.

The results of the qualitative analysis provided insights as to what transpired in effective schools to alleviate these undesirable behaviors. One theory posed by Virgilio et al., (1991), was that teacher selection/dismissal may have some affect on these differences. This was confirmed by three results of the interviews done for this study. One point was that ineffective schools hired from their student teachers, where effective schools never mentioned student teachers being a source of new teachers. These results coincide with reports of Teddlie and Stringfield (in press) that the effective schools may have a better chance of attracting more experienced, effective teachers, who are seeking to transfer into a more desirable teaching environment. Less effective schools may be
limited in their recruiting to novice teachers who have little choice in their job assignments.

When asked what they looked for when interviewing perspective teachers, the typical answers heard from the principals of ineffective schools were simply not those given by the principals of effective schools. Whereas principals in ineffective schools asked about philosophies of teaching and discipline, the principals in effective schools were looking for specific personality traits. It appears that, besides having better choices of potential teachers, these principals in effective schools may also have more insight into selecting teachers who will demonstrate better teaching behavior.

The dismissal procedures also yielded differences in effective versus ineffective schools. Both effective and ineffective school principals reported a few teachers that did not fit in with their philosophies of education. The way they were dealt with differentiated the two groups. The principals in the effective schools reported spending considerable time and effort on working with these teachers. On the other hand, the principals in ineffective schools more often reported trying to 'get rid of them'; an extremely difficult, time consuming process that is often unsuccessful (Bridges, 1986).
These responses not only reveal information about the dismissal process, but also about the teacher induction process and the overall atmosphere of the schools. Obviously, a place where inadequacies are directly addressed is much more likely to alleviate ineffective behavior than a place where the teacher is simply 'gotten rid of'.

Another school process believed to have an influence on eliminating inadequate teacher behavior is that of teacher socialization (Virgilio et al., 1991). The results of this study confirm this theory. The fact that teachers in effective schools were mentioned more often as helping new teachers fit in and informing them of school policies is consistent with the findings of a recent study on teacher induction (TerHaar, 1992). TerHaar found that teachers in effective schools were much friendlier and more willing to help novices than those from ineffective schools.

TerHaar also reported that the new teachers in effective schools rated their schools higher regarding shared teaching goals. The fact that this study found more mention, from both teachers and principals, of the teachers working together on the school’s goals substantiates the findings.

The support and help of other teachers would definitely have an effect on teacher behavior. When a new teacher is having a problem, colleagues with whom they feel comfortable are the most likely to be sought out for
help and advise. Without this interaction between other teachers, ineffective teaching practices are more likely to continue.

The teacher induction study by TerHaar also reported differences in effective, typical, and ineffective schools on the school’s success at managing student behavior. This, also, is consistent with the findings of this study. Both teacher and principal interviews confirmed that effective schools have more schoolwide discipline programs. The interviews also revealed that effective schools are more likely to use an Assertive Discipline Program. The fact that more teachers in ineffective schools mentioned establishing classroom rules by discussing them with their students shows an attempt by the teachers to establish the order in the classroom that is lacking in the school.

Unfortunately, without effective classroom management, very little learning can take place. Good discipline practices is a technique that can be learned. Some teachers have a natural ability to effectively manage a classroom, but many other teachers have learned through workshops and help from their principals. Having schoolwide policies that enforce the policies of the classroom appears to be a very important factor in increasing the effectiveness of the school and improving the performance of the teachers.
Effective school leadership, generally by the principal, is believed by most researchers to be imperative to a school’s effectiveness (Geske & Teddlie, 1990). Good and Brophy (1986) reported that principals could directly influence the school in the realm of goal orientation, instructional leadership, and assignment of students and teachers to classrooms.

This study found differences in the principals of effective schools versus ineffective schools in both their goal orientation and instructional leadership. Besides the above mentioned differences found in the teacher’s participation in goal setting, the future goals of the principals support the apparent differences in goal orientation of effective schools versus ineffective schools. The finding that the effective school principals had future goals for the school that more directly involved student learning agrees with the results reported by Good and Brophy (1986).

When asked what the most important role of the principals was, all responded "instructional leadership". This was one of those responses that might be a conditioned response, rather than a true belief. The fact that three of the principals in ineffective schools went on to add that they had very little time for this role indicates that the ineffective schools are lacking in instructional leadership.
Stronger instructional leadership in effective schools is also substantiated by the findings of a greater frequency of informal visits to the classroom by the principals in effective schools. Both districts in the sample had mandatory teacher evaluation by the principals; consequently, there was no difference found in that respect. The principal's presence in the classroom is not a direct indicator of instructional leadership (they could be there only to check on discipline or other matters), but the absence of the principal in the classroom in ineffective schools is an indication of a lack of instructional leadership. Whether the principal is in the classroom to observe instruction, for classroom management, or simply for social reasons, the more he/she is in the classroom and aware of what goes on, the more likely he/she is to see and implement changes of undesirable behavior of all sorts.

The number and type of changes made by principals when taking over effective schools versus ineffective schools also points to the stronger overall leadership provided by principals in effective schools. If a school were extremely effective when the principal took over, it might be considered a good policy to not make many changes. It is interesting to note, however, that ineffective schools had the fewest reported changes made by new principals. Also, the changes that were mentioned in
ineffective schools (i.e., less pressure, more organization, more committees) had very little to do with factors that are known to directly influence the school's effectiveness. On the other hand, the goals of principals of effective schools (more learning time, increased achievement test scores) are directly related to effective schooling.

No differences were found in the assignment of teachers or students to the classroom. The most common and approximately equally occurring response in both school effectiveness groups was that the principal did the selection along with the teacher's recommendation. There was one effective school where the principal did the selection alone. Two negative responses came from teachers in ineffective schools complaining about the group that they had. One reply to how the students were assigned to the classroom was "I sure would like to know after this group". Perhaps the wording of the question was not specific enough to discern the differences reported by Good and Brophy (1986).

All of the above differences which were found in effective schools versus ineffective schools are believed to have a direct influence on both the means and variances of teacher behavior. The help and support of both colleagues and the principal of the effective schools obviously had a
direct affect on teacher behavior; especially in the elimination of incompetent behavior.

**Conclusions Regarding Issue II**

The differences found in the means of student achievement for effective, typical, and ineffective teachers shows that teacher behavior does have an influence on the student achievement for both middle- and low-SES students. The results of this study indicate that the teacher behavior has a greater impact on low-SES students than on middle-SES students. The means for each teacher effectiveness group were different for low-SES students, where the only differences found for middle-SES students was between the effective teachers versus the typical and ineffective. There was no difference between the typical and ineffective teacher categories. Only the most effective of teachers appear to make a difference to the middle-SES students.

This seems reasonable when one considers that the middle-SES students are more likely to have help and support at home. Hence, the effect of the teacher is not as essential to their learning as it would be to the low-SES students whose primary chance to learn is at the school.

The larger gap between SES subgroups when taught by ineffective teachers also points to the lack of help in the home for the low-SES
students. The middle-SES students taught by ineffective teachers are still managing to score quite close to average (presumably with the help they are getting outside of the classroom); whereas, without the help of either the home or the teacher, the low-SES students’ scores are far below the average.

It is apparent that, even though low-SES students scored lower than the middle-SES students, the teacher does have a definite influence on the achievement of these low-SES students. The low-SES students taught by effective teachers scored the same (or, to be exact, .42 points higher) than the middle-SES students taught by ineffective teachers. This substantiates the school level findings of Mortimore and Sammons (1987) and Teddlie et al., (1984) in which they both reported that effective low-SES schools outscored ineffective middle-SES schools.

The results of the analysis on the variation of achievement scores for the different teacher groups indicates that effective teachers are not doing significantly better at reducing the gap between low- and middle-SES subgroups than the ineffective teachers. As mentioned in chapter 4, however, there is a noticeable difference in the average variance of student achievement for the effective teachers: it being smaller than the variance for typical and ineffective teachers. This, at least, points in the direction
that would be hoped for by equity advocates. This also goes along with the findings regarding Issue I where there was less variance of teacher behavior in effective schools. The effective teachers are doing, to some extent, what the effective schools did for the teachers: elevating the students to a higher quality of achievement.

When comparing the variance of scores between low- and middle-SES, there is again no significant difference. The within class variance of student achievement is the same for both SES subgroups.

Conclusions Regarding Issue III

The results of the means of student achievement at the school level are again as expected. The mean for typical schools is higher than the mean for ineffective schools, and the mean for effective schools is higher than for typical schools. There is approximately the same amount of difference between each school effectiveness group.

Also as expected, the low-SES group scored lower than the middle-SES group within each school effectiveness category. Consistent with Teddlie et al., (1984) and Moritmore and Sammons (1987), the low-SES students in effective schools outscore the middle-SES students in ineffective schools. At the school level there is about 1.5 points difference
between the low-SES effective school group and the middle-SES ineffective school group.

In this study, the largest gap between SES subgroups is found in the effective school category. In the ineffective and typical schools there is around 4 points difference between the two groups. In the effective schools the gap expands to 5 points. When looking at the low-SES group alone, the biggest difference is seen between the means of student achievement for the ineffective schools compared to the typical schools (3 points). There is only 2.4 points difference between the typical and effective school groups for low-SES students. The opposite happens for the mean student achievement for middle-SES students. The biggest difference is between the typical and effective schools.

Also when examining the lowest and highest scores, one can see that there is very little difference between the typical and effective schools for the low-SES students, and that the greatest difference is found between the typical and effective schools for the middle-SES students. In other words, the results indicate that the effective schools are being differentially more effective for the middle-SES students than for the low-SES students. Where the low-SES students are falling slightly behind, the middle-SES are moving ahead to further the gap between the subgroups.
The analysis of the variance between the three school effectiveness categories yielded no significant differences. As just discussed, the effective schools, if anything, are widening the dispersion of scores. This does, to some extent, concur with the findings of the Lark et al., (1984) study. Although there was no significant difference found in this study, the only analysis done in the Lark study was descriptive. When examining the average variance of student achievement for the middle-SES students, the effective schools have a noticeably larger dispersion of scores than the typical or ineffective schools.

This is, obviously, the opposite of what supporters of equity would hope to see. Lark argues for maximizing each student’s potential, which would increase the dispersion of scores. Those who believe in equity in education would like to see the scores of the minority and low-SES students brought up to more closely reflect the achievement of the population, which would decrease the dispersion of scores. Although much talk has been done concerning equity, this study does not indicate that the effective schools are doing anything to reduce this gap. This is also consistent with the report by Lee and Bryk (1989) that "individual school factors that produce high average achievement and contribute to
internal differentiation are more common than those that are simultaneously associated with high achievement and social equity". (p. 190)

The fact that there was a difference found in the variance of middle-SES students versus low-SES students is again attributable to the large differences seen in the effective school category. Where the middle-SES variance is around 4 points larger than the low-SES variance in the ineffective and typical school categories, there is over 9 points difference between the variances in the effective school group (with the middle-SES variance again being the larger of the two). Possible explanations for this difference will be explored in the section which compares Issue II and Issue III results.

Conclusions Regarding Issue IV

Although it is a known fact that socioeconomic status has a direct affect on student achievement, the findings of Issue IV were somewhat surprising. This involved the comparison of student achievement for middle- and low-SES students in high-, middle-, and low-SES schools. The average SES level of the school has as much affect on the achievement as the actual SES level of the students. When looking at the difference in achievement scores between the low-SES and high-SES
schools, there is between 4.5 and 5 points difference for both SES levels of students. On the other hand, the largest difference between middle- and low-SES students within these categories is 4.5 points in the low-SES schools.

Three possible explanations for these inequities between SES levels of the schools are that it is caused by (1) the student composition of the schools, (2) the differences in financial aid for the schools, or (3) the process within the schools. In the high-SES schools, the low-SES students make up 22% of the entire population. In the low-SES schools, the low-SES students constitute 79% of the entire population. Since low-SES has a strong correlation with low ability (Coleman, 1966; Teddlie et al., 1989; Useem, 1990), the research findings on the advantages and disadvantages of ability grouping can be considered here to offer explanations for these findings.

 Depending on the study, or group of studies, chosen one can find reports that ability grouping has no affect on student achievement (Kulik & Kulik, 1982; Slavin, 1987); has a positive affect on high ability student achievement (Feldhusen, 1989); or, in most cases, has a negative effect on student achievement (Dawson, 1987; Oakes, 1985; Winn & Wilson, 1983). There is enough research to support the fact that low ability students have
higher achievement when in a class with higher ability students (Cotton & Savard, 1981). If you support the fact that low ability is synonymous with low-SES, then a school with all low-SES (hence, low ability) students would also be detrimental to raising the achievement of these students.

Although this state does have more state funding of schools than many other states, there are still additional local funds provided to schools that make the low-SES schools less financially advantaged. Even though research has been found that teacher salaries and new buildings do not have that much relationship to student achievement (Purkey & Smith, 1983), the inequities in financial aid and capabilities of these schools could certainly not help the low-SES schools. If anything, more money needs to go into the low-SES schools, not less. More money does not ensure effectiveness, but may provide ways to attract better administrators and teachers in order to improve these schools.

Edmonds (1979) demonstrated 13 years ago that it is possible to have effective low-SES schools. The factors that he discusses that go into making a school effective for the urban poor are what would be considered process variables. These factors include strong administrative leadership, high expectations, orderly atmosphere, emphasis on acquisition of basic skills over other school activities, and frequent monitoring of progress.
Most of these processes can be instated without additional funds or new buildings. They do, though, take considerable desire and effort. Edmonds (1979) stated that "there has never been a time in the life of the American public school when we have not known all we needed to in order to teach all those whom we chose to teach". (p. 16). It appears that the choice still has been to not teach these students in low-SES schools.

Comparison of Findings for Issue II and Issue III

Although there was no hypothesis concerning any connection between Issue II results and Issue III results, it was expected that the same patterns would be found in the means and variances of student achievement at both the teacher and school level. There were some differences found that warrant exploration. Figure 5.1 summarizes these differences.

The differences in the results of the group means for Issue II compared to Issue III show that the differential effectiveness of the teacher and the school have similar influences on student achievement. There is slightly more differentiation in school effectiveness categories than in teacher effectiveness categories, as there was no significant difference between ineffective and typical teachers for middle-SES students. In other words, the schools have more of an effect on both SES subgroups than the teachers do.
When examining differences between SES subgroups, Issue II had the largest gap in mean achievement scores between the subgroups in the ineffective teacher group. In Issue III the biggest gap was found in the effective schools. This difference is consistent with the results of the variance analysis where the smallest variance was in the effective teacher groups and the largest variance was in the effective school group. Also, the largest gap in variances between the two SES subgroups is found in the effective schools; the middle-SES group having the greater dispersion of scores. The differences found in means and variances point to the same question: what is happening in effective schools to widen this gap in scores that does not happen within the effective classroom?

<table>
<thead>
<tr>
<th></th>
<th>MEANS</th>
<th>ISSUE II</th>
<th>ISSUE III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td></td>
<td>No differences between ineffective and typical teachers for middle-SES</td>
<td>Differences between all three groups</td>
</tr>
<tr>
<td>SES</td>
<td></td>
<td>Difference between SES subgroups. Largest gap for ineffective teachers</td>
<td>Differences between SES subgroups. Largest gap for effective schools</td>
</tr>
<tr>
<td>VARIANCE</td>
<td>Group</td>
<td>Less variance in effective teacher group</td>
<td>More variance in effective school group</td>
</tr>
<tr>
<td>SES</td>
<td></td>
<td>No difference between SES subgroups</td>
<td>Differences between middle-and low-SES</td>
</tr>
</tbody>
</table>

**Figure 5.1.** Comparison of Results for Issue II and Issue III.
In seeking possible explanations for this difference, the first concern was that perhaps the sample for Issue II was not representative of the entire state from which the sample for Issue III was taken. Hence, a school level analysis was done on the 11 schools used in the sample for Issue II. The results of this school level analysis were the same as the results for Issue III; the largest average variation of scores was found in the 5 effective schools, with middle-SES having greater variation, especially in the effective schools.

As the results from Issue II demonstrated that there is no difference in within class variance for effective, typical, or ineffective teachers, it was determined that the differences found in the school level variance must be due to differences in between class variance. One possible explanation, then, could be connected with the fact that some schools make assignments to classes based on ability grouping. It is not known how widespread this practice is. In schools where between class ability grouping is practiced, greater between class variance would be expected. It would be likely, though, to find ability grouping in the typical and ineffective schools as well, which would equally affect the variance in these groups. Consequently, this consideration alone does not seem to be a plausible solution.
Another consideration is that some schools have gifted and talented programs in which entire classes consist of gifted and talented. Unfortunately, many of the lower SES students are not as likely to be found in these classes. When schools have both a gifted and talented and regular education program this would certainly cause larger between class differences in variance.

One other possibility to consider is that some teachers in effective schools are challenging the middle-SES students to excel above the other classes. Even though there is less variance of teacher effectiveness in effective schools, it is still to be expected that some are far better than others. This, in combination with ability grouping, could be a plausible explanation. An excellent teacher with a higher ability group (which usually means higher SES group) could enable the students to achieve far beyond the average for that school.

Summary of Variance Results

As shown in Figure 5.2, the findings concerning variance of teacher behavior for effective versus ineffective schools parallels the findings of both Teddlie, et al., (1989) and Virgilio, et al., (1989). These studies found less variation in teacher behavior in the effective schools, with the truncation of range being caused by less ineffective teacher behavior.
There was no previous study found that compared variance of student achievement for differentially effective teachers. This study did find a slight trend toward smaller variance of student achievement in the effective teacher category.

When examining the variance of student achievement for differentially effective schools, this study more closely reflects the findings of the Lark et al., (1984) study. When doing a descriptive analysis, as Lark et al., (1984) did, this study also found evidence of a larger dispersion of scores in the effective schools than in the typical or ineffective schools. The Crone and Tashakkori (1992) study found no difference in standard deviations of scores until data was disaggregated by SES. Then the high-SES schools had smaller variation of scores in the effective school group, and the low-SES schools had smaller variation of scores in the ineffective school group. Two major differences are to be considered with the Crone and Tashakkori study: (1) it was conducted in middle school, and (2) it was comparing different SES levels of schools, rather than students.
<table>
<thead>
<tr>
<th>Issue</th>
<th>Lark</th>
<th>Teddlie, Kirby, &amp; Stringfield</th>
<th>Virgilio</th>
<th>Crone &amp; Tashakkori</th>
<th>Crone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issue I</td>
<td>Less variance of teacher behavior in effective schools</td>
<td>Less variance of teacher behavior in effective schools</td>
<td></td>
<td></td>
<td>Less variance of teacher behavior in effective schools</td>
</tr>
<tr>
<td>Issue II</td>
<td></td>
<td></td>
<td></td>
<td>Less variance of student achievement for effective teachers</td>
<td></td>
</tr>
<tr>
<td>Issue III</td>
<td>More variance of student achievement in effective schools</td>
<td></td>
<td>Less variance of student achievement in effective high SES schools, more variance of student achievement in effective low SES schools</td>
<td></td>
<td>More variance of student achievement in effective schools</td>
</tr>
</tbody>
</table>

Figure 2. Comparison of Variance Results to Other Studies.
Implications

Although the results of variance of student achievement are inconclusive, it is felt that the examination of variance is a helpful practice in exploring the equity issue. This should not take the place of disaggregation of data, but could be a useful tool in addition to data disaggregation. If disaggregation of data into subgroups had not occurred, it appears that some of the same conclusions could have been reached by examining the overall variance in scores.

The main implication to these findings is that equity is not being achieved. The effective teachers come the closest to reaching some degree of equity by reducing the dispersion of scores for the students in their classrooms, but the schools are not accomplishing this goal of some educators. It appears that concerted efforts are going to be necessary in order to find ways of reaching both low-SES students and low-SES schools. The fact that high-SES schools are achieving so much better than the low-SES schools, regardless of the SES level of the students, indicates that some serious inequity is occurring in our educational system. Whether these inequities are caused by differences in funding for the high-SES schools versus the low-SES schools, or whether it has to do with the attitudes and differential treatment in high-SES schools compared to low-
SES schools is unknown. The only certain conclusion that can be made here is that inequities in education certainly continue to exist.

**Recommendations for Further Research**

More qualitative research is needed to further investigate the process involved in making a school effective. Both the actions of the principal and the interaction of the teachers warrant more in-depth analysis. The interview questions used in this study yielded some interesting results. Some of the questions that, based on the literature, were expected to show differences (but didn’t) need further examination, such as assignment of students to classrooms and teacher evaluation procedures. Conducting interviews similar to the ones done in this study, but in other schools and with rewording of some questions is recommended. Separate questions about formal evaluations, and frequency of principal visits to the classroom may be enlightening. It also might be better to directly ask about changes the principal made when first coming to the school, then ask how the teachers found out about the changes. Finding a way to ask the questions to avoid getting 'pat' answers is also suggested. The question asked of the teachers regarding the school goals and of the principals regarding their future goals were the two that received the most insincere, homogeneous
responses. It is also advised that quantitative questions of the same nature be used to further substantiate the findings.

The variance issue needs much more investigation in different grade levels, and different areas. Further research is also recommended on the sample of schools used in Issue III. The identification and study of the particular effective schools that are widening the dispersion of achievement test scores is definitely needed. This would require sorting out those effective schools that have the largest variance and looking for any common characteristics (i.e., gifted and talented programs, ability grouping, urbanicity, size, etc.) that might contribute to this dispersion of scores.

Identification of the effective teacher characteristics that help reduce the variation of student scores would be extremely helpful in advancing the equity cause. Some additional work could be done with the existing data, finding those dimensions of teacher behavior that yielded the strongest relationships to the variance of student achievement. Qualitative observations of those effective teachers who had relatively small variation in achievement scores would also be recommended.

Also, additional research would be suggested regarding the findings of Issue IV. In this case, the group of schools with more that 70% low-SES students should be pinpointed for more extensive research. The best type
of research possible, if time and funds were available, would be to go into some of these schools, apply the correlates of effective schools set forth by Edmonds (1979) or some more recent derivation thereof (e.g., Levine and Lezotte, 1990), and see what can be done to improve the achievement scores of these low-SES students.

As these low-SES schools are in suburban and rural areas, as well as urban areas, comparison of effective and ineffective schools for the different community types would also be suggested. Although Edmonds (1979) feels that we know all that we need to know to make these schools effective, his studies were conducted for the urban poor. There may be other factors involved when dealing with suburban and rural areas, such as this sample includes.

It is also suggested that longitudinal studies be conducted on the variance, the gap reduction, and the equity issues. This study only examined what effective teachers and effective schools might be doing to reduce the gap in student achievement. It would be interesting to take national data bases, such as High School and Beyond or National Education Longitudinal Study, and look for any changes in the dispersion of scores over time.
References


Appendix A

Operational Definitions of Independent and Dependent Variables

School Effectiveness: Student achievement was used as the indicator of school effectiveness. This included language arts and math scores for grades three and five of the criterion-referenced test administered by the Louisiana Educational Assessment Program. These test scores were utilized in a multiple regression analysis, with indices of socioeconomic status used as the predictor variables. The residuals from the regression were used to label the schools as effective or ineffective. For this study +/-.15 was used as the cutoff point.

Teacher Effectiveness: The rating for teachers on seven dimensions of teacher behavior were used to operationalize teacher behavior. Two time-on-task (interactive and overall) indices were obtained with the use of the Classroom Snapshot portion of the Stallings Observation System. The other five dimensions (classroom management, instructional strategies factor, instructional presentation/questioning factor, classroom social/psychological environment factor, and classroom physical environment factor) were from the Virgilio Teacher Behavior Instrument.

Student Achievement: Scores on the grade three and five criterion-referenced test administered by the Louisiana Education Assessment Program were used as the measure of student achievement. Both language arts and math scores were used. For Issue II these two subject areas were transformed and combined. For Issue III both subject areas and grade levels were combined to produce an overall school mean and standard deviation.

Socioeconomic Status: Within each classroom (for Issue II) and each school (for Issue III) students were divided into two SES levels; middle and low. Family income and family size were the variables used for dissagregation.
Appendix B

Virgilio Teacher Behavior Inventory
(Virgilio, 1987)

The Virgilio Teacher Behavior Inventory was designed as an observational tool to measure specific teacher behaviors consistently described in teacher effectiveness research. The observation should be conducted in a regular classroom setting and last for an entire class period (50-60 minutes). The observer should rate each behavior according to the following rating scale:

1 - Poor
2 - Below Average
3 - Average
4 - Good/Above Average
5 - Excellent
6 - Not applicable/unable to observe

1. **Demonstrate Routine Classroom Management Techniques**
   1. The teacher clearly states rules and consequences.  
   2. The teacher uses time during class transition effectively.  
   3. The teacher uses student assistants to save time.  

2. **Maintains Appropriate Classroom Behavior**
   5. The teacher uses behavior incentive systems to manage student behavior (uses charts, tokens,... to keep students on task).  
   6. The teacher promptly handles inappropriate behavior.  
   7. The teacher continuously monitors the entire classroom.

3. **Focuses and Maintains Student Attention on Lesson**
   8. The teacher uses a motivating technique to focus on the lesson.  
   9. The teacher clearly states objective of the lesson.  
  10. The teacher presents new skill/material accurately.  
  11. The teacher presents detailed directions and explanations.  
  12. The teacher emphasizes key points of the lesson.

183
4. **Provides Students With Review and Practice**
   13. The teacher provides seatwork that is relevant to the lesson.
   14. The teacher guides individual practice.
   15. The teacher checks for understanding.
   16. The teacher summarizes the lesson.
   17. The teacher reteaches if student error rate is high.

5. **Demonstrates Skill in Questioning**
   18. The teacher uses a high frequency of questions.
   19. The teacher asks questions in an appropriate sequence.
   20. The teacher responds appropriately to students' questions/comments.
   21. The teacher probes further when responses are incorrect.
   22. The teacher uses appropriate wait time between questions and responses.

6. **Establishes Strategies of Evaluating Student Needs/Progress**
   23. The teacher identifies learners who need more assistance/training.
   24. The teacher assigns homework and provides feedback.
   25. The teacher provides a variety of activities to meet individual needs.

7. **Demonstrates a Variety of Teaching Methods**
   26. The teacher uses flexible grouping where appropriate.
   27. The teacher uses a variety of explanations that differ in complexity.
   28. The teacher uses a variety of teaching methods (peer tutoring, individual/small group instruction)
   29. The teacher uses manipulative materials/instructional aids/resources effectively (computers, manipulatives, fieldtrips...)

8. **Establishes a Positive Classroom/Learning Climate**
   30. The teacher communicates high expectations for students.
   31. The teacher exhibits personal enthusiasm.
   32. The teacher uses positive reinforcement techniques (nods, praises, avoids criticism or negative remarks...)
9. **Promotes Positive Self-Concepts in Learners**

33. The teacher encourages student interaction and communication.  

34. The teacher conveys genuine concern for students (empathetic, understanding, warm, friendly).  

35. The teacher knows and uses students’ names.  

10. **Creates Positive Classroom Environment**

36. The teacher displays students’ work in the classroom (ample amount, attractively displayed, current).  

37. The teacher prepares an inviting and cheerful classroom.  

38. The teacher prepares bulletin boards that are attractive, motivating, and current.
Appendix C

Teacher Interview Questions

1. How do new teachers get selected for this schools?

2. If your principal thinks a teacher is not doing a satisfactory job, what action does he/she take?

3. Did you come to this school before or after the present principal?
   If "after", go to question 4 & 5 (omit 6 & 7).
   If "before", go to questions 6 & 7 (omit 4 & 5).

4. When you first came to this school, how did you find out the school policies and procedures?

5. How were you made to "fit in" when you first came to this school?

6. When the new principal came, how did you find out the policies and procedures of the new principal?

7. In what way are new teachers made to "fit in"?

8. What kind of encouragement do the teachers get in the use of new teaching practices?

9. How do you know what your principal thinks about your teaching capabilities?
10. What is the primary goal of this school?

11. How did you find out about this goal?

12. What kind of input do you have regarding school goals?

13. Are you aware of any strategies that your principal uses in assigning students to classrooms? If so, what are they?

14. Do you notice any patterns in regard to the assignment of teachers to students?

15. How are the conduct rules established for your class?
Appendix D

Principal Interview Questions

1. How long have you been principal of this school?

2. When you first came, what kind of changes did you make in the policies and procedures?

3. How did you inform the teachers of these changes?

4. What kind of action did you take with teachers who did not subscribe to your philosophy of education?

5. How do new teachers get selected for this school? How much input do you have in their selection?

6. If a new teacher is having problems in the classroom, what do you do? (Probe-ask for specifics)

7. What do you do to encourage new teaching methods and practices?

8. How do you let the teachers know what you think of their performance? What kind of feedback system do you have in place?

9. What are the primary goals of this school? Do you feel that they have been attained?

10. How are those goals determined? Who had a role in the determination of those goals?
11. What future goals would you like to see accomplished for this school?

12. What kind of input do the teachers have regarding school goals, policies, etc.?

13. What strategies do you use to assign students to classrooms?  
   (Probe for specifics)

14. How are the conduct rules established for the schools?

15. Of the many roles a principal has to play, what do you consider the most important role of the principal?
Vita

The author of this dissertation, Linda J. Crone, graduated with a Doctor of Philosophy from Louisiana State University in December of 1992. She concentrated her graduate study in educational research and statistics. While serving as a graduate assistant at LSU she taught the graduate statistics lab (including ANOVA, ANCOVA, and Multiple Linear Regression), and assisted in research projects.

In addition to her Ph.D., the author has earned a bachelor of music education degree from Indiana University; and a master of music education degree from the University of Arizona. She is a certified music teacher in Indiana, and has taught music in the Indiana public schools for eight years.

The author is presently employed as a psychometrician with the Louisiana Department of Education. Her duties include analysis and interpretation of statewide test data for the Louisiana Progress Profiles Program (school report cards).
DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Linda J. Crone

Major Field: Education

Title of Dissertation: The Methodological Issues of Variance in Teacher Behavior and Student Achievement: The Relationship of Variance to School and Teacher Effectiveness

Approved:

[Signatures]

Dean of the Graduate School

EXAMINING COMMITTEE:

[Signatures]

Date of Examination: 8/20/1992