



SCHOOL of
GRADUATE STUDIES
EAST TENNESSEE STATE UNIVERSITY

East Tennessee State University
**Digital Commons @ East
Tennessee State University**

Electronic Theses and Dissertations

Student Works

12-2005

A Study of the Association between Multi-Age Classrooms and Single-Age Classrooms Regarding TCAP Reading/Language Gains.

Holly Irvin Flora
East Tennessee State University

Follow this and additional works at: <https://dc.etsu.edu/etd>

 Part of the [Educational Assessment, Evaluation, and Research Commons](#)

Recommended Citation

Flora, Holly Irvin, "A Study of the Association between Multi-Age Classrooms and Single-Age Classrooms Regarding TCAP Reading/Language Gains." (2005). *Electronic Theses and Dissertations*. Paper 1100. <https://dc.etsu.edu/etd/1100>

This Dissertation - Open Access is brought to you for free and open access by the Student Works at Digital Commons @ East Tennessee State University. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ East Tennessee State University. For more information, please contact digilib@etsu.edu.

A Study of the Association Between Multi-Age Classrooms and Single-Age Classrooms
Regarding
TCAP Reading/Language Gains

A dissertation

presented to

the faculty of the Department of Educational Leadership and Policy Analysis

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Doctor of Education

by

Holly Krysten Irvin Flora

December 2005

Dr. Louise MacKay, Chair

Dr. Mary Langenbrunner

Dr. Jasmine Renner

Dr. Terry Tollefson

Keywords: Nongraded, Multiage, Nongraded Education, Continuous Progress,
Developmentally Appropriate Practice

ABSTRACT

A Study of the Association Between Multi-Age Classrooms and Single-Age Classrooms

Regarding TCAP Reading/Language Gains

by

Holly Irvin Flora

The purpose of this study was to explore the differences between Reading/Language achievement gains of students in multi-age classrooms to the Reading/Language achievement gains of their peers in traditional, single-age classrooms. The causal-comparative quantitative approach to exploring cause-and-effect relationships was employed in this study. In this study, the effect of multi-age grouping and single age grouping was analyzed and compared using TCAP Reading/Language raw gain scores. Raw gain scores were used to determine the amount of progress children make from one year to the next regardless of their level of achievement. Findings in this study were mixed. Some significant differences were found in favor of single-age classrooms. However, the calculation of effect size showed no practical significance. Significance was also revealed in favor of males over females in both single-age and multi-age classrooms; although, effect size indicated only a small to moderate practical significance exists. This study provides an overview of the history of American educational structures. It might be helpful for the educational community in evaluating one dimension of the effectiveness of multi-age groupings. Teachers and administrators could benefit from the comparisons made in this study and as a result make better decisions regarding the delivery of instruction and the structuring of school classrooms.

DEDICATION

This study is dedicated:

To my husband Nathan and my daughter Elizabeth; The joy of your presence in my life has given me strength and encouragement to pursue my dreams.

To my parents, Rod and Ann Irvin, who have modeled lives of hard work, dedication, and passion.

To my sister, Emily Helphinstine, and to my brothers, Wally and Andy Irvin, who have provided me with support and laughter when needed most.

To my aunt, Becky Irvin, who has taught me to take time out to enjoy simple pleasures in life and to have fun along the way.

To my grandparents, Wallace and Inez Irvin and John and Leone Hill, who never seemed to be surprised by my accomplishments.

ACKNOWLEDGEMENTS

I would like to offer my most sincere thanks and appreciation to Dr. Louise MacKay, my chairperson, who has been a constant source of guidance, support, and encouragement. I also offer my gratitude to my committee members: Dr. Mary Langenbrunner, Dr. Jasmine Renner, and Dr. Terry Tollefson. I will always be grateful to have been taught by Dr. Russ West, whose gentle guidance made statistics a possibility for the “statistically challenged.”

I would like to thank Mr. Ed Abbott and Mrs. Helen Bennett, administrators at Abraham Lincoln Elementary School, for taking a risk in hiring me when I was young and inexperienced. Your trust allowed me to accept the challenge of teaching and turn that challenge into a life passion.

I am also grateful to Mrs. Cookie Greer and Dr. Lowell Biller, administrators at George Washington Elementary School, for staying true to the vision of a school that seeks ensure that there are no limits to a child’s learning.

I will be forever thankful to my dear friends and fellow teachers, Michele Arnold, Dena Porter, and Heather Wolf. I am blessed to work with teachers of such expertise and commitment! Working with you makes me a better teacher, and more importantly, a better person.

CONTENTS

	Page
ABSTRACT	2
DEDICATION	3
ACKNOWLEDGMENTS	4
LIST OF TABLES	8
Chapter	
1. INTRODUCTION TO THE STUDY.....	10
Purpose of the Study.....	12
Research Questions.....	13
Significance of the study.....	13
Limitations, Delimitations, and Assumptions.....	14
Definitions of Terms.....	15
Organization of the Study.....	16
2. REVIEW OF RELATED RESEARCH	17
History.....	17
Nongradedness Rationale.....	20
Features of Nongraded Programs	25
Philosophies that Undergird Nongraded Programs	30
Curriculum and Assessment.....	35
Supporting Research.....	37
Academic Achievement.....	38
Mental Health	39

Longitudinal Studies	40
Chapter Summary	41
3. RESEARCH METHODOLOGY	43
Research Design	43
Population	44
Sampling Method	44
Instrumentation.....	44
Data Collection.....	45
Data Analysis	46
Hypotheses	47
Chapter Summary	50
4. RESULTS	51
Demographic Information.....	51
Research Question #1	54
Research Question #2	57
Research Question #3	61
Research Question #4	64
Research Question #5	66
Chapter Summary	70
5. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS	72
Summary of the Study	72
Summary of the Findings.....	73
Research Question #1	74

Research Question #2	74
Research Question #3	75
Research Question #4	75
Research Question #5	76
Conclusions	76
Recommendations for Practice.....	78
Recommendations for Further Study.....	78
REFERENCES.....	80
APPENDIX.....	85
Letter of Permission.....	85
VITA.....	86

LIST OF TABLES

Table	Page
1. Demographic Profile of Ethnicity for 2005 Multi-age School.....	52
2. Demographic Profile of Ethnicity for 2005 Single-age School.....	53
3. Demographic Profile of the 2005 Fifth Grade Cohort.....	54
4. Differences in 2003 Reading/Language Raw Gains Between Multi-age and Single-age Students.....	55
5. Differences in 2004 Reading/Language Raw Gains Between Multi-age and Single-age Students.....	56
6. Differences in 2005 Reading/Language Raw Gains Between Multi-age and Single-age Students.....	57
7. Differences in 2003 Reading/Language Raw Gains Between Male Multi-age and Single-age Students.....	58
8. Differences in 2004 Reading/Language Raw Gains Between Male Multi-age and Single-age Students.....	59
9. Differences in 2005 Reading/Language Raw Gains Between Male Multi-age and Single-age Students.....	60
10. Differences in 2003 Reading/Language Raw Gains Between Female Multi-age and Single-age Students.....	61
11. Differences in 2004 Reading/Language Raw Gains Between Female Multi-age and Single-age Students.....	62
12. Differences in 2005 Reading/Language Raw Gains Between Female Multi-age and Single-age Students.....	63
13. Differences in 2003 Reading/Language Raw Gains Between Males and Females in a Multi-age Classroom.....	64
14. Differences in 2004 Reading/Language Raw Gains Between Males and Females in a Multi-age Classroom.....	65
15. Differences in 2005 Reading/Language Raw Gains Between Males and Females in a Multi-age Classroom.....	66

16. Differences in 2003 Reading/Language Raw Gains Between Males and Females in a Single-age Classroom.....	66
17. Differences in 2004 Reading/Language Raw Gains Between Males and Females in a Single-age Classroom.....	68
18. Differences in 2005 Reading/Language Raw Gains Between Males and Females in a Single-age Classroom.....	69

CHAPTER 1

INTRODUCTION TO THE STUDY

The *No Child Left Behind (NCLB) Act* (2003) has instituted a great deal of change in schools across America. Some believe it is a step in the right direction, with an increased emphasis on high stakes testing and accountability. Others have criticized NCLB and argue that its effects will lead to the end of America's public education. Regardless of one's personal opinions, the implications of NCLB cannot be ignored. Educators are faced with the challenge of striving to comply appropriately with its demands while at the same time staying true to what is best for students.

One of the components of NCLB is the issue of "closing the gap" that separates the achievement levels of children from low socioeconomic groups, English language learners, ethnic subgroups and special education students from their more privileged peers (U.S. Department of Education, 2004). The National Assessment of Educational Progress reported that the average eighth-grade minority student performs at about the level of the average white fourth grade student in the area of literacy (National Center for Education Statistics, 2003). NCLB emphasizes the need to reach all students. Public schools face high stakes testing to determine whether schools are meeting the needs of all children (U.S. Department of Education). Many schools are currently struggling to determine how to best achieve this goal of reaching every child in light of overcrowded schools and low funding.

In the midst of striving to meet the needs of all students, the multi-age structure is once again the subject of renewed interest (Gaustad, 1992b). Anderson stated "There are powerful forces for educational change in this country that are calling for structural as

well as instructional improvements that are wholly consistent with nongraded concepts and approaches” (Anderson & Pavan, 1993, p. 9).

A common theme for school reform is the need for continuous progress and changes in the practices of the retention and promotion of students. Goodlad and Anderson (1963) condemned America’s graded structure and argued there was little empirical research that supported the idea of segregating students by age and holding all students to the same standards. Pavan’s (1992) research concerning nongradedness, along with 64 other studies, reported that in most cases, students in nongraded classrooms had demonstrated higher academic achievement and mental health than their counterparts in age segregated classrooms. In addition, she found added benefits for males, African Americans, underachievers and students from low socioeconomic groups (Pavan, 1992).

Benjamin Bloom concluded in his “mastery learning theory” that 95 % of students in our schools were capable of significant success (Bloom, 1976). This statement should shake educators to the core. School leaders must take a serious look at restructuring our familiar, graded structures as a means of reaching all children. Although there is much research to support the cognitive benefits of multi-age programs, more research needs to be conducted to determine whether nongraded configurations increase academic achievement as measured on standardized tests (Bernheisel, 1992). In our newfound age of accountability and high stakes testing, these data will be necessary for restructuring to occur.

Purpose of the Study

The purpose of this study was to explore the differences between Reading/Language achievement gains of students in multi-age classrooms to the Reading/Language achievement gains of their peers in traditional, single-grade classrooms. Both schools are located in Northeast Tennessee. The raw Reading/Language gain scores of students enrolled in multi-age classrooms in a multi-age school was compared to the scores of students in four single-aged classes in a single-aged school. The scores reported for all students on the TCAP Standardized Assessment were examined as the primary dependent variable. The study focused on students from multi-age classrooms as they progressed through what is typically known as the third, fourth, and fifth grade and their counterparts at a single-grade school. An additional examination explored the performance levels of females in comparison to males among both groups.

A wide range of accepted terminology is present in research to describe and discuss the multi-age concept. Nongraded, mixed-age, and multi-age are just a few. In multi-age schools children are not labeled as being in grades k, 1, 2, 3, 4, and 5. Instead, they are housed in a primary or intermediate learning center. For the purpose of consistency within this dissertation, the term multi-age was used to describe the configuration of students of varying ages housed in one classroom.

For the purpose of comparison, the term “grade” was also used to identify children in multi-age classrooms who would have traditionally, based on age, been in the fifth grade.

Although these children participated in multi-age classrooms that focused not on “graded” norms but on each child’s ability, the term grade has also been used to categorize students so that associations can be made to America’s traditional school structure.

Research Questions

1. To what extent, if any, are there differences between the TCAP Reading/Language raw gain scores between children in multi-age classrooms and the scores of children in single-graded classrooms?
2. To what extent, if any, are there differences in the TCAP Reading/Language raw gain scores between males in multi-age classrooms and males in single-graded classrooms?
3. To what extent, if any, are there differences in the TCAP Reading/Language raw gain scores between females in multi-age classrooms and females in single-graded classrooms?
4. To what extent, if any, are there differences in the TCAP Reading/Language raw gain scores between males and females in multi-age classrooms?
5. To what extent, if any, are there differences in the TCAP Reading/Language raw gain scores between males and females in single-graded classrooms?

Significance of the Study

This study was significant by providing useful information to school leaders concerning the relationship between multi-aged and single-graded groupings on Reading/Language gains. This study has the potential for providing quantitative

information that could be used by the educational community in evaluating one dimension of the effectiveness of multi-age groupings. Although there has been renewed interest in continuous progress education in recent years, there is still a need for more empirical data regarding the multi-age concept. Teachers and administrators could benefit from the comparisons made in this study and as a result make better decisions regarding the delivery of instruction and the structuring of school classrooms.

Limitations, Delimitations, and Assumptions

The research in this study was confined to third, fourth and fifth grade students who participated in multi-age classrooms at a multi-age school and their counterparts in traditional classrooms who attended a single-graded elementary school located in Northeast Tennessee during the 2004-2005 school year.

It was assumed in this study that TCAP scores reported for all students were accurate and indicative of student achievement and gains. The researcher also assumed that the TCAP was administered in a setting that was conducive to optimum performance by all students. Environmental factors such as lighting and room temperature were assumed to be satisfactory. It was assumed that distractions were kept to a minimum throughout the testing procedure. It was assumed that all teachers participating in the study (multi-age and traditional) were guided in their instruction by a framework of instructional objectives issued by the State Department of Education.

It was assumed that all teachers in the study were capable, competent, and comparable in skill and ability. It was also assumed that both schools were safe and comfortable and provided equitable opportunities for academic success.

Definitions of Terms

For the purposes of this study the following definitions will apply:

1. Continuous Progress. The instructional approach that recognizes the child as the baseline for where instruction begins. There are no ceilings on learning, and each child moves at his or her own pace, regardless of the age of the child (Hillson & Bongo, 1971).
2. Developmentally Appropriate Practice. Guidelines that direct teachers in how to create a child friendly learning environment focusing on factors that are age appropriate and individually appropriate for each child (Bredekamp, 1987).
3. Fifth Grade Cohort. A group of fifth graders who traveled together as a group through what are traditionally known as the third, fourth, and fifth grades.
4. Flexible Grouping. Combining students of similar ability for specific purposes with an understanding that the when objectives are met the group configurations will be reassembled (Gaustad, 1992b).
5. Heterogeneous grouping. A group of children consisting of varying abilities and interests (Goodlad & Anderson, 1963).
6. Homogeneous grouping. A group formation where selection is based on general achievement in a specific area (Goodlad & Anderson, 1963).
7. Individualized Instruction. The practice of modifying instruction for each child based on each individual's goals, learning styles and temperament (Musgrave, 1922).
8. Multi-age grouping. Classrooms where there are a group of children with mixed ages and abilities all sharing in learning. Instruction is individualized. (Katz, Evangelou, & Hartman, 1990).
9. Nongraded Education. A vertical facet of school organization wherein instructional groups are organized on an integrated basis in one or more curriculum areas depending on the developmental needs of the children (Otto, 1969).
10. Reading/Language gain score. The growth observed over one year's time in the area of Reading/Language based on the TCAP achievement test.
11. Retention. The practice of keeping a child in the same grade for more than one year due to perceived developmental or academic deficiencies.

12. Single-aged classrooms. A configuration of children whose birthdays lie within the same twelve month range; with the exception of children who have experienced promotion or retention (Tomlinson, 1999).

Organization of the Study

This study is composed of five chapters. Chapter 1 is the introductory chapter containing the purpose of the study, the research questions, the significance of the study, limitations, delimitations, assumptions, and definition of terms. Chapter 2 contains the review of related literature. Chapter 3 contains a description of the study, the population, the sampling method, the instrumentation, the procedures, the methods of data collection, and the methods of data analysis. Chapter 4 will contain a description of the data obtained, discusses how the data were prepared for analysis, and presents the analysis of data. Chapter 5 will present the summary, conclusions, implications, and recommendations for practice and for further study based on the analysis of data.

CHAPTER 2

REVIEW OF RELATED LITERATURE

The idea of grouping children into grades according to age has been controversial from its inception. McLoughlin (1967) argued that graded schools had always shared one common flaw, “They attempted to fit the child to the curriculum and never the curriculum to the child” (p.1). This chapter provides a review of nongradedness as a means of reaching all students and as an attempt to fit the curriculum to the child. It will focus on the history, definition, features, philosophies, and advantages of multi-age schools and provide a summary of previous research findings.

History

Although little work has been done to trace the history of nongradedness, its presence dates back to the dame schools of the 1600s. Dame schools served children from ages 3 to 10 years. These schools functioned without grade classifications. The specific needs of students were met by individual and small group instruction. Other examples of multi-age classrooms included reading and writing schools of the 1700s and the Lancastrian schools of the early 1800s (Otto, 1969).

Most often recognized for its multi-age feature was the one-room schoolhouse. The schoolhouse functioned as a multi-age classroom out of necessity, but its advantages have not gone unnoticed (Lodish, 1992). Learning was individualized. Teachers worked with children one-on-one as well as with small groups for part of the day. Children also worked cooperatively in mixed aged groups (Goodlad & Anderson, 1963).

Although many benefits arose from the one-room, multi-age schoolhouse, some difficulties emerged as well. The teachers of the one-room schoolhouse were often not adequately prepared to teach children of such a wide ranging age span. In order to ease the hardship, and create more structure, leveled textbooks, such as *The McGuffey Eclectic Readers*, were introduced. These textbooks were adopted by many schools as a step toward a more “progressive” and “orderly” way of teaching (Anderson & Pavan, 1993).

In Boston in the late 1800s males and females were separated for instruction. Learning began to be broken down into separate subjects, and grade “norms” began to be established. School leaders sought to make the job of the teacher more manageable. As more and more children began attending school as a result of mass migrations, the multi-age nature of many schools was seen as inefficient, and a factory model began to be imposed on schools (Goodlad & Anderson, 1963). People sought to teach children in an assembly line manner with identical treatment given to each child (Katz, 1995). Gone was the individualization of instruction of the one-room school. In its place would soon be a standardization of instruction for all. Goodlad and Anderson (1963) stated that, “in such a setting, the ideas of educational spokesmen like Mann, Stowe, and Barnard, while controversial, struck many receptive ears” (p.2).

Horace Mann is credited with bringing the graded school structure to America in the 1800s. Mann was the Secretary of the Massachusetts Board of Education at that time. Following a visit to Prussia, Mann came back to the U.S. discussing the superior structure of the German schools. He said that in Germany, teachers were responsible for only one class. He praised the genius of dividing students into classes segregated by age. Within a decade, his ideas had gained a widespread audience. Within 15 to 20 years following

Mann's proposal for a graded structure, schools had developed into the image of school that many have today (Goodlad & Anderson, 1963).

This factory model promised to mass produce knowledgeable students in the most efficient manner (Pratt, 1986). The most recognized example of Mann's graded schools was the Quincy Grammar School of Boston, founded in 1848. Many elementary schools in later years were patterned after this one. It was housed in a four-story building divided into individual classrooms placing pupils of like age in the same class (Goodlad & Anderson, 1963). Each group of students was taught by one teacher, and at the end of the school year the students either passed to the next grade or were retained for another year (Anderson & Pavan, 1993).

By the 1870s, the graded structure began receiving criticism. Critics pointed out that a graded structure was too rigid. W.T. Harris brought forth a plan that suggested educators view grades as flexible groupings. He recommended that teachers stop after a six-week period and reshuffle students who varied from the group norm. Harris suggested frequent promotion and reclassification of students.

John Dewey, who became known as the father of Progressivism, also challenged the graded structure. At his laboratory school he eliminated the graded structure, subject matter, and the use of textbooks. Dewey argued learning occurred from a child's experiences in and out of the classroom (Goodlad & Anderson, 1963).

Progressive ideals began to spread in the mid 1900s. Efforts such as the Individually Guided Education and Continuous Progress helped to spread models of team teaching, hands on learning, and multi-age classrooms into many school districts. The Russian launch of the Sputnik satellite in 1957 caused a knee-jerk reaction in the minds

of politicians and editorialists. Fear created rampant propaganda that the U.S. had fallen far behind the Russians in the areas of math and science (Pavan, 1992). Within five years after Sputnik, the entire second grade curriculum had been moved into first grade. Children who were in the first grade were now expected to learn what had previously been taught an entire year later. Records showed that many children thus failed to thrive with the higher and inappropriate expectations (Connell, 1987). The graded school, however, survived the criticism and remains the dominant school structure in America today (Goodlad & Anderson, 1963).

In spite of the fact that the graded school structure is still the most prevalent, teachers, parents, and politicians are again beginning to look back reflectively to the one-room school (Uphoff & Evans, 1992). An increase in the understanding of a developmentally appropriate education has caused many educators to once again explore the idea of a multi-age structure in an attempt to meet the diverse needs and abilities of students.

Nongradedness Rationale

The rationale behind multi-age classrooms is the belief that a child's chronological age does not dictate his or her mental capabilities (Pavan, 1992). Multi-age schools do not use grade level classifications to label a student's ability to achieve (Pavan, 1992). True multi-age groupings teach children of various ages and abilities together without dividing them into separate groups or grades (Gaustad, 1992b). This is in contrast to traditional classrooms that consist of children within a twelve-month age range (Katz, 1995).

The cognitive growth of children varies just as physical growth varies among children. In multi-age classrooms, these differences are understood to be the norm. This truth is often hard to manage in single-aged classrooms (Milburn, 1981). Katz (1995) argued that, “When children in a class are close in age, teachers and parents tend to expect them to be ready to learn the same things at the same time” (p.4).

Age-segregated classrooms are difficult for children who perform above or below the grade norm. Teachers cite that the most difficult task they face involves adjusting instruction to the wide range of needs within a class. For the exceptional performers, “skipping” a grade is sometimes allowed in order to meet the academic needs of students. For the child whose development is not quite up to the norm, retention is a possibility. Skipping a grade and retaining children are both pitfalls of our graded structure and cannot be substitutes for genuine differentiated instruction (Pratt, 1986).

Goodlad and Anderson’s research concerning students who were promoted versus those who were not reveal that promoted slower learning children achieve at higher levels and are less likely to engage in misbehavior than their peers who were retained. Furthermore, those promoted students are also shown to have more positive feelings of self-worth whereas their retained peers often struggle with feelings of fear and low self-esteem (Goodlad & Anderson, 1963).

Multi-age classrooms teach children according to their developmental levels without special arrangements for remediation or acceleration. Children are not aware they are above or below a set norm because all children are taught in an individualized fashion (Milburn, 1981).

Each child has a unique pattern of growth (Grant, Johnson & Richardson, 1995). In multi-age classrooms, a child's growth is reported in behaviors observed and tasks completed rather than by letter or numerical grades (Pavan, 1992). Children make continuous progress at their own rate from easier to more difficult material. There are no specific time factors designated for mastery of specific skills. Development and progress are seen on a continuum (Gaustad, 1992b). Tomlinson (1999) shared, "Differentiated instruction begins where students are, not from the front of a curriculum guide"(p.2).

Grant et al. (1995) stated that the most important element in a multi-age classroom was the manner in which teachers accepted children where they were functioning academically as they enter school. They applauded the practice of beginning instruction at the level of student performance rather than at a designated starting point designed for the average student. "It is the responsibility of educators to ensure that schools are ready for children instead of making children ready for school" (p.2).

McLoughlin (1967) hailed flexibility as the hallmark of a multi-age school. He argued that graded structures were too rigid for the good of students. Multi-age structures allow for flexible groupings of students with similar needs. Organizing students in flexible learning groups allows children to make continuous progress at their own rate without being locked into one specific group. In many graded configurations children are expected to meet specific graded norms and objectives. If children do not meet these standards they are viewed as deficient. A multi-age approach is not defined by preconceived ideas of what a child should know. A continuous progress approach means to celebrate what a child does know and focuses on what a child is ready to learn rather than on what they do not know. Cushman (1990) stated, "Multi-age classrooms

expect diversity among students, and adapt to the diverse abilities of the children” (p. 2). Movement among learning groups is fluid (McLoughlin). Multi-age groupings provide children the opportunity to work at their own level without any obvious remediation or acceleration (Milburn, 1981).

Anderson and Pavan (1993) have established 11 statements that have gained widespread acceptance as a way to define of the multi-age approach:

- 1) Individual differences in the pupil population are accepted and respected, and there is ample variability in instructional approaches to respond to varying needs.
- 2) Learning, which is the “work” of the child, is intended to be not only challenging but also pleasurable and rewarding.
- 3) Students are viewed as a whole; development in cognitive, physical, aesthetic, social and emotional spheres are nurtured.
- 4) The administrative and organizational framework, for example with respect to pupil grouping practices, is flexible and provides opportunities for each child to interact with children, and adults, of varying personalities, backgrounds, abilities, interests and ages.
- 5) Students are enabled through flexible arrangements to progress at their own best pace and in appropriately varied ways. Instruction, learning opportunities, and movement within the curriculum are individualized to correspond with individual needs, interests and abilities.
- 6) Curricular areas are both integrated and separate. Instruction, programmatic, and organizational patterns are flexible, with outcomes rather than mere coverage of content as the primary focus.
- 7) The expected standards of performance (in terms of outcomes) in the core areas of the curriculum are clearly defined, so that the points to be reached by the end of a designated (e.g., a three or four year) period are well known. However, the time taken to reach that end, and the path followed to that end, is allowed to vary for students with different histories and potentialities.
- 8) Within the curriculum and related assessment practices, specific content learning is generally subordinate to the understanding of major concepts and methods of inquiry, and the development of the skills of learning: inquiry, evaluation, interpretation, and application.

- 9) Student assessment is holistic, to correspond with the holistic view of learning.
- 10) Evaluation of the learner is continuous, comprehensive and diagnostic. Except for reference purposes as necessary to parental and staff understanding, chronological age and grade norms play a much smaller role in evaluation and reporting activities than does the child's own growth history and potential.
- 11) While there are some core components of the curriculum that are especially valued (as reflected in performance standards in the major content areas), the system is largely teacher managed and controlled. Thus, it empowers teachers to create learning opportunities and to use instructional strategies at their own discretion, based on the perceived needs of the students they are serving. Assessment procedures are similarly flexible, individualized and teacher managed (p.62-63).

Many people have gross misconceptions of what happens in multi-age programs. Some wrongly believe that multi-age programs are less structured than their graded counterparts (Lodish, 1992). Gaustad (1992b) has responded, "Nothing could be further from the truth" (p. 91). Those who are not familiar with developmentally appropriate practice might assume that classrooms housing children sitting quietly in rows are more structured than classrooms where children are active and cooperative. In truth, it takes much more structure to organize a classroom where learning is individualized and where children work independently (Lodish). In multi-age classrooms routines and procedures are practiced, and a high expectation for learning is the expectation. Movement and cooperation within a classroom do not mean structure is absent.

A second misconception is that multi-age programs are designed to group children by ability. Multi-age programs do not track students. In fact, multi-age programs are based on views that are adamantly opposed to such practices (Lodish, 1992). Multi-age configurations allow children of differing ages, socioeconomic status, interests and abilities to work and learn together (Gaustad, 1992b). Children learn from

others who are working at higher cognitive levels. The higher functioning students benefit by verbalizing what they know and organizing their thoughts in a manner necessary to teach their peers, thus strengthening their own skills. Cooperative learning has both academic and social advantages for high and low performing students. There is a benefit to this type of heterogeneous grouping (Grant et al., 1995).

Slavin (1992) argued that the curriculum and instructional methods used within a multi-age program were as important as the organization itself. To investigate the effectiveness of multi-age programs, the characteristics and features of nongradedness need to be understood.

Features of Nongraded Programs

Teachers can differentiate instruction within any structural configuration of students; however, the multi-age classroom offers added flexibility for grouping students who are working at similar levels without having to affect the promotion or retention of children within a specific grade (Hillson & Bongo, 1971). Goodlad and Anderson (1963) recommended that students belong to a combination of children (ideally 70-120 students) that consists of at least two age groups. Furthermore, they stated that configurations of three age groups may be even more preferable than two. Evanshen (2001) conducted a more recent study on the effectiveness of two and three year configurations. Her research supported that of Pavan and Anderson, which suggested that a multi-age, three-grade span configuration may be preferable (Evanshen).

Cycling is associated with non-graded programs. Simply stated, cycling means that teachers stay with approximately the same group of children for more than one year (Goodlad & Anderson, 1963). The cycling concept is similar to looping. Looping

teachers teach a group of students in a particular grade for the first year. The following year, the teacher and the children go to the next grade together. At the end of the looping cycle, the teacher goes back to the original grade and begins again with a new group of students. A looping cycle typically ranges from two to three years (Grant et al., 1995).

In multi-age classrooms, teachers of two or three grade configurations stay with the children as they pass through the grades taught by that teacher. This allows for fewer transitions for teachers and students. It also helps create a strong sense of community within the class. Multi-year placements provide increased student observation time for teachers. Cycling provides children with the gift of time to grow (Grant et al., 1995).

Team teaching is often associated with multi-age classrooms (Gaustad, 1992b). It involves the collaboration of teachers in planning, instructing, and evaluating students (Lewis, 1969). Team teaching requires teachers to work closely and share the responsibilities of teaching students assigned to each member of the team (Anderson & Pavan, 1993). Team teaching is defined by Anderson and Pavan as having the following features (p.95):

- 1) Long range as well as short range curriculum planning, including occasional review of each teacher's specific lesson plans, is a total group responsibility.
- 2) The team members, although much of their teaching is within a private or "solo" context, regularly work together (co-teach).
- 3) Assessment of the overall instructional program, as well as assessment of components with which individual teachers are concerned, is a team wide responsibility.
- 4) The pupils, although connected for advisory and other purposes to one another of the teachers, belong to and are regularly connected with all of the teaching members.
- 5) Assessment of each child's needs and progress is an activity in which all team members participate in some systematic manner.

- 6) Resources are shared (co-owned) for the benefit of all teachers and students.
- 7) Team teaching benefits students by exposing them to a diversity of teaching styles and strategies.

The curriculum is made richer by combining the knowledge and ideas of a team. In addition, teachers can share job responsibilities based on their strengths or interests (Grant et al., 1995). Cohen (1981) found that teachers who worked as teams tended to grow together and be more effective than they would have been if they had worked alone. When teachers agree to “team,” the children assigned to each of the teachers in the team are considered to be one single group (Goodlad & Anderson, 1963). Combining children for specific learning can be more effective because the pool of students is larger than the number of students in a single class. When grouping children from several classes, it is more likely that the needs of the students in each learning group will more closely align. Instruction can specifically target the needs of members in each group. It is important to note that learning groups are based on need as well as interest. Student groups are flexible and change often (Gaustad, 1992b). In addition to homogeneous groups, planned heterogeneous groupings are also worthwhile, for children to work together and learn from each other (Goodlad & Anderson).

In many cases teachers pool not only their students, but their space and resources as well. This is a sign of a healthy team (Anderson & Pavan, 1993). Teachers work together in what is known as a learning center. Some schools are created for this purpose. Learning centers are constructed instead of classrooms. These centers are large and open allowing space for as many as four individual classes to work comfortably. In other schools, originally created for single-aged classrooms, multi-age centers can be

created by cutting doors or openings through the walls of adjacent classrooms (Goodlad & Anderson, 1963). Learning centers are arranged so that children can move freely from one room to the next as they regroup for various learning experiences. Multi-age learning centers can function much like the one-room school house. Children of varying ages work together in an active and participatory manner (Shanker, 1993).

Hands-on activities and cooperative learning concepts help to guide the physical organization of a multi-age learning center. Rows of desks are not arranged in one direction facing a chalkboard or teacher desk. Children sit together at tables. If tables are not available, then desks are grouped together. Children are expected to work alongside and with their peers (Gaustad, 1992b). Students do not have a permanent seat where they store all of their textbooks. Instead, students usually have a cubby or drawer that contains folders and personal items to use as supplies or resources during their learning (Nachbar, 1989). Learning centers are designed for children to learn cooperatively through centers. Centers consist of tables, bookshelves, and carts that contain various materials to be used throughout the learning process.

The appearance of a learning center is drastically different from a traditional classroom. However, what happens within the learning center is of much more impact than the physical appearance of the room. Direct instruction is kept to a minimum with a focus on the learner (Lewis, 1969). Center work provides an inquiry based approach whereby students arrive upon concepts by their direct involvement with materials. Textbooks may be present but are used to supplement the learning as a resource rather than as the main instructional tool (Gaustad, 1992b).

Many learning experiences happen within the context of a small group. Small-group work is employed in many graded classrooms. The difference in the groupings within a multi-age program is the flexibility of the group. In multi-age classrooms, children are frequently rearranged according to different criteria such as ability, interest, or learning style (Gaustad, 1992b). Gaustad (1992b) stated that, “Even the greatest of supporters of mixed-age and mixed ability grouping agree some curricula are most effectively taught to children of similar experience and achievement” (p.97). Goodlad and Anderson (1963) reported that reading and math were often areas where children benefited from homogeneous grouping. This type of combination would be based on achievement rather than age. The purpose of grouping children by ability is based on the argument that children’s needs are better served when teachers work with children whose needs are similar (Goodlad & Anderson).

There are also heterogeneous groupings of students in multi-age programs. Goodlad and Anderson (1963) also described circumstances where children of differing abilities worked well in learning groups. They argued that in thematic subjects such as science and social studies students benefit from working together heterogeneously. Cooperative groups allow students to work together performing different tasks according to their ability (Gaustad, 1992b). The learning experiences that occur in multi-age programs are heavily supported by deep philosophical roots that provide a setting for developmentally appropriate practice and continuous progress (Gaustad, 1992b).

Philosophies that Undergird Nongraded Programs

Multi-age classrooms and the term nongradedness are still unfamiliar to many Americans. Most Americans attended what many consider to be a traditional age-segregated schools. In fact, the idea of a graded school system is so familiar that it is difficult for many to consider any other type of configuration. However, our graded ideals are not universal (Pratt, 1986). Many educators around the world cannot imagine our system of gradedness (Connell, 1987).

Throughout New Zealand, schools permit children to enter school on their fifth birthday at any time during that school year. When children enter, they are known as early entrants. At their point of entry, the students learn and progress at their own pace. Children are promoted to the next level of learning by achievement rather than by their chronological age. There are no mass promotions consisting of children who all share a birthday within the same 12 months. New Zealanders do not know the meaning of retention. Nor do they conceive of “skipping” a grade. Indeed, educators in New Zealand have contended that our graded structure of educating children seems unfair, because they do not expect children to perform academically at the same pace or time (Connell, 1987).

Likewise, British primary schools assemble children according to family groupings. In these family groupings, children of different ages are combined into one classroom where teachers have individual expectations for each child. In these classrooms children stay with the same teacher for three years rather than one school year or 10 month period. Like the New Zealand structure, promotion is by achievement rather

than by chronological age. Children learn in cooperative groups and often help each other (Connell, 1987).

Today in America, most teachers still work with single-graded classrooms made up of children approximately the same age. These classrooms are labeled from the beginning by grade (Tomlinson, 1999). It is difficult for many who graduated from age-segregated school systems to imagine what school would be like without grade labels. The idea of a multi-age school structure may strike some as an inefficient and unfounded model. However, in light of today's charge to "close the gap" and "leave no child behind," the idea of teaching each child rather than teaching subject matter is the very idea with which we must come to terms. Even though many classrooms today are made up of children who are almost the same age, the needs of the children within the classrooms are nearly as diverse as the needs once held within the walls of the one-room school (Tomlinson). Proponents of multi-age classrooms share the belief that children of the same age often have diverse abilities. Children do not learn at the same rate at the same time. Chronological age does not equal mental age (Milburn, 1981).

Bredenkamp and Copple (1997), along with the National Association for the Education of Young Children (NAEYC), authored a definition of what is known as "developmentally appropriate practice" (DAP) for primary-aged children. DAP is closely aligned with what takes place in multi-age classrooms (Gaustad, 1992b). Bredenkamp and Copple argued that children should be taught according to what was age appropriate and by what was individually appropriate. In helping to define what is individually appropriate, the NAEYC (1997) stated, "Each child is a unique person with an individual pattern and timing of growth, as well as individual personality, learning

style, and family background (p.2). DAP encouraged teachers to offer children adequate time and experience for learning (Bredenkamp & Copple). Many classrooms today offer little assistance to students who are advanced or behind the lockstep, systematic approach to learning of the graded structure.

Multi-age programs function with an understanding that a child may excel in one area and have difficulty in another. Young children do not automatically mature at the same time in every domain of their development (Katz, 1995). Howard Gardner encouraged schools to see children's full development by seeking to discover not whether children are smart but how they are smart (Gardner, 1993). Gardner challenged that one's Intelligence Quotient (IQ) is not the only measure that can predict potential success (Blythe & Gardner, 1990). In his book *Intelligence Reframed*, Gardner reported that humans used up to 10 distinctive modes of thinking: linguistic, musical, logical mathematical, spatial, bodily-kinesthetic, interpersonal, intrapersonal, naturalist, existentialist, and spiritual (Gardner, 1999). He described these modes of thinking as multiple intelligences (MI). Traditional classrooms often rely heavily on the linguistic and logical-mathematical modes. Multi-age programs tap into children's multiple intelligences through constructivist and collaborative learning experiences. By allowing children the opportunity to use their personal intelligences, there is an increased likelihood of success in the classroom (Blythe & Gardner).

Schwebel and Raph (1973) stated:

If we are to improve our schools, we need to bring them more in harmony with the processes of development. This goal applies as much to the organization and climate of the school as a whole as to the social and intellectual character of classroom life (p.35).

Piaget (1896-1980), a Swiss psychologist, developed a well-respected theory of cognitive development. His focus was on the development of knowledge not the acquisition of skills or information. Piaget argued that memorizing facts alone was not true learning. Piaget stated that children should have direct interactions with the environment for learning to truly take place (Piaget, 1973). His theory of cognitive development described four stages: sensory-motor: approximately the first 24 months of life; preoperational: approximately 2 to 7 years; concrete operational: approximately 7 to 11 years; formal or abstract: adolescence and adulthood. Piaget suggested that all people progressed through those stages. However, it is not expected that all people pass through these stages at the same times in their lives (Piaget).

Multi-age programs naturally recognize and plan for diversity in abilities and rates of progress and also adjust to meet different emotional and social needs of students (Lodish, 1992). Traditional classrooms consist of students within a twelve- month age range. Often, such classrooms create tension for students because the expectation is that all students possess the same knowledge and skills. There is no evidence to support such a claim (Katz, 1995). Gaustad agreed as he outlined several drawbacks to the age stratified classroom. First, he noted that children of the same chronological age often vary in their cognitive development. Second, he acknowledged that children also differ in regard to their personal learning styles. Finally, he stated that teachers in traditional classrooms compare children with each other and consider children who do not meet the “norm” as deficient (Gaustad, 1992a). Multi-age classrooms provide a nurturing environment where children are not measured against one another but by their own personal successes (Cushman, 1990).

Bredenkamp and Copple (1997) included the importance of a child's social and cultural context into their vision of a developmentally appropriate classroom. Social interactions in multi-age classrooms positively affect children's development in all domains. Vygotsky's social-cognitive theory suggests that a child's development is enhanced by working with those of higher cognitive levels. He asserted that people learn more by working together than by working alone (Vygotsky, 1978). Vygotsky contended that more capable children ("experts") provide prompts to others ("novices") that lead them to more advanced solutions to problems. The collaboration that occurs in a multi-age classroom leads children to higher levels of thinking through the social interaction that takes place (Stone, 1994). Lodish (1992) describes the social advantages of the multi-age classroom:

1. A unidimensional classroom defines academic ability and work narrowly and uses a restricted range of performance criteria to evaluate children. In these classes, the assigned tasks tap only a limited range of children's abilities and interests. On the other hand, multidimensional classes, whether single-age or mixed, offer a comparatively wide range of activities in which varying levels of skills can be applied.
2. Although some parents express concern about the likelihood of competition in a multi-age group, research indicates the opposite – that greater cooperation is often the result. Because such grouping appears to minimize competitive pressure, discipline problems that seem inherent in competitive environments are often substantially reduced.
3. Since most young children are not equally mature in areas of development at a given time, mixed-age grouping can be an effective strategy for dealing with their different rates of development. This grouping can be very helpful for children functioning below age-group norms in some developmental areas.
4. As a child interacts with children at different levels of cognitive maturity, intellectual growth is stimulated. Some proponents of mixed-age classes argue that the cognitive conflict likely to arise in mixed-age interaction provides situations for significant learning for younger children as they strive to accommodate to the more advanced understanding of their classmates (p. 5).

Pratt (1986) reported that, while age was a factor in friendship formation, children were also likely to choose friends who are of a similar level in terms of development as well. The socialization that occurs within a multi-age setting is important for friendships and is also vital to the learning that occurs.

Curriculum and Assessment

“Successful multi-age classrooms require teachers to shift attention from teaching curriculum to teaching children” (Stone, 1994, p.4). Continuous progress must be the vehicle for instruction. In multi-age classrooms student needs are evaluated and seen on a continuum. There is no whole group instruction based on specific graded “norms.” The developmental needs of children drive instruction rather than skills of a prescribed curriculum (Stone). Hunter (1992) defined continuous progress by stating:

With continuous progress, students are challenged appropriately according to their ability to master intellectual, physical, emotional, and social tasks at progressively more difficult levels. Continuous progress mandates that students should neither spend time on what they have already adequately achieved, nor proceed to more difficult tasks if they have not yet learned materials or acquired skills essential to that new level of knowledge (p. 5).

In many classrooms, especially those of a single- graded nature, children learn according to a “time bound structure.” A time-bound structure refers to instruction that is paced and dictated by the calendar. Instruction is timed in a rigid manner without regard to the developmental levels of the children in the class. Efficiency is the goal rather than maximum learning and progress. Again, this concept is based on the false assumption that children of the same age learn at the same rate (Grant et al., 1995).

Goodlad and Anderson (1963) suggested that continuity and sequencing were fundamental to a well- planned curriculum. They proceeded to explain that learning must be meaningful to students. New learning must be linked to students' prior knowledge so that a foundation for future learning can be laid (Goodlad & Anderson). Recent brain research has reported that new learning takes place faster when the learner has had background knowledge or experience with the material being covered (Jensen, 2000). In addition, Goodlad and Anderson advocated that the well planned curriculum helps learners make connections between learning experiences.

Based on this research, many multi-age programs have implemented an integrated, thematic approach. Kavolik (2002) argued that children's education should be experience-based in order to give meaning to information. Kavolik implored, "Students can only dream to the limits of their awareness" (Kavolik, p. 1.14). Raley (2001) reported, "Experiences, thoughts, actions, and emotions actually change the structure of our brains" (p.18). Learning is more meaningful when students are involved in the process.

Assessments in multi-age programs are designed to measure student performance with methods that are congruent with the means of instructional delivery. Constructivist assessments are formative and ongoing. The assessments are used to guide future instruction. Student knowledge is measured as children are in the process of the learning experience (Wiggins, 1993). Examples of such assessments include anecdotal notes, student work, rubrics, and checklists (Clay, 1993). Anderson and Pavan (1993) advocated, "Assessment should be conducted in order to better understand the needs of the learner and to determine the direction of instruction" (p.164).

Supporting Research

Pavan is one of the most well known and respected names in non-graded, multi-age research. Her review of nongraded studies between 1968 and 1990 was conducted in attempt to determine what instructional configurations yielded the highest results for academic achievement and mental health.

Pavan gathered research for review through Research in Education, Current Indexes to Journals in Education, dissertation Abstracts, and Educational Index to Periodicals catalogs. Pavan used nine descriptors to search for nongraded studies: nongraded, nongradedness, nongrading, continuous progress, multiunit, individually guided education, multi-age, ungraded, and mixed age. Graded and nongraded programs with similar populations were compared.

In order to determine which research studies would be included in Pavan's review of research, the following criteria were used (Pavan, 1992):

1. Students in nongraded schools must be compared to those in graded schools or pre/post test of the same students, with the pre-testing conducted before entering a nongraded program.
2. Students must have been in a nongraded program for at least one academic year.
3. The nongraded label is assumed to be accurate unless either the researcher or a reading of the study very clearly indicates that the structure was not actually in operation.
4. There must be more than one nongraded classroom in operation. In cases where the sample size appears small, the study is reviewed to ascertain if matched pairs or a random sample had been obtained from a larger population
5. The entire school program must be nongraded, not just one subject area. Data

are acceptable from only one subject area, however, and multiple subject data are preferable.

6. Standard measures of academic achievement, mental health, and attitude are required.
7. Only studies conducted in the United States and Canada are included. students were in nongraded programs in their elementary school years commonly called grades one to six.
8. Some evidence was needed as to initial comparability of nongraded and graded schools in the study.
9. Studies should have been published since 1968.

Sixty-four research studies published between 1968 and 1990 were used to compare students in graded and nongraded classrooms. In addition to nongradedness, 13 studies were individually guided education programs, 11 include open space classrooms, and 11 refer to teacher teaming. As the studies of focus were narrowed, the review centered on academic achievement and mental health indicators (Pavan, 1992).

Academic Achievement

Fifty-seven studies out of the 64 reviewed by Pavan used standardized testing as a means of comparison. Many of the studies reviewed used more than one achievement test. In addition students of varying age groups were included in the study. Furthermore, testing was carried out over a long time span in some cases. As a result, comparisons of experimental groups (nongraded) and control groups (graded) were designed.

Out of 57 schools that used standardized testing as a means of comparison for achievement, 52 (91%) reported that nongraded groupings scored higher (58%) or as well as (33%) for all comparisons. Only 9 out of the 94 comparisons favored the graded

school classrooms. Overwhelmingly, students participating in nongraded programs performed as well and sometimes better than students in graded programs. Pavan suggested that the reason of the success of nongraded programs is that they responded to individual differences in students (Pavan, 1992).

Mental Health

In the area of mental health, 42 studies were performed out of the 64 being reviewed. Many studies used more than one instrument for the study. Therefore, as with the academic review, 81 comparisons of experimental (nongraded) and control groups (graded) were reported. Forty comparisons favored nongraded classrooms. Thirty-one indicated that both groups showed similar results. Ten studies found that students from nongraded programs were not as adequate (Pavan, 1992).

The Coopersmith Self-Esteem Inventory (Coopersmith, 1981) was used to assess self-esteem. Scores revealed that in all but one study, students in nongraded programs had higher self-esteem than students in graded structures (Pavan, 1992). In the one study that did not show higher results for the nongraded group, no significant differences were found.

Regarding mental health and school attitudes, 52% of the studies revealed that nongraded schools were better for students. Forty-three percent found that nongraded and graded schools had similar results. Five percent found that nongraded schools scored lower than graded schools.

The results of the studies indicated that students in nongraded schools were more prone to have positive self concepts, high self-esteem, and good attitudes toward school than their graded counterparts (Pavan, 1992).

Longitudinal Studies

Most of the studies included in Pavan's research reported data for just one year. However, 17 studies imparted data over a number of years. Longitudinal data from those schools revealed that students completing nongraded primary programs had higher academic achievement, lower retention rates, and more positive student attitudes than students from graded programs. Furthermore, the longer a student participates in a nongraded program, the better the results as with the studies by Carter, 1974; Eells, 1970; Killough, 1971; Morris, Proger, and Morrell, 1971; Pavan, 1977, 1992; Perrin, 1969; Ramayya, 1972; and Walker, 1973.

Five studies followed students who had spent 3 or 4 years in a primary multi-age classroom in contrast to the single-aged classrooms consisting of grades k through 3.. McLoughlin's (1967) study computed data for percentage of yearly deceleration and showed that 5 to 10 % more children enter fourth grade after three years of school (kindergarten not included) than students in graded programs. This means that fewer students are retained in nongraded programs.

Walker (1973) found similar results as students were followed over a 12-year period. Estimates showed that slower students from a nongraded background would be within one year of grade placement of their normal classmates upon graduation. It was found that the achievement gap begins to decrease at about the fifth grade level.

Eells (1970) reported that the longer students remained in multi-age/nongraded programs, the higher their achievement scores were in relation to their ability. Likewise, Perrin (1969) reported that students increase in achievement after three years in a nongraded program. Evanshen (2001) also found that student achievement is even better after three years in a multi-age setting than two years. Morris et al. (1971) reported that students in two and three year multi-age programs perform better academically than their peers in graded programs.

There were 58 comparisons made and two of those favored graded schools over nongraded schools (three percent). Sixteen reported similar results between the two types of schools (twenty-eight percent) compared to forty of the schools (sixty-nine percent) that favored the nongraded schools (Pavan, 1992).

In addition, many students in non-graded programs said they felt more positive or the same toward school as students from graded programs. Furthermore, students from nongraded programs had significantly fewer discipline problems as reported referrals to the office than students in graded programs (Pavan, 1992). Finally, the results revealed the results are more favorable for students the longer they participate in the nongraded program (Pavan, 1992).

Chapter Summary

This chapter has presented a literature review focusing on nongraded educational programs. A description of nongradedness was provided along with a brief history of its evolution. Features and philosophies of nongraded programs were described. Common

approaches to curriculum were explained. Previous research findings were also reported revealing the benefits for multi-age programs.

CHAPTER 3

RESEARCH METHODOLOGY

The purpose of this causal comparative study was to explore the associations between multi-age classrooms and single-age classrooms and TCAP Reading/Language gains. This chapter describes the research design, the population, instrumentation, data collection methods, and methods of analysis used in the study.

Research Design

The causal-comparative quantitative approach to exploring cause-and-effect relationships was employed in this study. The purpose of this analysis was to detect an association between variables. This method is sometimes referred to as ex-post facto research because causes are studied after they have exerted their effects on other variables (Gall, Borg, & Gall, 1996). Even though this design does not provide for a direct test of causation, it will provide information that will support or refute causal explanations. In this study, the effect of multi-age grouping and single age grouping was analyzed and compared using TCAP Reading/Language raw gain scores. Raw gain scores were used to determine the amount of progress children make from one year to the next regardless of their level of achievement. Findings could suggest a link between program design and increased Reading/Language gains.

Population

The focus population of this study consisted of 81 students who comprised the fifth-grade cohort in a multi-age elementary school located in Northeast, Tennessee, along with their 62 counterparts at a single-age school in the same school system. One hundred forty-three students in all were included in the study. Those who have not attended the school for a minimum of two years will not be included. The fifth grade cohort was chosen due to research supporting the proposal that students reach their peak of achievement at the end of the third year in a multi-age setting (Evanshen, 2001).

Sampling Method

Cluster sampling was used to select the population for this study. Gall, Borg, and Gall (1996) defined cluster sampling as a sample selected due to a group of naturally occurring groups in the population. Cluster sampling is used when it is more practical to select groups of individuals rather than individuals from a defined population (Gall et al.).

Instrumentation

Academic achievement between the two groups was compared through the use of scores on the Tennessee Comprehensive Assessment Program (TCAP). This state-mandated test is designed to measure achievement in the basic skills in grades 3 through 12. The test used multiple-choice questions and had set time limits. The subject area chosen for comparison was Reading/Language. Raw gain scores were used to obtain a

more comprehensive examination of the progress made by students during the school year.

The TCAP, published by the *Comprehensive Test of Basic Skills* (CTB McGraw-Hill, 2005) is a nationally normed achievement test and is used as an accountability measure by the state of Tennessee. The *TCAP CTB* provided criterion-referenced data along with corresponding scale scores that demonstrate growth over time. Criterion-referenced information also allows the comparison of student achievement against a specified level of performance. CTB-McGraw-Hill reported that its measure of achievement has a high degree of content, criterion, and construct validity (Tennessee Department of Education, 2005).

Data Collection

Approval to initiate this study was obtained from the Institutional Review Board at East Tennessee State University prior to any data collection. Written permission to conduct this study was obtained from authorized personnel in the involved school district (see Appendix A). School principals were subsequently contacted and briefed concerning the specifics of the study.

Classroom teachers administered the CTBS to the fifth grade students in the same week in April during the 2003-2005 school years. The test forms were sent to Nashville, Tennessee. They were scanned in Nashville and exported to CTB/McGraw Hill for scoring.

The source of data comparison was the Normal Curve Equivalent score (NCEs). These raw scores were used to calculate gains from one test to the next. The NCE was an

equal-interval score that can be treated arithmetically as interval data (Cannon, 2000). NCE raw scores for Reading/Language were used to make comparisons for differences between groups. Comparisons were made to determine if differences in Reading/Language gains existed between students in the multi-age classrooms and the single-age classrooms. Reading/Language raw gain scores (NCEs) were examined for the years 2003-2005. Secondly, comparisons were made between students completing all three years in multi-age classrooms and students who came to the multi-age classroom after the third grade year. The same comparisons were made between students completing all three years in single-age school and students who came to the single-age school after the third grade year

Comparisons were also made to determine if a difference existed between males and females in the population and to determine if there were interactions between gender and program design. Data collection forms included designations for gender and program design in the format for this purpose.

Data Analysis

As an initial step in the data analysis, descriptive statistics were computed to provide a profile of the population being studied. Data used in the statistical analyses for this study came from the TCAP CTBS. The Statistical Program for the Social Sciences (SPSS) was used to analyze data. t-tests for independent means were used to identify differences in the Reading/Language gains of students in the multi-age classrooms and single-age classrooms after what are traditionally known as the third, fourth, and fifth grade school years. t-tests for independent means were analyzed to determine if a

difference in gains achieved existed between students who remained at each school for all three years and those who did not. Gender differences were analyzed using t-tests for independent means to answer the final research questions and to identify interactions between student gender and the type of instructional program design.

All statistical tests were conducted using a preset alpha level of .05 to determine if statistically significant differences occurred in the Reading/Language gain scores of students in multi-classrooms and single-age classrooms by program design, gender, or an interaction of the two.

Hypotheses

For each research question, null hypotheses have been developed and examined at the .05 alpha level:

- HO₁ There are no differences in the Reading/Language raw gain scores on the 2002-2003 TCAP Standardized Assessment between students traditionally known as third graders in a multi-age school program and their counterparts in a graded classroom.
- HO₂ There are no significant differences in the Reading/Language raw gain scores on the 2003-2004 TCAP Standardized Assessment between students are traditionally known as fourth graders in a multi-age school program and their counterparts in a graded classroom.
- HO₃ There are no differences in the Reading/Language gain scores on the 2004-2005 TCAP Standardized Assessment between students traditionally known as fifth

graders in a multi-age school program and their counterparts in a graded classroom.

HO₄ There are no differences in the Reading/Language gain scores on the 2003 TCAP Standardized Assessment between male students traditionally known as third graders in a multi-age school program and their counterparts in a graded classroom.

HO₅ There are no differences in the Reading/Language gain scores on the 2004 TCAP Standardized Assessment between male students traditionally known as fourth graders in a multi-age school program and their counterparts in a graded classroom.

HO₆ There are no differences in the Reading/Language gain scores on the 2005 TCAP Standardized Assessment between male students traditionally known as fifth graders in a multi-age school program and their counterparts in a graded classroom.

HO₇ There are no differences in the Reading/Language gain scores on the 2003 TCAP Standardized Assessment between female students traditionally known as third graders in a multi-age school program and their counterparts in a graded classroom.

HO₈ There are no differences in the Reading/Language gain scores on the 2004 TCAP Standardized Assessment between female students traditionally known as fourth graders in a multi-age school program and their counterparts in a graded classroom.

HO₉ There are no differences in the 2005 TCAP Standardized Assessment between

female students traditionally known as fifth graders in a multi-age school program and their counterparts in a graded classroom.

HO₁₀: There are no differences in the Reading/Language gain scores on the 2005 TCAP Standardized Assessment between male and female students during what is traditionally known as the third grade in a multi-age school.

HO₁₁: There are no differences in the Reading/Language gain scores on the 2004 TCAP Standardized Assessment between male and female students during what is traditionally known as the fourth grade in a multi-age school.

HO₁₂: There are no differences in the Reading/Language gain scores on the 2003 TCAP Standardized Assessment between male and female students during what is traditionally known as the fifth grade in a multi-age school.

HO₁₃: There are no differences in the Reading/Language gain scores on the 2003 TCAP Standardized Assessment between male and female students during what is traditionally known as the third grade in a single-age school.

HO₁₄: There are no differences in the Reading/Language gain scores on the 2004 TCAP Standardized Assessment between male and female students during what is traditionally known as the fourth grade in a single-age school.

HO₁₅: There are no differences in the Reading/Language gain scores on the 2005 TCAP Standardized Assessment between male and female students during what is traditionally known as the fifth grade in a single age school.

Chapter Summary

Chapter three presented the methodology and procedures to be used in this study. The population and selection method were described. The casual-comparative research method was chosen and explained. TCAP CTBS along with its reliability and validity was presented. The methods of data collection and data analysis were detailed.

CHAPTER 4

RESULTS

The findings of this study along with the research questions and hypotheses presented in Chapter 1 are addressed in this chapter. The purpose of this study was to explore the differences between Reading/Language achievement gains of students in multi-age classrooms to the Reading/Language achievement gains of their peers in traditional, single-age classrooms. Further analyses were conducted to determine if specific groups benefited more by specific grade configurations. Five research questions and 15 null hypotheses were tested.

Demographic Information

The multi-age school in this study served children who are traditionally understood to be in kindergarten through fifth grade. Five hundred fourteen students were enrolled in the school at the time of the study. Approximately 31.3% of students enrolled qualified as economically disadvantaged. According to CRT data reported on the State Report Card of Tennessee, for the 2004 school year, the school received A's in the areas of Reading and Math and B's in the areas of Science and Social Studies (Tennessee Department of Education, 2004). The 2005 CRT report card achievement data had not been published at the time of the study. Specific demographic information concerning the ethnic composition of the student body is presented in Table 1 and Figure 1.

Table 1

Demographic Profile of Ethnicity for 2005 Multi-age School

Ethnicity	<i>n</i>	<i>Percent</i>
White	491	86.6
African American	46	8.1
Hispanic	20	3.5
Asian	6	1.1
Native American	4	.07

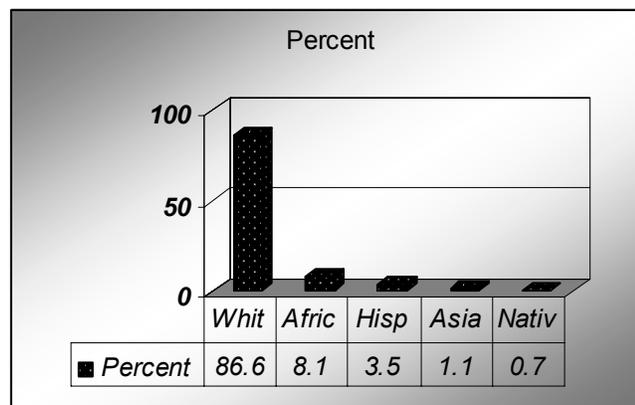


Figure 1. Demographic Profile of Ethnicity for 2005 Multi-age School

The single-age school in this study served children who are traditionally understood to be in kindergarten through fifth grade. Five hundred seventeen students were enrolled in the school at the time of the study. Approximately 30.6% of students

enrolled qualified as economically disadvantaged. According to CRT data reported by the State Report Card of Tennessee, the school received A's in the areas of Reading and Math and B's in the areas of Science and Social Studies. Specific demographic information concerning the ethnic make up of the student body is presented in Table 2 and Figure 2.

Table 2

Demographic Profile of Ethnicity for 2005 Single-age School

Ethnicity	<i>n</i>	<i>Percent</i>
White	494	91.7
African American	14	2.6
Hispanic	13	2.4
Asian	7	3.3
Native American	1	.2

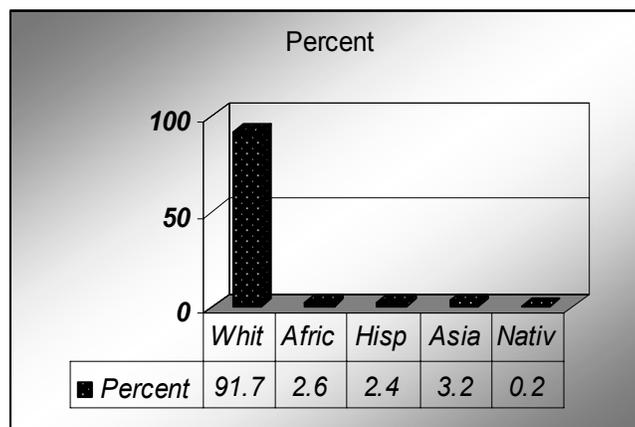


Figure 2. Demographic Profile of Ethnicity for 2005 Single-age School

Eighty-one students comprised the fifth graders in multi-age classrooms. Sixty-two students comprised the fifth graders in single-age classrooms. According to CRT data reported on the State Report Card of Tennessee, for the 2004 school year, the school received A's in the areas of Reading and Math and B's in the areas of Science and Social Studies (Tennessee Department of Education, 2004). The 2005 CRT report card achievement data had not been published at the time of the study. The demographic information for the 2005 fifth grade cohort is presented in Table 3.

Table 3

Demographic Profile of 2005 Fifth Grade Cohort

Student Cohort (2005) (5 th Grade)	Male	Female	Total
Multi-age Students	33	48	81
Single-age Students	40	31	71
Total	73	79	152

As shown in Table 3, there were 81 students included from multi-age classrooms and 71 students included from single-age classrooms. There were 152 students total.

Research Question #1

To what extent, if any, are there differences between the TCAP Reading/Language raw gain scores between children in multi-age classrooms and the

scores of children in single-graded classrooms? A t-test for independent means was used (alpha level of .05) to address question 1 and null hypotheses 1 through 3.

HO₁. There are no differences in the Reading/Language raw gain scores on the 2002-2003 TCAP Standardized Assessment between students traditionally known as third graders in a multi-age school program and their counterparts in a single-age classroom.

Table 4

Differences in 2003 Reading/Language Raw Gains Between Multi-age and Single-age Students

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Multi-age	65	54.09	19.37	2.2	*.03	.38
Single-age	70	61.47	19.52			

* Significance found at the .05 alpha level

Table 4 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as third grade students in multi-age and single-age classrooms in 2003. The test was significant, $t(133)= 2.2$, $p=.03$. The NCE raw gain scores for Reading/Language on the *TCAP* achievement test were higher in the single-age classrooms during the 2002-2003 school year. The 95% confidence interval for the difference in means ranged from 14.00 to .75. However, the effect size value of .38 indicates that only a small effect of the variance between means can be accounted for by the independent variable.

HO₂. There are no significant differences in the Reading/Language raw gain scores on the 2003-2004 TCAP Standardized Assessment between students who are

traditionally known as fourth graders in a multi-age school program and their counterparts in a graded classroom.

Table 5

Differences in 2004 Reading/Language Raw Gains Between Multi-age and Single-age Students

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Multi-age	72	54.09	19.45	2.05	*.04	.34
Single-age	71	61.46	18.92			

* Significance found at the .05 alpha level

Table 5 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as fourth grade students in multi-age and single-age classrooms in 2004. The test was significant, $t(141) = 2.05, p = .04$. The NCE raw gain scores for Reading/Language on the *TCAP* achievement test were higher in the single-age classrooms during the 2003-2004 school year. The 95% confidence interval for the difference in means ranged from 12.91 to .22. However, the effect size value of .34 indicates that only a small effect of the variance between means can be accounted for by the independent variable.

HO₃. There are no differences in the Reading/Language gain scores on the 2004-2005 TCAP Standardized Assessment between students traditionally known as fifth graders in a multi-age school program and their counterparts in a graded classroom.

Table 6

Differences in 2005 Reading/Language Raw Gains Between Multi-age and Single-age Students

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Multi-age	81	57.67	16.42	2.65	*.009	.43
Single-age	73	64.92	17.58			

* Significance found at the .05 alpha level

Table 6 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as fifth grade students in multi-age and single-age classrooms in 2005. The test was significant, $t(152) = 2.65$, $p = .009$. The NCE raw gain scores for Reading/Language on the *TCAP* achievement test were higher in the single-age classrooms during the 2004-2005 school year. The 95% confidence interval for the difference in means ranged from 12.66 to 1.84. However, the effect size value of .43 indicates that only a small effect of the variance between means can be accounted for by the independent variable.

Research Question #2

To what extent, if any, are there differences in the TCAP Reading/Language raw gain scores between males in multi-age classrooms and males in single-graded classrooms? A t-test for independent means was used (alpha level of .05) to address question 2 and null hypotheses 4 through 6.

HO₄. There are no differences in the Reading/Language gain scores on the 2003 TCAP Standardized Assessment between male students traditionally known as third graders in a multi-age school program and their counterparts in a graded classroom.

Table 7

Differences in 2003 Reading/Language Raw Gains Between Male Multi-age and Single-age Students

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Multi-age Male Students	27	49.85	21.47	1.05	.30	.26
Single-age Male Students	38	55.00	17.93			

* Significance found at the .05 alpha level

Table 7 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as third grade male students in multi-age and single-age classrooms in 2003. The null hypothesis was retained $t(63) = 1.05, p = .30$. No statistical differences were found between the means of male students in multi-age and single-age classrooms in 2003. The 95% confidence interval for the difference in means ranged from 14.95 to 4.64. The effect size value of .26 indicates that only a small effect of the variance between means can be accounted for by the independent variable.

HO₅. There are no differences in the Reading/Language gain scores on the 2004 TCAP Standardized Assessment between male students traditionally known as fourth graders in a multi-age school program and their counterparts in a graded classroom.

Table 8

Differences in 2004 Reading/Language Raw Gains Between Male Multi-age and Single-age Students

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Multi-age Male Students	29	54.50	18.72	.06	.96	.01
Single-age Male Students	38	54.76	16.72			

* Significance found at the .05 alpha level

Table 8 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as fourth grade male students in multi-age and single-age classrooms in 2003. The null hypothesis was retained $t(65) = .06, p = .96$. No statistical differences were found between the means of male students in multi-age and single-age classrooms in 2004. The 95% confidence interval for the difference in means ranged from 8.92 to 8.42. The effect size value of .01 indicates that a very small effect of the variance between means can be accounted for by the independent variable.

HO₆. There are no differences in the Reading/Language gain scores on the 2005 TCAP Standardized Assessment between male students traditionally known as fifth graders in a multi-age school program and their counterparts in a graded classroom.

Table 9

Differences in 2005 Reading/Language Raw Gains Between Male Multi-age and Single-age Students

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Multi-age Male Students	34	56.56	18.18	.75	.46	.18
Single-age Male Students	39	59.62	16.72			

* Significance found at the .05 alpha level

Table 9 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as fifth grade male students in multi-age and single-age classrooms in 2005. The null hypothesis was retained $t(71) = .75, p = .46$. No statistical differences were found between the means of male students in multi-age and single-age classrooms in 2005. The 95% confidence interval for the difference in means ranged from 11.02 to 5.09. The effect size value of .18 indicates that only a small effect of the variance between means can be accounted for by the independent variable.

Research Question #3

To what extent, if any, are there differences in the TCAP Reading/Language raw gain scores between females in multi-age classrooms and females in single-graded classrooms? A t-test for independent means was used (alpha level of .05) to address question 2 and null hypotheses 7 through 9.

HO₇. There are no differences in the Reading/Language gain scores on the 2003 TCAP Standardized Assessment between female students traditionally known as third graders in a multi-age school program and their counterparts in a graded classroom.

Table 10

Differences in 2003 Reading/Language Raw Gains Between Female Multi-age and Single-age Students

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Multi-age Female Students	40	58.15	17.88	2.36	*.02	.57
Single-age Female Students	30	68.57	18.84			

* Significance found at the .05 alpha level

Table 10 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as third grade female students in multi-age and single-age classrooms in 2003. The test was significant, $t(68)=2.36$, $p=.02$. The NCE raw gain scores for Reading/Language on the *TCAP* achievement test were higher for females in single-age classrooms during the 2002-2003 school year. The 95% confidence interval for the difference in means ranged from 19.23 to 1.6. The

effect size value of .57 indicates that a moderate effect of the variance between means can be accounted for by the independent variable.

HO₈. There are no differences in the Reading/Language gain scores on the 2004 TCAP Standardized Assessment between female students traditionally known as fourth graders in a multi-age school program and their counterparts in a graded classroom.

Table 11

Differences in 2004 Reading/Language Raw Gains Between Female Multi-age and Single-age Students

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Multi-age Female Students	45	56.36	29.61	2.59	*.01	.61
Single-age Female Students	31	68.35	18.63			

* Significance found at the .05 alpha level

Table 11 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as fourth grade female students in multi-age and single-age classrooms in 2004. The test was significant, $t(74)=2.59$, $p=.01$. The NCE raw gain scores for Reading/Language on the *TCAP* achievement test were higher for females in single-age classrooms during the 2003-20054 school year. The 95% confidence interval for the difference in means ranged from 21.22 to 2.78. The effect size value of .61 indicates that a moderate effect of the variance between means can be accounted for by the independent variable.

HO₉. There are no differences in the 2005 TCAP Standardized Assessment between female students traditionally known as fifth graders in a multi-age school program and their counterparts in a graded classroom.

Table 12

Differences in 2005 Reading/Language Raw Gains Between Female Multi-age and Single-age Students

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Multi-age Female Students	49	59.10	15.18	3.21	*.002	.72
Single-age Female Students	32	70.81	17.28			

* Significance found at the .05 alpha level

Table 12 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as fifth grade female students in multi-age and single-age classrooms in 2005. The test was significant, $t(79) = 3.21$, $p = .002$. The NCE raw gain scores for Reading/Language on the *TCAP* achievement test were higher for females in single-age classrooms during the 2004-2005 school year. The 95% confidence interval for the difference in means ranged from 18.96 to 4.46. The effect size value of .72 indicates that a moderate effect of the variance between means can be accounted for by the independent variable.

Research Question # 4

To what extent, if any, are there differences in the *TCAP* Reading/Language raw gain scores between males and females in multi-age classrooms? A t-test for independent means was used (alpha level of .05) to address question 1 and null hypotheses 11 through 13.

HO₁₀. There are no differences in the *TCAP* Reading/Language gain scores on the 2003 TCAP Standardized Assessment between male and female students during what is traditionally known as the third grade in a multi-age school.

Table 13

Differences in 2003 Reading/Language Raw Gains Between Males and Females in a Multi-age Classroom

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Multi-age Male Students	27	49.85	21.47	1.72	*.09	.42
Multi-age Female Students	38	58.15	17.88			

* Significance found at the .05 alpha level

Table 13 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as third grade male and female students in multi-age classrooms in 2003. The test was significant, $t(65)=1.72$, $p=.09$. The NCE raw gain scores for Reading/Language on the *TCAP* achievement test were higher for females in multi-age classrooms than males in multi-age classrooms during the 2004-2005 school year. The 95% confidence interval for the difference in

means ranged from 17.95 to 1.35. The effect size value of .42 indicates that a moderate effect of the variance between means can be accounted for by the independent variable.

HO₁₁: There are no differences in the Reading/Language gain scores on the 2004 *TCAP* Standardized Assessment between male and female students during what is traditionally known as the fourth grade in a multi-age school.

Table 14

Differences in 2004 Reading/Language Raw Gains Between Males and Females in a Multi-age Classroom

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Multi-age Male Students	29	54.53	18.72	.39	.70	.03
Multi-age Female Students	45	56.36	20.61			

* Significance found at the .05 alpha level

Table 14 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as fourth grade male and female students in multi-age classrooms in 2004. The null hypothesis was retained $t(72) = .39, p = .70$. No statistical differences were found between the means of male and female students in multi-age classrooms in 2004. The 95% confidence interval for the difference in means ranged from 11.28 to 7.61. The effect size value of .03 indicates that a small effect of the variance between means can be accounted for by the independent variable.

HO₁₂. There are no differences in the Reading/Language gain scores on the 2005 *TCAP* Standardized Assessment between male and female students during what is traditionally known as the fifth grade in a multi-age school.

Table 15

Differences in 2005 Reading/Language Raw Gains Between Males and Females in a Multi-age Classroom

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Multi-age Male Students	34	56.56	8.18	.70	.50	.15
Multi-age Female Students	49	59.10	15.18			

* Significance found at the .05 alpha level

Table 15 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as fifth grade male and female students in multi-age classrooms in 2005. The null hypothesis was retained $t(81) = .70, p = .50$. No statistical differences were found between the means of male and female students in multi-age classrooms in 2005.

Research Question #5

To what extent, if any, are there differences in the *TCAP* Reading/Language raw gain scores between males and females in single-graded classrooms? A t-test for independent means was used (alpha level of .05) to address question 1 and null hypotheses 14 through 16.

HO₁₃. There are no differences in the Reading/Language gain scores on the 2003 *TCAP* Standardized Assessment between male and female students during what is traditionally known as the third grade in a single age school.

Table 16

Differences in 2003 Reading/Language Raw Gains Between Males and Females in a Single-age Classroom

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Single-age Male Students	38	55.00	17.93	3.03	*.003	.74
Single-age Female Students	30	68.57	18.83			

* Significance found at the .05 alpha level

Table 16 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as third grade male and female students in single-age classrooms in 2003. The test was significant, $t(66)=3.03$, $p=.003$. The NCE raw gain scores for Reading/Language on the *TCAP* achievement test were higher for females in single-age classrooms than males in single-age classrooms during the 2002-2003 school year. The 95% confidence interval for the difference in means ranged from 22.51 to 4.6. The effect size value of .74 indicates that a moderate to high effect of the variance between means can be accounted for by the independent variable.

HO₁₄. There are no differences in the Reading/Language gain scores on the 2004 TCAP Standardized Assessment between male and female students during what is traditionally known as the fourth grade in a single-age school.

Table 17

Differences in 2004 Reading/Language Raw Gains Between Males and Females in a Single-age Classroom

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Single-age Male Students	38	54.76	16.71	3.2	*.002	.77
Single-age Female Students	31	68.35	18.63			

* Significance found at the .05 alpha level

Table 17 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as fourth grade male and female students in single-age classrooms in 2004. The test was significant, $t(67)=3.19$, $p=.80$. The NCE raw gain scores for Reading/Language on the *TCAP* achievement test were higher for females in single-age classrooms than males in single-age classrooms during the 2003-2004 school year. The 95% confidence interval for the difference in means ranged from 22.09 to 5.09. The effect size value of .77 indicates that a moderate effect of the variance between means can be accounted for by the independent variable.

HO₁₅. There are no differences in the Reading/Language gain scores on the 2005 TCAP Standardized Assessment between male and female students during what is traditionally known as the fifth grade in a single-age school.

Table 18

Differences in 2005 Reading/Language Raw Gains Between Males and Females in a Single-age Classroom

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Single-age Male Students	39	59.62	16.72	2.76	*.007	.66
Single-age Female Students	32	70.81	17.28			

* Significance found at the .05 alpha level

Table 18 presents the results of a t-test for independent means used to evaluate Raw Reading/Language gains for what are traditionally referred to as fifth grade male and female students in single-age classrooms in 2005. The test was significant, $t(69)=2.76$, $p=.007$. The NCE raw gain scores for Reading/Language on the TCAP achievement test were higher for females in single-age classrooms than males in single-age classrooms during the 2004-2005 school year. The 95% confidence interval for the difference in means ranged from 19.27 to 3.12. The effect size value of .66 indicates that a moderate effect of the variance between means can be accounted for by the independent variable.

Chapter Summary

The results of data collected were presented in Chapter 4 with accompanying analyses. A series of independent samples t-tests were performed to determine if significant differences in NCE raw scores for Reading/Language existed between students in multi-age classrooms and students in single-age classrooms. The results were mixed.

Significant differences were found in the scores of children in multi-age and single age classrooms in the 2003, 2004, and 2005 school years. The scores in the single-age classrooms were higher than those of multi-age classrooms during this time. Although significant differences were found, the effect size of each was small. Therefore, no practical significance was determined.

There were no significant differences found between males in multi-age classrooms and males in single age classrooms in the years 2003, 2004, and 2005.

There were significant differences found in the NCE raw gain scores for Reading/Language for females in multi-age and single-age classrooms for the years 2003, 2004, and 2005. The differences favored the females in single-age classrooms. Effect size was calculated and determined to be in the moderate category. There might be a moderate level of practical significance favoring females in single-age classrooms.

The results of the scores of males and females within multi-age classrooms were analyzed to determine whether significant differences existed. The results were mixed. There was a significant difference found in the 2002-2003 scores that favored females. However, the effect size was calculated and indicated no practical significance. For the

years 2004 and 2005, there were no significant differences found between males and females in multi-age classrooms.

Male and female scores were also analyzed within single-age classrooms to determine whether significant differences existed. Significant differences were found in the 2003, 2004, and 2005 scores that favored females. Effect size was calculated for each and revealed moderate practical significance.

Chapter 5 presents an analysis of the results of the study highlighted in this chapter. It provides a summary of the study and presents the specific findings associated with each research question. Additionally, the final chapter presents a summary of conclusions that might be drawn from the study as well as recommendations for further study and practice.

CHAPTER 5

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

This study was conducted to explore the differences between Reading/Language achievement gains of students in multi-age classrooms to the Reading/Language achievement gains of their peers in traditional, single-age classrooms. Chapter 5 provides a summary of the findings of the study and provides conclusions and recommendations for further study and practice.

Summary of the Study

The *No Child Left Behind Act* challenges schools to “close the gap” that separates the achievement levels of children from low socioeconomic groups, English language learners, ethnic subgroups and special education students from their more privileged peers (U.S. Department of Education, 2004). The increased emphasis on high stakes testing and accountability has caused many schools to search for researched based structures and methods of instruction that help meet this demand while staying true to what is developmentally appropriate for students.

A review of the literature presents the history of the American educational structure from the time of the one room schoolhouse to the accepted single-graded model of today. The literature challenges the graded structure by arguing that a child’s chronological age does not dictate his or her mental capabilities (Pavan, 1992). Recommendations for differentiating instruction for students is provided; adding however, that multi-age classrooms offer additional flexibility for grouping students who

are working at similar levels without having to affect the promotion or retention of children within a specific grade (Hillson & Bongo, 1971).

Research focusing on multi-age programs reveals that there are cognitive and social benefits for students. Furthermore, it encourages teachers to use more hands-on, developmentally appropriate methods of instruction (Gaustad, 1997). However, there is not a great deal of recent research regarding multi-age programs as measured by standardized tests (Pavan, 1992). This researcher attempted to add to recent research involving multi-age classrooms using standardized tests as the dependent variable.

The primary goal of this study was to explore the differences between Reading/Language achievement gains of students in multi-age classrooms to the Reading/Language achievement gains of their peers in traditional, single-age classrooms. In this study, the effect of multi-age grouping and single age grouping was analyzed and compared using TCAP Reading/Language raw gain scores. Raw gain scores were used to determine the amount of progress children make from one year to the next regardless of their level of achievement.

Summary of the Findings

The descriptive data associated with the research questions of the study were mixed. Few significant differences were found between students of the 2005 fifth grade cohort that participated in multi-age and single- classrooms from 2003-2005. In addition the calculation of effect size determined that there was little practical significance for students in either group as measured by this study. Each research question and its associated findings are summarized below.

Research Question #1

To what extent, if any, are there differences between the TCAP Reading/Language raw gain scores between children in multi-age classrooms and the scores of children in single-graded classrooms?

Students in single-age classes had higher Reading/Language NCE raw gain score means than their multi-age peers on the *TCAP* (CTB McGraw-Hill, 2005) as indicated by the independent samples *t* test. However, when effect size was calculated only a “small” practical effect was found.

These results were contrary to the majority of research conducted and reviewed by Pavan (1992) and Goodlad and Anderson (1963) whose results favored multi-age programs almost consistently.

Research Question #2

To what extent, if any, are there differences in the TCAP Reading/Language raw gain scores between males in multi-age classrooms and males in single-graded classrooms?

There were no significant differences found between Reading/Language NCE raw gain score means between males in multi-age classrooms and males in single-age classrooms on the *TCAP* (CTB McGraw-Hill, 2005) as indicated by the independent samples *t* test.

Research Question #3

To what extent, if any, are there differences in the TCAP Reading/Language raw gain scores between females in multi-age classrooms and females in single-graded classrooms?

Females in single-age classes had higher Reading/Language NCE raw gain score means than their female multi-age peers on the *TCAP* (CTB McGraw-Hill, 2005) as indicated by the independent samples *t* test. There was a “moderate” effect found between the scores of males and females for the females in single-age classrooms.

Research Question #4

To what extent, if any, are there differences in the TCAP Reading/Language raw gain scores between males and females in multi-age classrooms?

The results were mixed. There were significant differences in the 2003 Reading/Language NCE raw gain score means favoring females on the *TCAP* (CTB McGraw-Hill, 2005) as indicated by the independent samples *t* test. When effect size was calculated a “moderate” practical effect was found. There were no significant differences between the Reading/Language NCE raw gain scores of males and females in multi-age classrooms in 2004 and 2005 on the *TCAP* (CTB McGraw-Hill, 2005) as indicated by the independent samples *t* test.

Although the female scores were consistently higher than male scores in multi-age classrooms, the male scores grew significantly the longer the males participated in the multi-age classroom. The 2003 scores indicated a significant difference favoring the females with effect size indicating a moderate practical effect. Two years later, there

were no significant differences between male and female scores. These results seem to indicate an agreement with Pavan's research in 1992 that suggested added benefits for males in multi-age programs as well as increased gains for students the longer they participate in multi-age programs (Pavan, 1992).

Research Question #5

To what extent, if any, are there differences in the TCAP Reading/Language raw gain scores between males and females in single-graded classrooms?

Females in single-age classes had higher Reading/Language NCE raw gain score means than their single-age male peers on the *TCAP* (CTB McGraw-Hill, 2005) as indicated by the independent samples *t* test. Effect size calculations were mixed. For the 2003 school year, effect size calculations indicated a "moderate to high" practical effect. There was a "moderate" effect found between the scores of males and females for the 2004 and 2005 school year.

Conclusions

Educators are faced with the difficult task of teaching children of varying backgrounds and ability levels. Developmentally appropriate practices (DAP) result from professionals considering the individual intellectual, social, emotional, and cultural needs of each child (Bredekamp & Copple, 1997). The *No Child Left Behind Act* holds schools accountable to meet the needs of children from low socioeconomic groups, English language learners, ethnic subgroups, and special education students from their more privileged peers (U.S. Department of Education, 2004). Research is available that

indicates multi-age programs lead to higher academic success than single-age programs (Goodlad & Anderson, 1963; Pavan 1992). However, in this small population, the results did not yield the same consistent results.

Some conclusions can be drawn from this study that are consistent with past research focusing on multi-age programs. Longitudinal studies such as Carter, 1974; Eells, 1970; Evanshen, 2001; Killough, 1971; Morris, Proger, Morrell, 1971; Pavan, 1977, 1992; Perrin, 1969; Ramayya, 1972; and Walker, 1973 reveal that the longer a student participates in a multi-age program, the better the results. Evanshen (2001) suggests that student scores reach their highest after three years in a multi-age program. In each analysis performed from 2003 to 2005, the Reading/Language NCE raw gain score mean for each multi-age group improved. Thus, the yearly gains led to the highest mean peaking after the 2005 fifth grade year.

Pavan (1992) reported that there are added benefits for males in multi-age programs. Although there were no significant differences in the Reading/Language gains of males from multi-age and single-age classrooms in this study, there were interesting results in the male/female comparison within the multi-age classroom. The 2003 school year resulted in a significant difference favoring females in the multi-age classroom. In fact, effect size indicated that the practical effect of female over male performance was moderate. However, these results changed over time. The scores of the males improved. By 2004, there were no significant differences between the same group of males and females and likewise, no practical effect. These results seem to correspond with Pavan's research indicating that males may benefit more than females from their participation in multi-age programs (Pavan, 1992).

The majority of the results of this specific study did not correlate with past studies. Therefore, a series of recommendations are provided for the researcher interested in following up on the findings of this study.

Recommendations for Practice

As a result of this study the following recommendations are offered;

Regardless of what structure is in place in schools, the need to individualize student instruction is paramount. Developmentally Appropriate Practices (Bredekamp & Copple, 1997) should guide decision making regarding students instruction in classrooms. These decisions should be based on the following:

1. What is known about child development and learning-knowledge of age related characteristics that permits general predictions within an age range about what activities, materials, interactions, or experiences will be safe, healthy, interesting, achievable and also challenging to children;
2. What is known about the strengths, interests, and needs of individual child in the group to be able to adapt for and be responsible to inevitable individual variation;
3. Knowledge of the social and cultural contexts in which children live to ensure that learning experiences are meaningful, relevant, and respectful for the participating children and their families (pp. 8-9).

Administrators and teachers need to be aware of the need to participate in appropriate training and staff development regarding developmentally appropriate instructional strategies.

Recommendations for Further Study

As a result of this study the following recommendations are offered;

1. Further quantitative studies should be conducted in other multi-age classrooms to expand research that uses standardized measures to determine one aspect of the effectiveness of multi-age programs.
2. Further qualitative studies should be conducted to expand research concerning the effectiveness of multi-age programs using measurements other than standardized achievement test data.
3. Attitudinal surveys should be performed to expand research suggesting that students in multi-age classrooms have better attitudes toward school.
4. Longitudinal data should be collected to determine whether a comparison exists between the graduation rates for students from multi-age classrooms and single-age classrooms.
5. Quantitative studies should be conducted to investigate the comparison of retention and promotion rates between students from multi-age and single-age classrooms.
6. Quantitative studies should be conducted to determine whether there is a relationship between the number of discipline referrals of students and their school structure (multi-age/single-age).
7. Quantitative studies should be conducted to compare the academic growth of multi-age and single-age students as related to their socioeconomic status.
8. Qualitative studies should be conducted to compare satisfaction of parents of children from multi-age and single-age classrooms.

REFERENCES

- Anderson, R. & Pavan, B. (1993). *Nongradedness: Helping it to happen*. Lancaster, PA: Technomic.
- Bernheisel, S. (1992). Give it a try. *Principal*, 72, 22.
- Bloom, B. (1976). *Human characteristics and school learning*. New York: McGraw-Hill.
- Blythe, T., & Gardner, H. (1990). A school for all intelligences. *Educational Leadership*, 47, 33-37.
- Bredenkamp, S., & Copple, C. (1997). *Developmentally appropriate practice in early childhood programs*. Washington, DC: NAEYC.
- Cannon, T. (2000, Fall). Information about norm-referenced scores. *Tennessee Assessment Newsletter*, 1, 2.
- Carter, J. (1974). A study of the effects of multi-grade grouping on the attitudes toward self, others and school of selected third and fifth grade students. *Dissertation Abstracts*, 34, 5469A.
- Clay, M. (1993). *An observation survey of early literacy achievement*. Hong Kong: Heinemann.
- CTB/McGraw-Hill (2005). CTB/McGraw-Hill. Retrieved March 5, 2005 from <http://www.tams.net/TAMSJourSpr05.pdf>
- Cohen, E. (1981). Sociology looks at team teaching. *Research in Sociology of Education and Socialization*, 2, 163-193.
- Connell, D. (1987). The first 30 years were the fairest: Notes from the kindergarten and ungraded primary. *Young Children*, 42, 30-39.
- Coopersmith, S. (1981). *The antecedents of self-esteem*. Palo Alto, CA: Consulting Psychologists Press.
- Cushman, K. (1990, Summer). The whys and hows of the multi-age primary classroom. *American Educator*, 14, 20-25.
- Eells, C. (1970). A study of academic ability of immigrant pupils over five years in a suburban nongraded school system. *Dissertation Abstracts*, 31, 579A.

- Evanshen, P. (2001). *An analysis of configurations in a nongraded elementary school in Northeast Tennessee*. Doctoral Dissertation: Johnson City, TN: East Tennessee State University
- Gall, M., Gorg, W., & Gall, J. (1996). *Educational research*. White Plains, NY: Longman, 12.
- Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.
- Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. New York: Basic Books.
- Gaustad, J. (1992a). Nongraded education: Mixed age, integrated, and developmentally appropriate education for primary children. *Oregon School Study Council*, 35(7), 1-38.
- Gaustad, J. (1992b). Nongraded primary education. *Research Roundup*, 9(1), 96-99.
- Gaustad, J. (1997). *Building support for multi-age education*. Eugene, OR: College of Education. University of Oregon. (ERIC Digest. 114, EDO-EA-97-6).
- Goodlad J., & Anderson, R. (1963). *The nongraded elementary school*. New York: Harcourt Brace.
- Grant, J., Johnson, B., & Richardson, I. (1995). *Multi-age Q&A – 101 practical answers to your most pressing questions*. Peterborough, NH: Crystal Springs Books.
- Hillson, M., & Bongo, J. (1971). *Continuous progress education: A practical approach*. Palo Alto, CA: Science Research Associates.
- Hunter, M. (1992). *How to change to a nongraded school*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Jensen, E. (2000). *Brain based learning*. San Diego, CA. Brain Store Publishing.
- Katz, L. (1995). The benefits of mixed-age grouping. *ERIC Digest*. Available: <http://ericps.ed.uiuc.edu/ece/pubs/digests/1995/lkmag95.html>.
- Katz, L., Evangelou, D., & Hartman, J. (1990). *The case for mixed age grouping in early education*. Washington, DC: National Association for the Education of Young Children.
- Kavolik, S. (1994). *ITI: The model integrated thematic instruction*. Kent, WA: Books for Educators.

- Kavolik, S. (2002). *Exceeding expectations: A user's guide to implementing brain research in the classroom*. Covington, WA: Susan Kavolik and Associates.
- Killough, C. (1971). *An analysis of the longitudinal effects that a nongraded elementary program, conducted in an open-space school, had on the cognitive achievement of pupils*. Doctoral Dissertation, Houston, TX: University of Houston.
- Lewis, J. (1969). *A contemporary approach to nongraded education*. West Nyack, N.Y: Parker.
- Lodish, R. (1992). The pros and cons of mixed age grouping. *Principal*, 71, 20-22.
- McLoughlin, W.P. (1967). *The nongraded school: A critical assessment*. NY, NY: The State Education Department.
- Milburn, D. (1981). A study of multi-age or family grouped classrooms. *Phi Delta Kappan*, 62, 513-514.
- Morris, V., Proger, B., & Morrell, J. (1971). Pupil achievement in a nongraded primary plan after three and five years of instruction. *Educational Leadership*, 29, 621-625.
- Musgrave, G. (1922). *Individualized instruction*. Boston: Allyn and Bacon.
- Nachbar, R. (1989). A k-1 class can work-wonderfully. *Young Children*, 44, 67-71.
- National Center for Educational Statistics (2003). *National Assessment of Educational Progress, the nation's report card: 2003 mathematics and reading results*. Retrieved March 12, 2005, from <http://nces.ed.gov/nationsreportcard/reading/results2003/raceethnicity.asp>
- No Child Left Behind, (2004). *No child left behind: A guide for policy makers*. Retrieved November 6, 2004, from <http://www.nclb.gov/>
- Otto, H.J. (1969). *Nongradedness: An elementary school evaluation*. University of Texas at Austin.
- Pavan, B. (1992). The benefits of nongraded schools. *Educational Leadership*, 50, 182-185.
- Pavan, B. (1977). The nongraded elementary school: research on academic achievement and mental health. *Texas Tech Journal of Education*, 4.

- Perrin, J. (1969) A statistical and time change analysis of achievement differences of children in a nongraded and graded program in selected schools in the Little Rock Public Schools. *Dissertation Abstracts*, 30, 530A.
- Piaget, J. (1973) *Memory and intelligence*. New York: Basic Books, Inc.
- Pratt, D. (1986). On the merits of multi-age classrooms. *Research in Rural Education*, 3, 111-115.
- Ramayya, D. (1972). Achievement skills, personality variables and classroom climate in graded and nongraded elementary schools. *Psychology in the Schools*, 9, 88-92.
- Ratey, J. (2001). *A user's guide to the brain: perception, attention, and the four theaters of the brain*. New York: Pantheon Books.
- Schwebel, M. & Raph, J. (1973). *Piaget in the classroom*. New York: Basic Books.
- Shanker, A. (1993). Multi-age classrooms: Children learning at their own speed. In D. Sumner (Ed.), *Multiage classrooms: The ungrading of America's schools*, 36-38. Peterborough, NH: SDE.
- Slavin, R. (1992). Response: The nongraded elementary school. *Educational Leadership*, 50, 24.
- Stone, S. (1994). Strategies for teaching children in multi-age classrooms. *Childhood Education*, 71, 102.
- Tennessee Department of Education (2002). *Tennessee Comprehensive Assessment Program*. Retrieved October 5, 2005 from <http://www.tennessee.gov/education/assessment/tsachtam.pdf>
- Tennessee Department of Education (2004) *Tennessee State Report Card*. Retrieved November 15, 2005 from www.k-12.state.tn.us/rptcrd04/school1.asp
- Tomlinson, C. (1999). *The differentiated classroom*. Alexandria, VA: ASCD.
- Uphoff, J. & Evans, D. (1992). The country school comes to town: A case study of multi-age grouping and teaching. In D. Sumner (Ed.), *Multiage classrooms: The ungrading of America's schools*, 36-38. Peterborough, NH: SDE.
- U.S. Department of Education (2004). News: No Child Left Behind: Closing the Achievement Gap in America's Public Schools. Retrieved March 5, 2005, from <http://www.ed.gov/nclb/overview/welcome/closing/edlite-slide002.html>
- Vygotsky, L. (1978). *Mind in society: The development of psychological processes*.

Cambridge, MA: Harvard University Press.

Walker, W. (1973). *Long term effects of graded and nongraded primary programs*.
Doctoral Dissertation, Nashville, TN: George Peabody College of Teachers.

Wiggins, G. (1993). *Assessing student performance*. San Francisco: Jossey Bass
Publishers.

APPENDIX
Letter of Permission

Holly Flora
XXXXXXXXXX
XXXXXXXXXX
XXXXXXXXXX
XXXXXXXXXX

XXXXXXXXXXXX
Director of Schools
XXXXXXXXXXXX
XXXXXXXXXXXX

Dear Dr. XXXXXX,

As a student at East Tennessee State University, I am currently involved in my dissertation phase of the Educational Leadership and Policy Analysis doctoral program. My dissertation will explore the relationship between multi-age programs versus single-age programs on Reading/Language gains using the TCAP achievement test.

I would like your permission to access and utilize non-identifiable scores on the TCAP from the years 2004-2005 for the classrooms selected for the study. Random numbers will be used to protect the identity of all participants.

In preparation for the study, I will contact the principal at each participating school and arrange for the collection of all necessary data with a minimum of disruption.

I believe the results of my study will be helpful in evaluating just one dimension of the success of these two programs within your school system. The results may also be helpful for those teachers or administrators who are considering the possibility of implementing a multi-age program.

Thank you for your cooperation.

Sincerely,

Permission is hereby granted to Holly I. Flora to access and use TCAP scores for third, fourth and fifth grade students who have participated in a multi-age and single-age program in 2004-2005.

Signature

Date

VITA

Holly Krysten Irvin Flora

- Personal Data: Date of Birth: April 10, 1974
Place of Birth: Johnson City, Tennessee
Marital Status: Married
- Education: Public Schools: Kingsport, Tennessee
Milligan College, Johnson City, Tennessee: Elementary and Early
Childhood Education, B.A., 1996
East Tennessee State University, Johnson City, Tennessee: Early
Childhood Education, M.A., 2001
East Tennessee State University, Johnson City, Tennessee:
Educational Leadership and Policy Analysis, Ed.D., 2005
- Professional
Experience: Teacher, Kingsport City Schools, Abraham Lincoln Elementary
School, August 1996-2002
Teacher, Kingsport City Schools, George Washington Elementary
School, August 2002-2005
Adjunct Faculty, East Tennessee State University, Johnson City,
Tennessee, 2001
- Additional
Experience: *(2005) ETSU Early Childhood Conference: Session Facilitator
*(2004) Facilitator of New Teacher Orientation, Kingsport City
Schools
*(2004) Staff Development Committee
*(2005) ETSU Early Childhood Conference: Session Facilitator
*(2002) Staff Development Intern