Fall 2002

Predictor variables of performance on the Louisiana Educational Assessment Program for the 21st Century

June Adeline Thomas
Louisiana Tech University

INFORMATION TO USERS

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

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

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

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

PREDICTOR VARIABLES OF PERFORMANCE ON THE LOUISIANA EDUCATIONAL ASSESSMENT PROGRAM FOR THE 21ST CENTURY

by

June A. Thomas, B.S., M.A.

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Education in Curriculum and Instruction

COLLEGE OF EDUCATION
LOUISIANA TECH UNIVERSITY

November, 2002

LOUISIANA TECH UNIVERSITY

THE GRADUATE SCHOOL

November 22, 2002
Date

We hereby recommend that the dissertation prepared under our supervision by June A. Thomas entitled Predictor Variables of Performance on the Louisiana Educational Assessment Program for the 21st Century be accepted in partial fulfillment of the requirements for the Degree of Doctor of Education.

Cathy Stockton
Supervisor of Dissertation Research

Head of Department

Curriculum, Instruction & Leadership
Department

Recommendation concurred in:

Thomas P Springer (Jr)

Advisory Committee

Approved:

Director of Graduate Studies

Dean of the College

Approved:

Dean of the Graduate School

ABSTRACT

Using the Student Characteristics: Iowa Test of Basic Skills (ITBS) composite national percentile scores, repeating fourth grade, spring third grade Developmental Reading Assessment (DRA) scores, socioeconomic status, race, gender, and preschool attendance and the School Characteristics: teacher degree level, teacher experience, teacher professional development, and school Title I status, this study employed a stepwise multiple regression analysis to determine the best predictors of fourth grade scores on the Louisiana Educational Assessment Program for the 21st Century (LEAP 21). Students must pass the English Language Arts (ELA) and Mathematics portions of this test to be promoted to fifth grade. Having previously failed either portion of the LEAP 21 or having scored at or below the 30th percentile on the ITBS has identified a student for the LEAP Tutoring Program. This small-group, pull-out program provides 10 weeks of instruction prior to the LEAP 21. ITBS and DRA scores were the strongest predictors of all LEAP 21 scores. Repeating fourth grade, attending preschool, attending a Title I school, and being taught by a teacher who participated in professional development also predicted positive ELA LEAP 21 scores. Being male and being White were added to the ITBS and DRA scores for the Mathematics LEAP 21 predictive model. Combining the ELA and Mathematics LEAP 21 scores caused the gender and Title I variables to drop out leaving ITBS, DRA, race, preschool, and repeating fourth grade from the student characteristics and teacher professional development from the school characteristics as the predictive model.

TABLE OF CONTENTS

ABSTRACT iii

LIST OF TABLES vii

LIST OF FIGURES viii

CHAPTER 1 1

Problem 5

Purpose of the Study 6

Justification for the Study 7

Theoretical Framework 7

Jeanne Chall's Stages of Reading Development 8

The Fourth Grade Slump 12

The Matthew Effect 12

The National Council of the Teachers of Mathematics Principles and Standards 14

Research Questions 16

Assumptions and Limitations 17

Definition of Terms 19

Summary 20

CHAPTER 2: LITERATURE REVIEW 22

Socioeconomic Status 22

Race 28

Gender 31

Retention 34

Teacher Degree 37

Teacher Experience 39

Teacher Professional Development 42

Title I Designation of School 45

Preschool Attendance 48

Norm-referenced and Criterion-referenced Tests 53

Summary 59

CHAPTER 3: METHODOLOGY

Research Design 60

Sample 61

Instrumentation 65

LEAP 21 65

Iowa Test of Basic Skills-Form M 66

Developmental Reading Assessment Test (DRA) 67

Survey of Pre-First Grade Experience 68

Survey of Fourth Grade Teachers 68

Procedural Details 68

Data Collection 69

Pilot Study 71

Data Analysis 71

CHAPTER 4: RESULTS

Descriptive Results 74

Bivariate Relationships 76

Stepwise Linear Multiple Regression Analysis. 80

Summary 89

CHAPTER 5: CONCLUSIONS. IMPLICATIONS, AND RECOMMENDATIONS

Conclusions 91

Implications 99

Recommendations Based on Findings 100

Recommendations for Further Study 101

REFERENCES 102

APPENDIXES

Appendix A: Survey of Pre-First Grade Experience 116

Appendix B: Participant Consent Form 117

Appendix C: Directions for Administering the Survey of Pre-First Grade Experience 118

Appendix D: Survey of Fourth Grade Teachers 119

LIST OF TABLES

1. Jeanne Chall's Stages of Reading Development 10
2. Comparison of LEAP 21 and NAEP Reading Scores 58
3. State and District LEAP 21 Scores for 2000 62
4. State and District Accountability Results for Elementary Schools in 2001 ... 63
5. District Accountability Summary with Louisiana State Averages 64
6. Descriptives for Selected Variables 74
7. Frequency Distributions for Dichotomous and Discrete Variables 75
8. Pearson Correlations Between Dichotomous Variables and LEAP 21 Scores. .78
9. Kendall's Tau-c Associations Between Ordered Variables and LEAP 21 Scores 79
10. Regression Analysis of Predictor Variables and LEAP 21 ELA Scores 80
11. Model Summary for English/Language Arts LEAP 21 Scores 82
12. Model Summary for Mathematics LEAP 21 Scores 83
13. Regression Analysis of Predictor Variables and LEAP 21 Mathematics Scores 84
14. Model Summary for Combined ELA and Mathematics LEAP 21 Scores ... 85
15. Regression Analysis of Predictor Variables and Combined ELA and LEAP 21 Mathematics Scores 86

LIST OF FIGURES

Figure 1. Interrelationship of Variables 8

Figure 2. Revised Interrelationship of Variables 99

CHAPTER 1

INTRODUCTION AND PROBLEM

Increasing attention has been paid to the quality of education in the United States as international reports compare the academic achievement of this country with others in the world. Education has been seen as the key to having a competitive edge in today's global economy, and most employers have been dissatisfied with the level of basic skills of young people entering the work force (Public Agenda, 2002b). Political and social pressure has been applied to hold the educational system more accountable for the academic progress of all students (Olson, 2001).

This trend has been in place for several decades, beginning with the launch of the Russian satellite, Sputnik, when the federal government began pouring millions of dollars into education. States used nationally norm-referenced tests such as the California Achievement Test, the Iowa Test of Basic Skills (ITBS), and the Comprehensive Test of Basic Skills to compare the achievement levels of students in their own states with students across the country. These tests were usually given in timed, multiple-choice format. Results were used to rank students for identification for gifted programs and Title I compensatory programs, since norm-referenced tests compare scores with those of a norm group and rank them accordingly.

The equity of educational opportunities also came into question in the 1960s. Coleman's study (1966) served as a starting point for much debate about racial and economic issues related to the quality of education available to various subgroups of students. The passage of Title IX federal legislation introduced issues of gender equity.

During the 1970s, behaviorists proclaimed that stating educational goals in terms of measurable objectives and then measuring goal attainment with criterion-referenced tests would enable educators to determine whether an appropriate level of education had been achieved. Stories of young people graduating from high school unable to read precipitated the additional requirement by many states of a minimum competency exam before graduation. These criterion-referenced tests were mainly pass/fail for only the very basic skills deemed necessary by each state (Council of Chief State School Officers, 1996, p. 25). Also at this time, a national criterion-referenced test, the National Assessment of Educational Progress (NAEP), was implemented as a means for the federal government to obtain comparative data of educational achievement across the states. The NAEP test introduced a performance assessment component by requiring students to write an essay from a given prompt (Council of Chief State School Officers, 1996, p. 23).

As a result of the *Goals 2000: Educate America Act* of 1984 mandate for all states receiving federal funding for education, many states became involved with developing statewide assessments that included varying degrees of norm-referenced, criterion-referenced, and performance assessment components. This act required that states set challenging standards for all students, write state improvement plans, and provide evidence on how such plans would include students with disabilities (Erickson, et al.,

1995). The 1995 Improving America's Schools Act (IASA) replaced its requirement of using norm-referenced tests to monitor individual student progress with a requirement for "comprehensive assessment systems" capable of reporting "adequate yearly progress" for both individual students and school programs (Council of Chief State School Officers, 1996).

The state of Louisiana became involved in these significant reform efforts in 1993 (Louisiana Department of Education, 2000a), beginning with setting state standards. In May 1997, the State Board of Elementary and Secondary Education (BESE) approved content standards in English Language Arts (ELA), mathematics, science, social studies, foreign languages, and the arts. These standards were clustered in groups of more specific benchmarks for grades K-4, 5-8, and 9-12 to fulfill the Goals 2000 mandate for three accountability groups. The Fordham report, *The State of State Standards*, gave Louisiana a C- for these standards, which placed it twelfth among the fifty states (Finn, Petrilli, & Vanourek, 1998).

In 1996, the BESE initiated the development of the criterion-referenced testing program that aligns with these standards (Advanced Systems in Measurement & Evaluation, 1999). This testing program was called the Louisiana Educational Assessment Program for the 21st Century, or LEAP 21, and was scheduled to be completely phased-in by the school year 2001-2002. The first phase included the development of the Mathematics and ELA tests for fourth and eighth grades. Phase II added the Science and Social Studies components to those grades. The final phases included a requirement for 10th graders to take Mathematics and ELA tests in the spring

of 2001 and a requirement of a passing score on either the Science or Social Studies components of the Graduate Exit Examination 21 (GEE 21) for 11th graders in 2002.

The development of the LEAP 21 at fourth and eighth grades and the GEE 21 at 10th and 11th grades are part of Louisiana's total accountability program, which has the goal of educational improvement (Louisiana Department of Education, 1999). Students are required to pass the Mathematics and ELA components of the LEAP 21 in order to be promoted to the next grades and to pass the GEE 21 in order to graduate from high school. Students in third, sixth, seventh, and ninth grades are required to take the ITBS. Only the LEAP 21 and GEE 21 are high-stakes for individual students. However, each school is held accountable for all of these tests because the schools can be placed in various levels of corrective action if they do not make adequate progress.

Scores from the tests are combined in a mathematical formula with attendance and dropout rates to give individual schools a School Performance Score (SPS) ranging from 0-100 and beyond. The LEAP 21 and GEE 21 account for 60% of the total SPS while the ITBS composes only 30%. The School Performance Score assigns each school a performance label of *Academically Unacceptable*, *Academically Below Average*, *Academically Above Average*, *School of Academic Achievement*, *School of Academic Distinction*, or *School of Academic Excellence*. These performance scores are compared to a 10-year goal of 100, about the current national average of the NAEP, for every school. Louisiana set its 20-year goal at 150, which is the average of the nation's higher performing states. The difference between the School Performance Score and the 10-year goal is divided by the number of 2-year cycles remaining to meet that goal. Each school is expected to improve by that Growth Target. Based on their progress with this growth,

schools received Growth Labels in the fall of 2001 of *Exemplary Academic Growth*, *Recognized Academic Growth*, *Minimal Academic Growth*, or *School in Decline*. A Performance Label of *Academically Unacceptable* and Growth Label of *School in Decline* trigger one of three levels of corrective action by the state. District Assistance Teams of highly trained Distinguished Educators are assigned to those schools to help them improve or be reconstituted. Thus the criterion-referenced LEAP 21 and GEE 21 are high stakes for both individual students and schools.

Problem

Since students in the fourth and eighth grades have been required to score at least *Approaching Basic* on both the ELA and the Mathematics portions of the LEAP 21 in order to pass to the next grade, the state has addressed the issue of identifying students who may be in danger of receiving an *Unsatisfactory* score in either ELA or Mathematics. These students have been provided the LEAP 21 Tutoring Program in order to increase their chances for success (Louisiana Department of Education, 2001b). In this program, students are tutored at least two hours per week with no more than five students per tutor. Following state mandates, the tutoring is to take place before school, during the school day, or after school. Curriculum for the tutoring sessions is planned at the district level, but the tutors at the fourth grade level and tutors at the eighth grade level must have at least a high school diploma and at least two years of college. The state's required measurable objectives for each school or school system who participates are:

1. At least 80% of the eligible students will participate in the LEAP 21 Tutoring Program.

2. Of those students who participated in the English Language Arts component of the LEAP 21 Tutoring Program, 50% will score at or above the *Approaching Basic* achievement level on the English Language Arts component of the spring LEAP 21.
3. Of those students who participated in the Mathematics component of the LEAP 21 Tutoring Program, 40% will score at or above the *Approaching Basic* achievement level on the Mathematics component of the spring LEAP 21 (Louisiana Department of Education, 2001, Part III: Program Description, C).

Criteria for selection to participate in the LEAP 21 Tutoring Program include having been retained in the fourth or eighth grades because of an *Unsatisfactory* score in either ELA or Mathematics on the previous year's LEAP 21 or having a composite score at or below the 30th percentile on the third grade ITBS. Data are needed to determine if these criteria do indeed best identify students who might experience difficulty passing the LEAP 21.

Purpose of the Study

The first purpose of this study was to determine whether the two above-mentioned criteria (ITBS score or repetition of fourth grade) currently selected for student inclusion in the LEAP 21 Tutoring Program were indeed predictive of students' scores on either the ELA or Mathematics portions of the LEAP 21. The second purpose was to explore the relationship among other student characteristic variables (students' Developmental Reading Assessment [DRA] score, socioeconomic status [SES], race, gender, and preschool attendance) together with school characteristic variables (teacher degree level,

teacher experience, teacher professional development participation, and school Title I status) and students' scores on the LEAP 21. Finally, this study proposed to determine which combination of these variables had the strongest predictive value for academic success as measured by the ELA and Mathematics portions of the LEAP 21.

Justification for the Study

In the current LEAP 21 Tutoring Program, considerable personnel, financial, and instructional time resources are expended. Students and schools are under a great deal of individual and collective pressure to demonstrate optimal levels of academic achievement. Therefore, proper identification of students who would benefit most from educational interventions is imperative. To exclude students who are not currently identified and to include students not in need of special services both miss the intended purpose of the program and misuse valuable educational resources. The high-stakes nature of this test makes proper intervention for students at risk of failing the LEAP 21 critical to the state of Louisiana's goal for educational improvement. This study provides evidence about the validity of the method currently in place to select these students. It also serves as a starting place to determine which additional student and school characteristics best predict students' scores on the LEAP 21 at the fourth grade level.

Theoretical Framework

The researcher used an ex post facto design to explore the relationships among student and school characteristics and academic achievement as measured by the LEAP 21. A model for this relationship is illustrated in Figure 1. Two of Jeanne Chall's theories (her stages of reading development and her concept of the fourth grade slump), the

Matthew effect of the rich getting richer and the poor getting poorer, and the principles and standards of the National Council of Teachers of Mathematics served as the theoretical bases for this study.

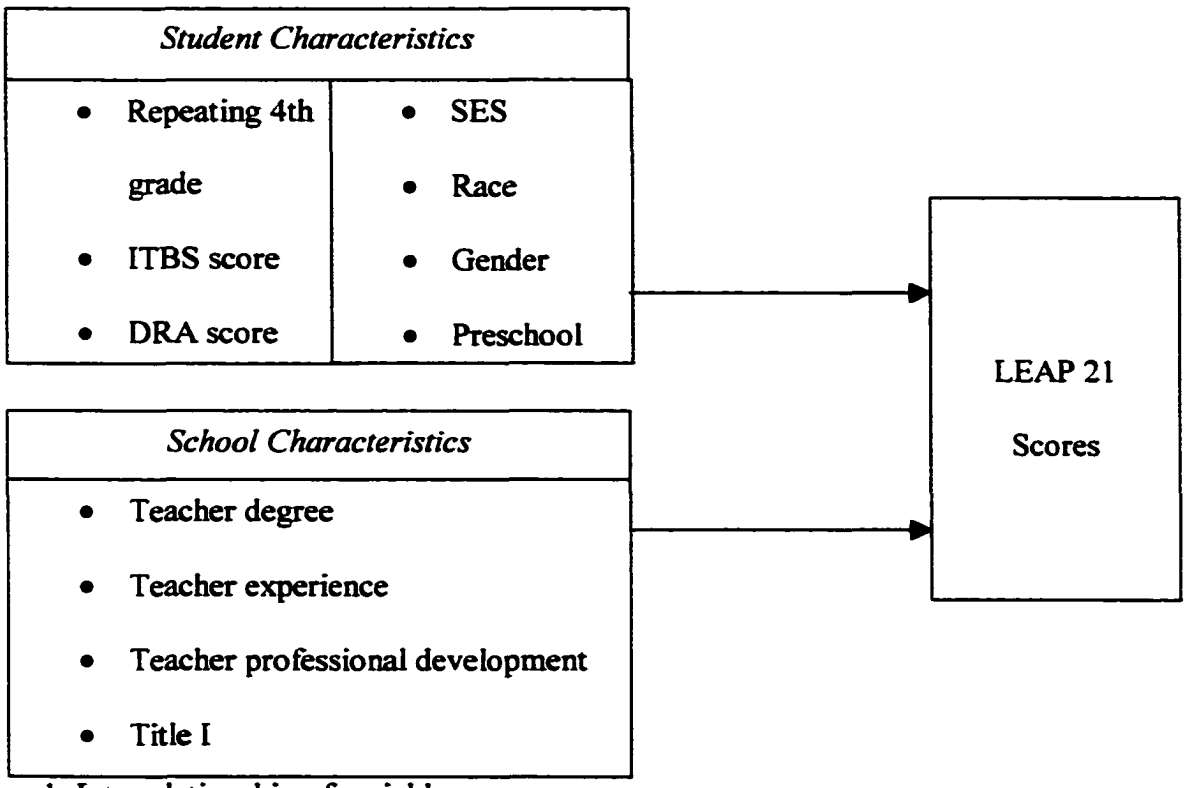


Figure 1. Interrelationship of variables.

Jeanne Chall's stages of reading development. The emphasis in reading instruction over the years has swung between a concentration on the decoding aspect of reading (a phonics/linguistic approach) and a concentration on meaning in context (a whole language approach). Many teachers take a balanced approach, choosing activities for large and small group settings that enhance children's ability both to decode written symbols and to gain meaning and insight from written material (Zemelman, Daniels, & Bizar, 1999). Chall's Stages of Reading Development Theory (1983), which states that people go through developmental stages while learning to read, incorporates both aspects

of reading into a sequential process that leads to a variety of levels of reading ability (see Table 1). Children must have a beginning understanding of language, linguistic patterns, and phonetic sounds—a set of literacy schemata—as a basis for reading literacy (Baker, 1995).

During the first stage, the reader matches an arbitrary set of symbols to his or her inherent literary schemata to decode written language for meaning. Fluency and word recognition accuracy must be developed in the second stage or the student becomes at risk for failure to attain full literacy (National Research Council, 1998). The complexity of factors involved in these first two stages—acquired proficiency in language, verbal memory, lexical and syntactic skills, overall language skills, phonological awareness, acquired knowledge of literacy, reading readiness, letter identification, concepts of print, and verbal and overall IQ—allow for children to learn to read in spite of problems in one area. However, multiple risk areas can be difficult for children to overcome (National Research Council, 1998). In any case the awareness that reading is being done must fade into the background so that the reasons for reading can be fulfilled.

Transition to Stage 3 reading is critical in the ability to be able to learn by reading. As the International Reading Association posits in its *Children's Right to Excellent Reading Instruction*: "Children have a right to reading instruction that builds both the skill and the desire to read increasingly complex materials" (2000, Right 2). This includes such abilities as synthesizing information from various sources, recognizing how a text is organized and using that organization as a tool for learning, judging the reader's own understanding, and evaluating authors' ideas and perspectives. The study *How*

Table 1. *Jeanne Chall's Stages of Reading Development*

Stage	Grade Range	Major Qualitative Characteristics and Masteries
Stage 1: Initial reading or decoding	Grades 1 & 2 (Ages 6 & 7)	Child learns the arbitrary set of letters and associates these with the corresponding parts of spoken words, why different letter arrangements make different words, and how to know when a mistake is made.
Stage 2: Confirmation and fluency	Grades 2 & 3 (Ages 7 & 8)	Child reads simple, familiar stories and selections with increasing fluency. This is done by consolidating the basic coding elements, sight vocabulary, and meaning context in the reading of familiar stories and selections.
Stage 3: Reading for learning new ideas	Grades 4-8 (Ages 9-13)	Reading is used to learn new ideas, to gain new knowledge, to experience new feelings, and to learn new attitudes, generally from one viewpoint.
Stage 4: Multiple viewpoints	High School Grades 10-12 (Ages 15-17)	Reading widely from a broad range of complex materials, both expository and narrative, with a variety of viewpoints.
Stage 5: Construction and reconstruction	College and beyond (Age 18+)	Reading is used for one' own needs and purposes (profession and personal); reading serves to integrate one's knowledge with that of others, to synthesize it and to create new knowledge. It is rapid and efficient.

Note: From *Stages of Reading Development* by Jeanne S. Chall, pp. 86-87. Copyright 1983 by McGraw-Hill Book Co., New York.

People Learn (National Research Council, 1999a) found that one distinguishing characteristic between experts and novices was automatic and fluent retrieval of relevant knowledge. Experts' pattern recognition triggers conditions for accessing knowledge that is relevant to the task. At Stage 3, readers begin to engage in reading contexts that will enable them to abstract relevant features of concepts and develop new knowledge. Students who do not have the decoding automaticity and fluency necessary for this transition undergo what Chall (1983, p. 67) has termed the "fourth grade slump."

Before World War II, a fifth grade reading level was the criterion for functional literacy (Chall, 1983). The army designed basic literacy training for recruits below this level. Since that time, many literacy programs have increased that criterion to eighth grade level, the end of Stage 3. Stage 4 reading includes dealing with multiple viewpoints and layers of facts and concepts added on to those acquired earlier. This level of ability is developed through work with reference works, original and other sources, secondary textbooks, and the free reading of books, newspapers, and magazines. At this stage, the reader becomes concerned with more than one set of facts, various theories, and multiple representations to understand new concepts and points of view.

Readers at Stage 5 development have achieved versatility in suiting reading styles to a variety of purposes and ranges of reading material. The reader knows what not to read, as well as what to read. From reading what others say, the reader uses analysis, synthesis, and judgment to construct knowledge for himself or herself at a high level of abstraction and generality. This level is more a qualitative rather than quantitative approach to knowledge. One creates one's own "truth" from the "truth" of others.

The fourth grade slump. Fourth grade has traditionally been the grade in which the systematic study of subjects such as science and social studies is introduced, so students at this age must depend more on acquiring knowledge from texts than they did in the primary grades (Chall, 1983). Besides decoding ability and fluency, students must develop a more sophisticated vocabulary and knowledge of syntax. In a recent survey of 126 primary-grade teachers (Yopp & Yopp, 2000), only 14% of the materials teachers reported reading aloud on any given day were informational in nature. Students entering fourth grade with few experiences in reading nonfiction, who may also be in an at-risk population that did not allow for sufficient development of background understandings, can experience a slump in reading achievement.

As a result of this phenomenon, the United States has ranked poorly in international reading comparisons. Results of the Organization for Economic Cooperation and Development's Program for International Student Assessment of 15-year olds in reading (Hoff, 2001) showed the United States with a score of 504, just barely above the 32-nation average of 500. Reading experts say that this supports the fact that reading basics are competently taught in the primary grades but are not refined and extended in complexity. This study did show an improved ranking for the United States in Mathematics just below the middle. The Third International Mathematics and Science Study had shown a decline, so the fourth grade slump may be slowing or reversing in mathematics (Hoff, 2001).

The Matthew effect. Originally, Merton (1968) applied this phrase to the reward system in science whereby famous scientists become even more famous because others want to associate with them in their work, while unknown scientists who may do equally

valid work do not gain recognition. The Bible states in Matthew 25:29 (King James Version), "For unto everyone that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath." Stanovich (1986) first applied this to reading when he pointed out that children who begin reading with good vocabularies increase their vocabularies by being good readers. This leads to increased comprehension in a reciprocal relationship. Children who struggle with reading do not read as frequently or with as much understanding and thereby gain less from the reciprocal vocabulary/comprehension relationship. In a study on derivational morphology, Moats and Smith (1992) also concluded that adults who read poorly have a cumulative deficit in their word store.

Several studies have been undertaken to verify this phenomenon. Juel (1988) focused on the literacy development of 54 children in first through fourth grades and found that the probability that a child would remain a poor reader in fourth grade if he or she was a poor reader in first grade was .88. Cunningham and Stanovich (1997) looked at an even longer time frame and compared first grade scores with 11th grade outcomes. Multiple measures were used at both grades. They found that early reading acquisition in first grade could predict 11th grade cognitive outcomes, even after 11th grade comprehension ability is partialed out.

Still other researchers have tried to determine which particular factors might contribute to the Matthew effect. A study of the Title I program (Pogrow, 1999), which the federal government has instituted to ameliorate the effect of poverty on the achievement gap, found that this intervention has made some difference in the early grades. However, failure to sustain early gains or to continue narrowing the gap after the

early grades is particularly evident in high poverty schools. "Simply put, grades 4-8 are the black hole of American education which seems to suck in whatever progress has been made" (Pogrow, 1999, ¶ 7).

The California Budget Project (2001) found another situation in which the rich get richer and the poor get poorer in the California educational accountability system. Those schools in the lowest two deciles based on test score results received $1,141 per point improvement, while those in the highest two deciles received $1,423. Since schools with the highest percentage of poor students are the most likely to be low-performing, the poor get poorer in California's accountability system, and that of many others. Louisiana's system of rewards and consequences is also tied to achievement in reading and mathematics, so applying valid theories of learning is important in both areas.

The National Council of Teachers of Mathematics Principles and Standards. In 1989, the National Council of Teachers of Mathematics (NCTM) created a system of principles and standards that became a model for many of the other national curricular standards that followed (2000). The six principles for school mathematics, added in 2000, address the following overarching themes:

> Equity. Excellence in mathematics education requires equity—high expectations and strong support for all students.
>
> Curriculum. A curriculum is more than a collection of activities: it must be coherent, focused on important mathematics, and well articulated across the grades.

- Teaching. Effective mathematics teaching requires understanding what students know and need to learn and then challenging and supporting them to learn it well.
- Learning. Students must learn mathematics with understanding, actively building new knowledge from experience and prior knowledge.
- Assessment. Assessment should support the learning of important mathematics and furnish useful information to both teachers and students.
- Technology. Technology is essential in teaching and learning mathematics; it influences the mathematics that is taught and enhances students' learning. (chap. 2, ¶ 2)

In classrooms that provide the NCTM envisioned environment:

> teachers provide students with numerous opportunities to solve complex and interesting problems; to read, write, and discuss mathematics; and to formulate and test the validity of personally constructed mathematical ideas so that they can draw their own conclusions. Students use demonstrations, drawings, and real-world objects—as well as formal mathematical and logical arguments—to convince themselves and their peers of the validity of their solutions (Battista, 1999, How Mathematics is Taught, ¶ 2).

Thus, teachers' degree levels, years of experience, and professional development have a potential influence on creating this environment.

The reality is that the NCTM vision has not been happening in schools in the United States. Wiley and Yoon (1995) found that 31% of California 10th grade teachers of mathematics were familiar with NCTM standards, while only 3.95% in fourth grade

and 16.09% in eighth grade were familiar with the standards. An analysis of questions asked of students taking the 1996 NAEP mathematics test (Steen, 1997) revealed that only about a fifth of Louisiana students were asked to write about mathematics solutions and only about a fifth were asked to discuss their solutions with other students every day. Transition from novice to expert involves not just mastering an objective, but rather making connections among objectives (National Research Council, 1999a). The content of U.S. mathematics classes requires less high-level thought than classes in Germany and Japan (U.S. Department of Education, 1996b). A typical U.S. teacher's goal is to teach how to do something rather than to understand mathematical concepts. Cohesive implementation of the NCTM Principles and Standards is not yet a reality in all American schools.

Research Questions

Theoretically, the ideal educational environment should have enabled all students to meet academic standards that have been set forth as the state goal. However, experience has taught that all students do not demonstrate mastery on the statewide assessment associated with these standards. Therefore, it was necessary to identify students most at risk for failing to demonstrate mastery and provide intervening strategies that maximize their potential success. The questions that this study addressed were:

1. Are the current criteria for being categorized as at-risk for participation in the LEAP 21 Tutoring Program (grade retention in fourth grade or 30th percentile or below on the third grade ITBS) actually related to LEAP 21 scores?
2. To what extent are other student characteristics (Developmental Reading Assessment [DRA] score, socioeconomic status, race, gender, and preschool attendance) and

school characteristic (teacher degree level, teacher experience, teacher participation in professional development, and school Title I status) also related to student scores on the LEAP 21?

3. What combination of variables has the strongest predictive value?

These questions were considered in the context of the model of student and school characteristics previously illustrated in Figure 1. These variables were identified in many previous studies as being related to academic achievement and their relationship is further discussed in the review of literature. The above referenced model of student and school characteristics as potential predictor variables related to the ELA and Mathematics scores on the LEAP 21 was used as the basis for the following research hypothesis: Both student characteristics (students' grade retention in fourth grade, core total national percentile on the third-grade ITBS, final third-grade DRA scores, SES, race, gender, and preschool attendance) and school characteristics (teachers' degree level, teachers' experience, teacher's participation in professional development, and students' school Title I status) are related to LEAP 21 scores.

This hypothesis was applied to three outcome variables—ELA LEAP 21 scores only, Mathematics LEAP 21 scores only, and a combined total of the ELA and Mathematics LEAP 21 scores—because this gave more precise information for different weightings of ELA and Mathematics LEAP 21 scores.

Assumptions and Limitations

The researcher assumed that the 2002 LEAP 21, the ITBS, and the DRA were administered as instructed, and that all responses on the Pre-First Grade Experience Survey were accurate. This study included all fourth grade students in a particular school

district, both those who participated in the LEAP 21 Tutoring Program and those who did not. Some of them were identified at the beginning of their fourth grade year as being qualified to participate in this program. A major portion of the students who have been identified as being at or below the 30th percentile composite score on the ITBS or who have failed either the Mathematics or the ELA portion of the LEAP 21 have received a pullout tutoring intervention in both subjects in the LEAP 21 Tutoring Program. Therefore, this study also assumed that this intervention had little impact on the outcome of these students' scores on the 2002 LEAP 21. This view is corroborated in the review of literature under the Title I section (Mullin & Summers, 1983; Puma. 1999). Because of the LEAP 21 Tutoring Program intervention, the study could possibly be skewed to show a weaker correlation between the lower than 30th national percentile core total score on the ITBS or fourth grade retention than existed without this intervention among those students who participated in the LEAP 21 Tutoring Program. The use of national percentile equivalent scores for students' core totals of the ITBS may also limit the internal validity of the results.

The use of only one grade in only one school district limits the generalizability of this study. However, the reader may use the data in the Sample section of Chapter 3 that compares the selected school district with Louisiana state averages to further evaluate the extent to which the results of this study may be generalized to the state as a whole.

Scoring errors on any of the assessment instruments may also be a limitation. The impact of the high-stakes nature of the LEAP 21 may also influence the students' test taking ability due to stress or other emotional or physical constraints. Professional

development programs vary in length and by subject. Additional professional development opportunities may have been omitted that should have been included.

Definition of Terms

Grade retention: Students were classified as being enrolled in fourth grade for the first time or as being enrolled in the fourth grade for the second time. No students were enrolled in fourth grade for the third time.

ITBS score: The student's composite ITBS national percentile score was used.

Developmental Reading Assessment (DRA) score: This score was the one recorded on the form turned in to the state at the end of the students' third grade year identifying the student as below grade level, on grade level, or above grade level in reading.

Socioeconomic status (SES): Students who qualify for the free- or reduced-federal lunch program were considered low SES, and students who do not qualify were considered high SES.

Race: Students were classified as either White or African American as recorded in the Louisiana Student Information System. Since Hispanic and Asian American students comprised less than 1% of the students in this study, they were not included.

Gender: Students were classified male or female as recorded in the Louisiana Student Information System.

Preschool attendance: Students were identified as having attended preschool if any other responses were selected on the Survey of Pre-First Grade Experience than the last one on each of items three, four, and five.

Teachers' degree: Teachers with temporary certificates were coded as *0*, teachers with bachelor's degrees were coded as *1*, teachers with master's degrees were coded as *2*, and teachers with 30 graduate credits beyond the master's degree were coded as *3*.

Teachers' experience: Number of years of experience in the school district archives was recorded.

Professional development: Teachers who were participants in the state-sponsored Intech and Louisiana Systemic Initiative Program professional development experiences or in the Math Their Way, Math A Way of Thinking, or Balanced Literacy district-sponsored professional development opportunities were recorded as having this predictor variable. These professional development opportunities were selected because they involved a minimum of 20 hours of formal instruction, included follow-up commitments during the school year, and were available to all teachers in the study. They also were free of charge or included a financial or material stipend.

Title I school: Any school receiving school-wide or targeted assistance Title I funds qualified for this designation.

LEAP 21 score: Students' numeric scores on the English Language Arts portion of the LEAP 21 and students' numeric scores on the Mathematics portion of the LEAP 21 qualified for this score.

LEAP 21 Tutoring Program: The tutoring program prescribed by the state of Louisiana specifically for defined students at-risk of failing the LEAP 21.

Summary

Because students who receive an *Unsatisfactory* score on either the English Language Arts (ELA) or the Mathematics portions of the Louisiana Educational

Assessment Program for the 21st Century (LEAP 21) must repeat fourth grade, the state has instituted the LEAP 21 Tutoring Program, which is conducted during the fourth grade school year. Criteria for inclusion in this program include either having repeated the fourth grade due to an *Unsatisfactory* score on the previous year's LEAP 21 on either ELA or Mathematics or achieving a composite score at the 30th percentile or below on the third grade Iowa Test of Basic Skills (ITBS). This study determined if repeating fourth grade and third grade ITBS scores were, indeed, predictive of the LEAP 21 achievement score, or whether other criteria such as socioeconomic status (SES), race, gender, Developmental Reading Assessment (DRA) score, preschool attendance, teacher degree level, teacher experience, teacher professional development, and school Title I status were more predictive. A further purpose was to determine which combination of variables had the highest predictive value for better identification of an at-risk population of students.

CHAPTER 2

LITERATURE REVIEW

Socioeconomic Status

Since Coleman's landmark study on *Equality of Educational Opportunity* (1966), socioeconomic status has been seen as a strong predictor of student achievement. Coleman asserted that the influence of student background was greater than anything that goes on within schools. Poverty is indeed a factor among children in the United States. Rainwater and Smeeding (1995), in their 18 nation *Luxembourg Income Study,* found that during the 1990s families of children in the United States had lower real income than families of children in almost every other nation. Although the poverty rate for people under 18 years old dropped from 16.9% in 1999 to 16.2% in 2000 (U.S. Census Bureau, 2001), American children remained the poorest population by age group. Of these approximately 12 million children, one third live in extreme poverty in families with incomes below 50% of the poverty line. The child poverty rate in Louisiana is the second highest in the United States (Hoff, 2002) and the highest in the South at 29% (Bennett & Lu, 2000). The per capita personal income in Louisiana in 1998 was only 82% of the U.S. average, and that of the school district in this study was only 69% of the United States' average (Center for Business & Economic Research, n.d.). This means that another large

portion of children was very near the poverty level, adding to the total number of children in economic constraints at the time this study was conducted.

The issue of socioeconomic status and its relationship to student achievement is more complex than Coleman's (1966) report first intimated. First of all, the relationship can be explored on various unit levels, from that of nations and states, districts, and schools, and on to classes and individual students. Payne and Biddle (1999) commented in their study of data obtained from the Second International Mathematics Study (SIMS) that if the United States had been represented only by its school districts with low-level poverty, the United States would have ranked second out of the 23 nations involved. If the high-poverty district scores were used, the United States would have ranked only above Nigeria and Swaziland. Findings from Binkley and Williams' (1996) study of the International Association for the Evaluation of Educational Achievement (IEA) Reading Literacy Study supported a somewhat similar comparison between poverty and reading literacy. The low-poverty fourth-grade group in the United States faired better than any group in the 32 other countries. The high-poverty group scored much lower than the low-poverty group, but never fell below the international average. Since the IEA assessment measured only a basic comprehension level, low socioeconomic status was not as strong a detriment to U.S. students in an international reading comparison as that shown by the study of the SIMS higher level mathematics assessment.

At the national level, Chall (1996) analyzed a combination of NAEP reading results, Scholastic Aptitude Test scores over time, and a synthesis of research on beginning reading from 1910 to 1996. She also concluded that there are large differences

between higher- and lower-socioeconomic status children. The differences were smaller among younger children and increased in the higher grades.

Using the 1996 NAEP data for state-level mathematics achievement and for state-level poverty and *Education Week's* 1997 edition of *Quality Counts* for state-level funding of education, Biddle (1997) concluded that the child poverty/achievement correlation was $r = .700$ ($p < .001$) and that, together, school funding and child poverty predict 55% of the variance of state differences in mathematics achievement. The impact of child poverty was stronger at the state than the district level. Darling-Hammond (1999) also used NAEP data at the state level, this time from two years of fourth-grade mathematics results, two years of eighth-grade mathematics results, and two years of fourth-grade reading results. She also concluded that poverty was significantly and negatively correlated with student outcomes at the state level.

In a district level study of urban schools belonging to the Council of Great City Schools (2001), the results of the Stanford Achievement Test indicated that the greater the concentration of poverty in the school districts, the lower the student achievement. Of the three grades—4th, 8th, and 10th—selected to report poverty data, achievement gaps between districts of high and moderate concentrations of poverty were generally greatest in fourth grade in both reading and mathematics. Caldas (1999) compiled the results of all Louisiana 10th graders in 1990 who took the Louisiana Graduation Exit Examination to correlate both district- and school-level effects of poverty on achievement. SES accounted for 45.5% of the variation between districts and 41% of the variation among schools within districts. However, Caldas discovered that the percentage of one-parent families accounted for 96% of the variation in average school test scores among districts

and for 59% of the variation among schools within districts. He further found that even if a student came from a two-parent family, the domination of a school or district by one-parent families could have an overriding negative influence stronger than that of poverty or race.

In a study of West Virginia districts and schools in grades 3, 6, 9, and 11, Howley (1995) found a weaker level of correlation between SES and achievement at these levels. Additional analysis revealed that the smaller class sizes in most West Virginia schools tended to ameliorate the negative effects of poverty. The Matthew Project (Howley & Bickel, 1999) extended this study to four additional states: Ohio, Georgia, Texas, and Montana. The additional findings further supported the benefits of smaller class sizes for impoverished communities and the benefits of larger classes for more affluent communities. This was most evident at the school level.

The U.S. Department of Education conducted *The Longitudinal Evaluation of School Change and Performance (LESCP) in Title I Schools* (2001a) to determine the effectiveness of Title I schools. Key findings were that individual and school poverty had a clear, negative effect on student achievement and that students who attended schools with the highest percentages of poor students performed worse initially on both reading and mathematics tests. Gaps in reading remained the same from third to fifth grades, but gaps in mathematics partially closed. Teacher effects made the difference in this study and in another one by Fetler (1999) of California state high schools. Again, however, poverty had a strong relationship to achievement at the school level. Analysis of Stanford 9 scores in reading and mathematics from 2,000 fifth graders in Texas (Klein, Hamilton, McCaffrey, & Stecher, 2000) also showed a strong negative correlation at the school

level. The percentage of students at a school who were in the federal free- and reduced-lunch program predicted that school's mean on the test regardless of test type, multiple choice or open-ended. Sander (2001) compared Chicago schools with those in the rest of Illinois. Again, the low-income students had lower achievement, but Chicago grade schools were just as efficient as the others in teaching reading and mathematics after factoring out family background. Reading scores became significantly lower for impoverished students at the high school level, consistent with Chall's (1996) findings.

The strength of the district and school level influences of socioeconomic status on academic achievement is evident in a growing movement to integrate school districts on the basis of equitable economic status rather than on racial equity. LaCrosse, Wisconsin, was the first (Kahlenberg, 1999), followed by others including San Francisco, California, and most recently, Cambridge, North Carolina (Richard, 2002).

A study of more than 6,000 fourth-grade classrooms in Texas (Lopez, 1995) revealed that low SES classrooms had significantly lower gains on the Norm-referenced Assessment Program of Texas than non-low SES classrooms. At the classroom level, however, teacher factors influenced student achievement causing greater variance. Poverty played a significant role in the print environment and experience of students in first-grade classrooms in the greater Boston area (Duke, 2000). Poor classes had books and magazines, less print on the walls and other surfaces, less exposure to and experience with extended text, and less time engaged in activities in which students had a high degree of authorship. The reverse was true of classes with more financial support.

To further illustrate the degree to which individual schools and classes can reduce the effects of poverty on student achievement, The Education Trust (Jerald, 2001)

identified 4,577 schools nationwide that were in the top third of their state in reading or mathematics performance and that had at least 50% low-income or at least 50% minority students compared with other schools at their grade level. Louisiana had 96 of those schools, but none of them were in the school district in this study.

The influence of socioeconomic status at the individual level is still prevalent (Capraro, 2000) but less strong in much of the literature. Entwisle and Alexander (1996), in a study of mother-only, mother-extended family, and two-parent families with children in first through third grades, concluded that two measures of parent expectations had a somewhat stronger influence than did the economic variables. The effects of prior achievement were stronger than poverty on junior high and high school students in a study of data obtained from the Longitudinal Study of American Youth (Brookhart, 1997). Cultural effects of a race and gender interaction for African American males among elementary school-aged children (Diamond & Onwuegbuzie, 2001) were stronger than socioeconomic status in predicting reading achievement. SES became stronger for individuals at the postsecondary level (Trusty, 2000) because more family and individual resources are necessary to attain this level. However, low prior mathematics achievement can have a strong barrier effect as well, regardless of family or individual SES.

In a meta-analysis of socioeconomic status, White (1982) concluded that the utility and wisdom of using SES in conjunction with academic achievement depended largely on the unit of analysis and the validity of the way in which it was defined. This study used operational definitions of SES. First, individual SES was defined by participation in the federal free- and reduced-lunch program to show only a weak correlation with academic achievement at this level. Secondly, enrollment in a Title I

school designated school-level SES because these schools, by qualification, must have 70% or more of their student populations participating in the free- and reduced-lunch program.

Race

The factor of race or ethnicity is closely associated with that of poverty as a predictor of achievement. Harkreader and Weathersby (1998) found its influence much less than economic factors, whereas Bankston and Caldas (1998) concluded that minority status was more highly related to achievement than was socioeconomic status.

Coleman's report (1966) was the basis for the desegregation required in the civil rights acts of the 1960s. As a result, the South became the most highly integrated part of the nation with the most substantial contact between African American and White students (Orfield, 2001). During the 1990s, there were three major Supreme Court decisions authorizing a return to segregated neighborhood schools and limiting the reach and duration of desegregation orders. These decisions took the stance that positive policies taking race into account for the purpose of creating integration were suspect and had to demonstrate both a compelling reason and prove that the goal could not be realized without considering race. The 2000 Census showed a continuing return of African Americans to the South into more racially segregated situations. However, it is still more common for African Americans to attend school with Whites in the South than in any other part of the country (Orfield, 2001).

African American children (33.1%) are more likely to live in poverty than White children (13.5%). They are also more likely to have single parents, and more likely to be welfare dependent (Rector, Johnson, & Fagan, 2001). African American children are also

disproportionately represented in Title I schools (Puma, 2000). Racial minority status is more likely to be correlated with lower teacher qualifications such as certification and years of experience (Darling-Hammond, 1999).

Data collected at a national level have been analyzed in a variety of ways to determine if the achievement gap between White students and racial minority students has narrowed. The National Center for Education Statistics (Jacobson, Olsen, Rice, & Sweetland, 2001) used data from several cohorts of the Chapter I Prospects Study, a study commissioned by congress to evaluate the Title I program, to determine that mathematics and reading scores of African Americans were generally lower than corresponding scores of Whites even with similar levels of prior achievement one or two grades earlier. The gap narrowed during elementary school but widened during junior high school with little change in high school. Phillips, Crouse, and Ralph (1998) used these data plus the National Education Longitudinal Survey (NELS) data and came to the same conclusion about mathematics achievement. However, they determined that race had a stronger effect on reading growth than on mathematics growth, with the biggest gap generated in elementary school.

Grissmer, Kirby, Berends, and Williamson (1994) used data from NELS together with data from the National Longitudinal Survey of Youth and the National Assessment of Educational Progress (NAEP) test to determine that the rising test scores of minorities have resulted in a significant closing of the achievement gap between minority and non-minority youth. However, the gap that remained was still significant. Barron and Koretz (1994) maintained that the small sampling sizes for minorities used in the trend NAEP, one part of the NAEP test specifically designed to track national longitudinal

achievement data, cause unreliable conclusions to be reached in racial comparisons. Standard errors for minorities were twice that of Whites. The study determined that minorities would have to have much larger gains than Whites to achieve significance. The source for recording racial data also has a strong impact on the results. Young elementary students who self-reported race were more likely to respond inconsistently with adults who reported observed racial options. Since NAEP racial data are self-reported, the smaller sampling of minority students might indeed be affected by this inconsistency, particularly at the fourth grade, the lowest grade tested.

The particular achievement test used was also an influencing factor in determining racial gap scores. Klein et al. (2000) reported that the NAEP test showed a gap that was wide to begin with and got wider with time for Texas students. During the same time period, Texas Assessment of Academic Skills showed that the gap started off somewhat smaller and then became substantially smaller over a four-year period. In a study of the Ohio state proficiency test, Dimitrov (1999) determined that the response format, open-ended versus multiple-choice, did not make much difference by ethnicity. Only the low and high ability Hispanics had negative academic relationships to their response strategies for multiple-choice items, not the extended response items one might expect with limited English proficiency. The end-of-grade tests in North Carolina (North Carolina State Department of Public Instruction, 2000) reflected a racial disparity with African Americans, Hispanics, and Native Americans performing well below Multi-Racial, Asian and White groups. However, all groups continued to improve.

The Education Trust (2001) reported a closing of the gap in basic skills in mathematics by race during the 1970s and the 1980s, but the gap remained the same or

widened at higher levels of cognition. In Louisiana, in particular, African American fourth graders made more progress in mathematics from 1992 to 1996 than they did in most other states (Education Trust, 2001). However, achievement gaps remained constant between eighth grade White and African American students from 1992 to 2000 (National Education Goals Panel, 2001).

In reading, the National Center for Education Statistics (Donahue, Voelkl, Campbell, & Mazzeo, 1999) concluded that in 1998 eighth grade African American students achieved a significant gain over their NAEP scores in both 1992 and 1994, while fourth grade African American students achieved a significant gain over their 1994 results. Louisiana African American fourth graders, however, made reading gains five points less than the national average (Education Trust, 2001). At the eighth grade, the gap between Louisiana African American and White students' reading achievement was 27 points. Besides being influential in their own right, racial and ethnic cultural influences may also combine with smaller gender effects to predict achievement.

Gender

Some correlation appears to exist between gender and reading achievement. Disaggregation of the 1998 NAEP reading results by gender rather than race (Donahue et al., 1999) revealed that females outperformed males in 4th, 8th, and 12th grades, as they also did in 1992 and 1994. At the 4th-grade level, however, the males made a significant gain over their 1994 score while the females remained the same. A similar trend was noted in the North Carolina end-of-grade tests administered in grades three through eight (North Carolina State Department of Public Instruction, 2000). A similar phenomenon appears to be occurring in Great Britain as well (Salisbury & Rees, 1999).

Perhaps some of this gender difference can be explained by a national survey of reading attitudes conducted with 18,185 children across the United States in first through third grade (McKenna, Kear, & Ellsworth, 1995). Girls as a group possessed more positive attitudes than boys at all grade levels, both toward recreational and academic reading. These attitudes appeared unrelated to ability. A four-year longitudinal study of elementary school age children in Michigan (Eccles, Wigfield, Harold, & Blumenfeld, 1993) revealed that girls valued reading significantly more than boys and also saw themselves as being more competent readers than boys. Using data from the National Longitudinal Survey of Youth, Baharudin and Luster (1998) found that female children in the overall sample and in the Caucasian subsample appeared to receive more supportive care than male children. These same two groups scored significantly higher than males on reading achievement as well. Effects for gender in reading were seen as early as second grade (Entwisle & Alexander, 1996) and continued through high school (Binkley & Williams, 1996).

Gender as a predictor of mathematics achievement in Baharudin and Luster's study (1998) of six- to eight-year olds emerged again as significant for females in general and for the African American female subgroup. On the NAEP 2000 Mathematics Assessment (U.S. Department of Education, 2001b), however, a higher percentage of boys performed at or above *Proficient* than girls at 4th, 8th, and 12th grades, with the older two grades being significantly higher. The gap between the average scale scores of males and females was quite small at all three grades and has fluctuated only slightly over the past 10 years. There was no significant difference by gender at the fourth-grade level. In Louisiana, neither the scale scores nor the percentage of students scoring at or

above the *Proficient* level was significant for gender at fourth grade. At eighth grade, the difference in scale scores was not significant, but the difference in percentages scoring above the *Proficient* level was positively significant for males.

In an international comparison of Third International Mathematics and Science Study data in English-speaking countries, Webster, Young, and Fisher (1999) determined that in Australia and the United States very little of the student level variance was explained by gender and SES, although most of the variance was at the student level and not at the class level. The U.S. Department of Education's (2000a) analysis of that same data revealed that males outperformed females in 3 of the 25 countries at the fourth-grade level, in 8 of the 39 countries at the eighth-grade level, and in 18 of the 21 countries participating in their final year of secondary school. However, in the United States, males and females scored similarly at all three levels.

Results from an analysis of the National Educational Longitudinal Study of 1988 data (Catsambis, 1994) showed that male and female eighth graders attained similar achievement, but a larger portion of girls were placed in high-ability classes and a larger portion of boys were placed in low-ability classes. Racial/ethnic influences may have played a role in secondary mathematics course selections and judgment of academic performance. The chances of young African American women enrolling in high-ability mathematics classes were 48% greater than those of African American male students. Hispanic females reported lower participation in these classes and higher performance anxiety, while White females had the highest enrollment in high-ability classes. Females in general in this study, and also in Campbell & Beaudry's study (1998) of the

Longitudinal Study of American Youth data, revealed less confidence in their mathematical ability and greater exertion of effort in mathematics classes than males.

Mathematical ways of thinking may differ by gender according to Fennema, Carpenter, Jacobs, Franke, and Levi (1998). These researchers studied 82 children as they progressed from first through third grades. They identified gender differences in strategy use that was evident from the beginning of the study and persisted through the end. Girls tended to use more modeling or counting strategies, while boys tended to use more abstract strategies such as derived facts or invented algorithms. By the third grade, girls used significantly more standard algorithms than did the boys.

In an analysis of the Delaware Student Testing Program and the Stanford Achievement Test Series 9th Edition for students in 3rd, 5th, 8th and 10th grades, Zhang and Manon (2000) found that males had a larger variance in mathematics scores than females. In this study, females tended to outperform males among the low-achieving students and males tended to outperform females among the high-achieving students. This higher variance for males makes them more susceptible to rewards and sanctions in many state accountability systems.

Retention

Gender, race, and socioeconomic status also have a role to play in the issue of retention. McCoy and Reynolds (1998) used data from the Chicago Longitudinal Study of 1,164 low-income, mostly African American 14-year-old students who had all attended a federally funded kindergarten program. Retained children were most likely to be boys and most likely to have lower scores in reading and mathematics achievement. No national or regional agencies monitor grade retention. However, a report for the

National Research Council (1999b) used information from the U.S. Census Bureau to determine that, nationally, sex differential in retention gradually increases with age from five percentage points at ages 6 to 9 to ten percentage points at ages 15 to 17. Rates of retention are racially similar at the younger ages, but by ages 15 to 17 the rate is between 40% to 50% among African Americans and Hispanics, but only 25% to 35% among Whites. Hauser (1999) estimated that at least 15%, and probably 20%, of children have been held back at some time in their lives. In 1998 41% of teachers reported that their schools promoted students based on age, but in 2001, only 31% did so (Johnson, Duffett, Foleno, Foley & Farkas, 2001). The Louisiana Department of Education (2001a) analyzed its Student Information System (SIS) data from 1997-2001 in grades K-12 and found that male students were more likely to be retained than female students, and students on free lunch were twice as likely to be retained as students not receiving any food services. However, African American students receiving reduced lunch had significantly lower retention rates than those on free lunch or those not receiving any food services in 2000-2001.

This same study (Louisiana Department of Education, 2001a) revealed that the number of students retained more than tripled in fourth and eighth grades, reflecting the impact of high stakes testing on retention in the state of Louisiana. With eight states planning to base promotion in some grades on statewide assessment results by 2004 (Edwards, Chronister, & Olson, 2002), this impact is likely to increase. Cizek, Trent, Crandell, Hirsch, and Keene (2000) surveyed teachers and principals of a random stratified sample of fourth-grade students across the state of Ohio to determine if their assessment of students' readiness for fifth grade corresponded with the results of the

Ohio Proficiency Test administered at the end of fourth grade. Educator agreement was high, but varied by district in relation to the standards of the proficiency test. Since the number of students actually retained was considerably less than the number of students deemed unprepared, suggested further research includes discovering additional criteria on which to base decisions related to retention.

The results of retention were decreased academic progress and higher dropout rates. Roderick, Bryk, Jacob, Easton, and Allensworth (1999) conducted an analysis of the implementation of the first two years of the Chicago Public Schools' intensive effort to end social promotion and raise achievement, which began in 1996. Their analysis revealed that only one fourth of retained eighth graders and one third of retained third and sixth graders in 1997 made "normal" progress to pass the test cutoff the next May. Retention was therefore better for some students in the short term. However, the average ITBS score increase in the two years required to repeat a grade was 1.2 grade equivalents compared to 1.5 grade equivalents for students who had similar scores and were promoted prior to policy implementation. The performance of third graders was significantly poorer than that of sixth and eighth graders, indicating that remediation strategies may need to be different for younger children than for older children.

A later follow-up of the Chicago study (Roderick, Nagaoka, Bacon, & Easton, 2000) disclosed some additional negative results of retention. First, despite higher passing rates, retention rates have not fallen. This is due to the fact that, over the three year study, fewer students are being socially promoted as a result of the stricter guidelines for promotion. Secondly, retained students are struggling in their second time to face the promotion policy because they still do not do well in the next tested grade.

Finally, nearly a third of retained eighth graders in 1997 had dropped out by the fall of 1999. However, overall dropout rates were stable. Several positive results were also discovered. Passing rates improved in all three grades, more at-risk sixth and eighth graders are raising their test scores during the school year, and more students maintained positive test trajectories two years after promotion.

Students in the state of Texas were analyzed by Haney (2000), who found that a comparison of the cumulative total of 2.2 million students enrolled in sixth grade between the fall of 1984 and the spring of 1993 and of the cumulative total of 1.5 million graduates in the classes of 1992 and 1999 meant that during that nine year period around 700,000 children were lost or left behind before graduation. Haney attributed this to an increase in retention rates, particularly among African Americans and Hispanics, and an increase in the dropout rate. Only 50% of minority students have been progressing from ninth grade to graduation since the initiation of the Texas Assessment of Academic Skills (TAAS), again reflecting the impact of high stakes testing and accountability. With increased pressure on students to achieve comes increased pressure on teachers' ability to teach them what they need to know and to be able to do.

Teacher Degree

One factor that has been studied as being related to teachers' ability to teach is whether or not that teacher has an advanced degree. The National Center for Education Statistics (Lewis, Parsad, Carey, Bartfai, & Farris, 1999) used its Fast Response Survey System to collect data from a nationally representative sample of full-time public school teachers. Virtually all teachers had bachelor's degrees, and nearly half (45%) had master's degrees. More high school teachers (55%) had master's degrees than elementary

teachers (40%). Regional differences of teachers with master's degrees by percentage were Northeast 60%, Midwest 47%, South 42%, and West 37%. Only 30% of teachers with 4 to 9 years of experience had master's degrees, but 61% of teachers with 20 or more years of experience had master's degrees. Boe and Barkanic's (2000) study of predictors of advanced degrees using the Schools and Staffing Survey indicated that teacher experience was the strongest predictor.

One study that included teacher degree as part of an assessment of the influence of teacher quality on students' mathematics achievement in California (Fetler, 1999) was conducted in 795 high schools serving 1.3 million students. The average number of years of teaching experience was positively related to test scores to the same degree that education level of teachers in the school was negatively related to test scores. However, when Darling-Hammond (1999) used data from the 1993-1994 Schools and Staffing Surveys (SASS) along with NAEP reading and mathematics scores of several years in a regression analysis to determine which teacher quality variables were predictors of academic success, she found that certification status and degree in the field to be taught were very significantly and positively correlated with student outcomes. Education level, defined as percentage of teachers with master's degrees, showed a positive but weaker relationship. Miller-Whitehead (2001), in her study of Alabama's Education Report Card for the year 2000, also discovered a positive correlation between faculty with advanced degrees and student achievement. In addition, she found a negative relationship with teachers who had only bachelor's degrees.

Considering the academic major and minor associated with a teacher's degree along with its status as graduate or undergraduate, Ingersol (2002) found from the SASS

data that the combination of undergraduate status and lack of a major or minor in the field of teaching responsibility was more likely to occur in high poverty and high minority urban schools. Lee (1998) also determined that non-tenured teachers across the state of Maryland tended to be concentrated in higher numbers in either school districts with higher poverty or districts with higher numbers of minority students, and that either of these two conditions served as a predictor of higher teacher flight. This condition was true both between and within school systems.

Hanushek, Kain, and Rivkin (1998) used the extensive Texas database of student and teacher demographic data and of three statewide cohorts of student achievement in multiple elementary grades over three years to determine the boundaries of the effect that various aspect of teacher quality had on student achievement. They concluded that the effects of a master's degree were generally negative and always statistically nonsignificant. Lopez (1995), who used data from 6,000 fourth-grade students and their teachers, found no significant differences in student performance between classrooms with teachers who had bachelor's degrees and classrooms with teachers who had master's degrees. One of the policy recommendations that resulted from this study was to provide training programs for teachers that would extend the maximum potential to affect gains in student achievement.

Teacher Experience

A teacher qualification related to the type of degree held by teachers is the number of years of experience they have in teaching. Utilizing data from the National Center for Education Statistics, the U. S. Department of Education (2000a) determined that the percentage of teachers with master's degrees increased with years of experience,

and that teachers with three or fewer years of teaching experience were more likely than more experienced teachers to hold academic, as opposed to education, degrees. Ingersoll (2002) used data from the 1993-1994 SASS and concluded that just over one 10th of all public school teachers were beginners with 3 years or less of experience and just over one third were seniors with more than 20 years of experience. In disadvantaged schools, twice as many teachers were beginners and fewer teachers were seniors than in advantaged schools. Lee (1998), in his study of Maryland schools, discovered similar findings. Schools of affluent students were taught by experienced teachers while those of low-income and other disadvantaged students were taught by uncertified, beginning, and novice teachers, many of whom harbored negative feelings towards low-income and other disadvantaged students. Data collected in the 1998 Fast Response Survey System (Lewis, Parsad, Carey, Bartfai, & Farris 1999) indicated that emergency and temporary certification was higher among teachers with less experience. Less than 1% of elementary classroom teachers with ten or more years of experience had emergency or temporary certification, but 12% of elementary classroom teachers with three or fewer years of teaching experience had them.

A report to the National Education Goals Panel (Smrekar, Guthrie, Owens, & Sims, 2001) concerning the Department of Defense (DoDEA) system of schools revealed a strong link between teacher quality and student achievement. Only 10% of the DoDEA teachers had two or fewer years experience while over 42% had over twenty years experience. Students in the DoDEA system scored much higher than the national average in the NAEP although the percentage of African Americans and Hispanics in that system is also much higher than the national average. Another positive link between teacher

experience and exemplary teaching was discovered in the SASS data (U.S. Department of Education, 1998). The percentage of teachers who reported that they were currently a master or mentor teacher in a formal teacher induction program was 2.5% for 0-3 years of experience, 8.8% for 4-9 years of experience, 12.7% for 10-19 years of experience, and 13.8% for 20 or more years of experience.

A study by Hanushek, Kain, and Rivkin (1998) of the Texas database for teachers that selected teacher experience as one school variable associated with student scores on the Texas Assessment of Academic Skills revealed that the first and second years of experience significantly improved teacher quality, but additional years rarely had a significant impact. Darling-Hammond (2000) also found that the benefits of experience level off after about 5 years. In an analysis of data for all Illinois schools collected for the Illinois School Report Card, Sutton and Soderstrom (1999) determined that teachers' experience averaged 15.13 years. This average experience, along with elementary class size, contributed the least amount of unique variance to the model of school control factors, while per-pupil expenditure, elementary pupil-teacher ratio, and teacher salary had the most. However, in a review of Kentucky's teachers, Clements (1999, p.12) may have stated it best:

> While no *necessary* correlation exists between teacher quality and years of experience, many who study teacher quality issues argue that time in the classroom indeed relates to teacher competency...the teacher rank and pay scale reflects the belief that those with more experience are 'worth' more to schools and communities than those with less.

Teacher Professional Development

The National Research Council (1999c) gave the task of guiding states in implementing Title I standards-based reform to the Committee on Title I Testing and Assessment. The committee developed a model of setting high standards, assessing student performance against those standards, and holding schools accountable for meeting those standards. In its most recent evaluation of this theory of action, the committee added an additional component, professional development for improved teaching. Data on indicators of professional development were limited and moderate in quality, according to the National Center for Education Statistics (Mayer, Mullens, Moore, 2001). For new methods of teaching such as cooperative learning, 61% of teachers spent one to eight hours in professional development and only 39% spent more than eight hours on new methods of teaching. In-depth study in the subject area of the teacher's main teaching assignment was more sustained with 44% spending one to eight hours and 56% spending eight or more hours in professional development. Using the SASS data, the National Center for Education Statistics (Choy & Chen, 1998) found that significantly more elementary than secondary teachers attended professional development activities, and significantly more teachers with 10 or more years experience attended professional development activities than less experienced teachers. Teachers in the state of Louisiana were categorized with states that had the lowest participation rates in two of the five types of professional development activities indicated.

In a nation-wide assessment of all state-reported low-performing schools, the Office of the Under Secretary and Office of Elementary and Secondary Education (U.S. Department of Education, 2001c) determined that "quality professional development,

targeted toward specific needs identified by the school, is an often-neglected element of the academic program in low-performing schools" (p. 34). Another study (Smrekar, Guthrie, Owens, & Sims, 2001) of the system of Department of Defense (DOD) schools throughout the world determined that the reverse is also true. Scores on the NAEP and the Terra Nova Comprehensive Test of Basic Skills in these schools placed the DOD school system second only behind Connecticut in a comparison with each of the United States. Further, the gap in average scores by race/ethnicity was not significant. Access to integrated, extensive professional development opportunities for the teaching staff, often taking place over weeks or months so that teachers could practice strategies in the classrooms, was seen as one of the factors contributing to teacher quality. Another was the fact that 64% of the staff had master's degrees and 25% had doctorates.

Another study that identified a link between academic achievement and professional development (Kim et al., 2001) used data from urban schools involved in the National Science Foundation's Urban Systemic Initiative program. Schools that participated in this program for the longest time saw the greatest gains in science and mathematics achievement, while reducing achievement gaps among racial/ethnic groups. Results of the Enacted Curriculum Survey used in this study revealed that teachers with professional development in standards-based curriculum and instruction reported teaching practices that were more consistent with state and national standards. The Longitudinal Evaluation of School Change and Performance of 71 Title I schools (U.S. Department of Education, 2001a) utilized standardized achievement tests, surveys, focus groups, documents, and records of school staff, students, and parents to examine changes in student performances in reading and mathematics. Gains in both curriculum areas were

related to high teacher ratings of staff development that matched the school's reform plan, focused on standards and assessments, and added to teachers' confidence in using new approaches.

Harkreader and Weathersby (1998) examined characteristics of staff development in Georgia's highest and lowest performing schools to determine any differences between the two. The staff of higher performing schools engaged in greater collaboration, focused more highly on students and their needs, and more effectively used the training strategies with administrative support than those in the lower performing schools. Additional characteristics of professional development that correlate with student achievement were observed in a field test of the Contemporary Mathematics in Context curriculum (Schoen, Finn, Griffin, & Fi, 2001). The 10 teachers whose students' mean gain scores in the Iowa Test of Educational Development were in the lowest quartile were compared to the 10 teachers whose students' mean gain scores were in the highest quartile. Higher gains were associated with teachers who either had a strong preparation for curriculum reform or who had completed a workshop specific to the curriculum being field-tested.

Title I Designation of School

Title I was originally designed to provide more funding for educating impoverished students than local or state tax bases provided. This compensatory education was analyzed by Mullin and Summers (1983) in a synthesis of 47 studies that all included achievement as the output measure. The major findings were that the programs had a small effect that was greater in earlier years but not sustained, there was no association between dollars spent and achievement gains, and no particular approach or program was consistently found to be effective. The 1988 reauthorization of Title I

added an accountability component that schools found ways to circumvent (Palmaffy, 1999). Then the 1994 reauthorization act required standards, assessment, and accountability programs that schools are still struggling to implement. Thirty states have received or are on track for waivers, but only 16 states have fully approved standards and assessment systems (Robelen, 2001). Louisiana is one of those with full approval.

Out of the 1988 concern for more accountability came the "most comprehensive and authoritative study of Title I ever undertaken" (Palmaffy, 1999, Prospecting for Gold, Finding Coal, ¶ 2). Under the leadership of Puma (1999), the Prospects study also disclosed a pattern of progress whereby students served by Title I retained the same achievement ranking relative to their classmates in later grades as they had when they started. Variations in student outcomes were substantially larger between students than among schools. In a nationwide study, Jerald (2001) specifically identified schools as high-poverty and high-minority with reading or mathematics scores in the upper third of their states. Of the 4,577 schools so identified, 67% of were Title I, while 44% of schools nationwide are Title I. Variability was thus identified among schools as well as between students.

Borman and D'Agostino (1996), in a meta-analysis of 17 Title I studies and 657 observations, concluded that effect sizes by subjects taught varied as well. Mathematics participants held a significant achievement gain relative to reading participants, but the advantage disappeared around seventh grade. Gains were greater from fall to spring testing, as opposed to annual testing, suggesting a detrimental summer effect. This effect appeared greater for the older grades.

Another finding of the Prospects study (Puma, 1999) was that evidence about the effectiveness of intensive tutoring is inconclusive. According to the report, the focus for improvement should therefore be on the regular classroom, where the student spends more time. A small trend began to implement school-wide Title I programs rather than continue the targeted-assistance programs that tracked all resources separately for select groups of students.

The Prospects study (Puma, 2000) also identified no discernable difference between Title I participants and disadvantaged non-participants. This was not a true experimental study with a control group, however, so conclusions could not be made about whether or not Title I students would have fallen farther behind in the absence of this program. However, the case can be made that the negative correlation of SES on achievement at the school level that was discussed earlier is true of Title I schools, which by their very definition have a large low-SES population. In particular, the *Longitudinal Evaluation of School Change and Performance (LESCP) in Title I Schools* conducted by the U.S. Department of Education (2001a) revealed that individual and school poverty had a clear, negative effect on student achievement, and that students who attended schools with the highest percentages of poor students performed worse initially on both reading and mathematics tests. Caldas and Bankston III (1997) used achievement data from 42,041 Louisiana 10th graders who took the Louisiana Graduation Exit Examination and aggregated participation in the federal free and reduced lunch program to the school level to define the SES of the peer population. He found that peer family social status does have a significant and substantive independent effect on individual academic achievement.

Setting the stage for a later mandated study of Title I, *Mapping Out the National Assessment of Title I: The Interim Report* (U.S. Department of Education, 1996a) reported that students in high-poverty schools lost ground in reading relative to students in other schools with the gap changing from 20 points to 34 points. More than 11,000 Title I schools were identified as being in need of program improvement, with over 1,000 of them being in that category for over four years. *Promising Results, Continuing Challenges: Final Report of the National Assessment of Title I* (U.S. Department of Education, 2000c) reported an improvement for high-poverty schools in mathematics scores, however, with a gain of 10 points while low-poverty schools only gained 9 points in the same ten-year time period. Progress was very uneven from state to state. In Maine, 80% of fourth graders in high-poverty schools scored at or above the *Basic* level in mathematics, and only 25% of fourth graders in California did so. Louisiana had 36% score at that level, well below the national average of 62%. Teachers who used curricula with NCTM standards had students with higher gains in mathematics, but only 37% of teachers in high-poverty schools felt well prepared to do so. An overwhelming 70% of teachers in high-poverty schools divulged receiving less than nine hours of professional development per year related to content and performance standards.

The *National Assessment of Title I* also revealed that the use of the pull-out model of instruction has decreased, while in-class models, school-wide programs, and extended-time instruction have all increased. However, a Brookings Institute paper (Farkas & Hall, 2000) reported observing across a variety of schools and districts that the pull-out model is still by far the most common practice among Title I programs. Wong and Meyer's (1998) synthesis of findings on school-wide programs concluded that results are mixed

and inconclusive. Only a handful of the thousands of programs have disseminated reliable evaluation data.

Preschool Attendance

Another area that educators have targeted in their efforts to reduce the academic gaps among students is that of early childhood education. However, as a report by a team of researchers for the Organization for Economic Cooperation and Development (2001, ¶ 43) put it, "An important starting point for trying to understand the U.S. system of early childhood education and care is to realize that there is no 'system'." The study of 12 nations determined that almost all of them provided more extended family leave for the parents of young children, more generous child-care allowances, and greater support for high-quality, early-childhood programs than did the United States. The federally funded Head Start program, state-funded programs, and a wide variety of private care facilities comprise the U.S. system. In 1999, 60% of White children, 73.2% of African American children, and 44.2% of Hispanic children ages three through five were enrolled in early childhood programs (U.S. Department of Education, 2000b). Almost 74% of children whose mothers had bachelor's degrees or higher were enrolled, while 40.3% of children whose mothers had less than a high school education were enrolled in such programs. Of the 722,893 Head Start enrollments nationwide, 20,402 were in Louisiana (Public Agenda, 2002a). Louisiana also had 2,110 of the 105,564 licensed child-care centers and 11,000 of the 286,568 licensed family child-care homes in the United States. In addition, 2,877 children were enrolled in state-financed prekindergarten programs (Doherty, 2002). A Rand report, in an extensive analysis of NAEP scores, stated that when all other things

were equal, NAEP scores were higher in states that had more children in public prekindergarten programs (Grissmer, Flanagan, Kawata, & Williamson, 2000).

Since the quality of early child-care varies as much as the providers themselves, one longitudinal study by the National Center for Early Development and Learning (Clifford, Peisner-Feinberg, Culking, Howes, & Kagan, 1998) followed a random sample of 401 child care centers in California, Colorado, Connecticut, and North Carolina over a three year period from preschool through second grade. They found that children with high-quality care—as defined by measures that included classroom environment, teacher sensitivity, child-centeredness, and teacher responsiveness—had better language skills and mathematics skills over this time than students in schools with lower-quality care. Children in higher-quality classrooms had better reading skills during the first year of preschool, but there was no difference in reading skills after that. A national survey of 1,902 teachers of preschoolers (Early, Clifford, & Howes, 1999) discovered that teachers reported that they were generally able to engage in the practices they endorsed. One particular practice in the survey was called group-centered beliefs in which all children engage in the same activity at the same time, as opposed to child-centered beliefs in which children choose from a variety of activities at their own pace. Teachers in public schools, Head Start, and other non-profit centers endorsed group-centered beliefs significantly less than did teachers in religiously affiliated or for-profit settings. Also, teachers with more education endorsed group-centered beliefs less.

A longitudinal study by Marcon (1995, 2000; Marcon, Randall, & Brooks, 1997) tracked 249 children from 67 schools from preschool through sixth grade. The children were clustered according to their preschool teachers' instructional beliefs and practices.

Teachers labeled Model CI utilized an active child-centered approach, teachers labeled Model AD were more didactic with academically-directed programs, and teachers labeled Model M fell in between the other two models. The study determined that students with Model AD teachers did significantly better in first grade. Students with Model CI teachers made the transition to fourth grade more successfully than the other two models. Sixth grade academic achievement was enhanced by early learning experiences that emphasized the child-centered approach over the academically-centered approach, particularly for males.

One study (Randle, 1997) of 30 third graders in a 100% minority Chicago public school who had attended preschool and 30 randomly selected third graders in the same school who had not attended preschool found no significant difference in ITBS reading scores. This study did not differentiate among types of preschool experiences.

A statewide study (Roth, Carter, Ariet, Resnick, & Crans, 2000) of 12,098 children who had participated in Florida's Prekindergarten Early Intervention (PKEI) program in 1992-1993, who remained in Florida public schools for the next 5 years, and who participated in the federal free or reduced lunch program and the 62,927 students with the same criteria who served as a control group because they had not participated in the PKEI revealed different effects by ethnic groups. African American and Hispanic students who participated in PKEI had a significantly greater chance of being in the highest, second-highest, and third-highest categories for both reading and mathematics achievement on the statewide assessment, but White students did not. Female PKEI participants were significantly higher in both reading and mathematics than male PKEI

participants. PKEI participants were more likely to be in the testable category rather than having been retained or being assigned to Exceptional Student Education.

Two studies followed the effect of preschool education on into young adulthood. One was the Abecedarian Project (Campbell & Pungello, 2000), which followed 105 of the 111 original participants to age 21. These participants were randomly assigned to an intensive preschool group, a group that received home visits, and a control group that received free formula and disposable diapers. All participants received free health care until they entered kindergarten. The school included child-directed and teacher-directed activities. Initially, preschool treatment had a positive effect on IQ performance, learning and cognitive performance, language development, social responsiveness, and academic locus of control, with the greatest effect occurring in children of mothers with IQ's below 70 (Ramey & Ramey, 1994). At kindergarten entry, the preschool and control groups were re-randomized to form four groups, two with a home-school resource teacher for the first three years in elementary school (Campbell & Pungello). Reading scores showed a consistent increase as a function of years of treatment. Mathematics scores showed an identical increase for both preschool groups, and the group in the primary-treatment-only group slightly outscored the untreated group. In addition, individuals treated in preschool completed significantly more years of education by age 21 than did the preschool controls, particularly females, with almost three times as many enrolled in a four-year college.

Another longitudinal study, the High/Scope Preschool Curriculum Comparison Study (Schweinhart & Weikart, 1997), randomly assigned 68 children to one of three preschool instructional models: the High/Scope Educational Research Foundation

preschool in which children learned actively through plan-do-review and group times, Direct Instruction preschool in which teacher-directed scripts focused on academics, or traditional Nursery School in which children learned through play. For a decade, no curriculum group differences in intellectual or academic performances were found. At age 15, the Direct Instruction group had three times as many arrests per person. In the Direct Instruction group 47%, as compared to 6% of the other two groups of the students, were treated for emotion impairment or disturbance. At age 27 the High/Scope group had graduation rates more than 30% higher than their counterparts and significantly higher average achievement scores and literacy scores. These results corroborated Marcon's findings that child-initiated curricula were more beneficial for students in the long run.

The Louisiana state superintendent of education estimated that about 14,000 four-year-olds in Louisiana now living in poverty were not receiving prekindergarten services through programs such as Head Start (Edwards, Chronister, & Olson, 2002). In January 2002 the Louisiana legislature initiated a program to expand its existing preschool program. Under the program, public schools operate preschools within existing elementary school buildings or form partnerships with private childcare facilities. Each classroom is required to have a lead teacher certified in early childhood education. Updated state standards in early childhood education raise expectations for preschool experiences. The estimated annual cost of sustaining full participation in this program is $70 million. A variety of assessments of costs and benefits discussed by Kasely, et al. (2001) detailed many savings by stakeholders as a result of preschool participation, such as fewer students repeating grades, reduced special education services, reductions in involvements with the criminal justice system, higher employment, and less welfare

dependence. Other, less tangible, benefits were more difficult to quantify, such as fewer child abuse cases, less caregiver stress, and improved quality of home environment. Much evidence existed that showed the cost-effectiveness of early interventions versus later public expenses. If preschool programs, such as the one in Louisiana, are carefully designed and evaluated, they can provide additional evidence to justify these initially great expenses.

Norm-referenced and Criterion-referenced Tests

Assessments are often used as evidence of academic achievement. In particular, two main types of tests, norm-referenced and criterion-referenced tests, have been most commonly used, with a third, standards-referenced, being added as a result of the 1995 Improving America's Schools Act. A norm-referenced test (NRT) does not indicate mastery of particular skills, acceptable progress, or measurement of desired outcomes (Cizek, 1998). Instead, it focuses on ascertaining where the test-taker stands in relation to a control group of students for whom the test was originally normed, or averaged. Appropriate comparisons are limited to other students who have taken the same test, rather than students taking other norm-referenced tests or other more global comparisons. However, a comparison of item difficulty indices for the major batteries within NRTs reveals that the ITBS is comprised of slightly more difficult items and the California Test of Basic Skills of slightly easier items, with other major batteries falling in between. One important note about NRTs is added by Popham (2001). Since these tests were originally designed by the army in World War I to identify potential officers, test items are selected on their basis to create scores that are widely spread rather than their ability to assess what should be taught in school.

Criterion-referenced tests (CRT) are designed to determine whether a student knows or can do specific things (Cizek, 1998). The student either has performed up to expectations or not, and consequently either passes or fails the test. This type of test does not indicate whether a student is better or worse than average or how appropriate the criteria are that are being tested. The Developmental Reading Assessment (DRA) is an example of this type of test. The third type of test, the standards-referenced test (SRT), is similar to the CRT in that both attempt to describe the knowledge, skills, or abilities that students possess. The SRT does so in reference to content standards that are developed to represent academic statements of what students should know and be able to do in specific subjects. This test consists of sets of items or tasks designed to measure the student's knowledge and abilities over a range of performance standards, which describe how well students need to be able to perform on a set of standards in order to meet pre-defined specified levels of expected performance. The term *standards-referenced* does not have widespread use, and most of the literature continues to refer to these tests as criterion-referenced. The NAEP and the LEAP 21 are examples of this type of test.

The LEAP 21 Tutoring Program uses a student's percentile ranking on the ITBS, a NRT, to predict vulnerability on the LEAP 21, a SRT. This has also been the case in several studies of statewide assessments. In one study (Janiak, 1999) using a stratified random sample of 600 out of 15,960 mostly White third-grade students in one Florida county, the Terra Nova norm-referenced reading subtest was correlated with the reading section of the Florida Comprehensive Assessment Test (FCAT) to determine whether the Terra Nova was a good predictor of students who achieved the lowest rank on the FCAT. The Terra Nova score had the strongest relationship ($r = .73$) out of nine predictor

variables. A statewide study (Beard, Kevan, Posrie, & Sheridan, 1990) was conducted to determine the feasibility of predicting success on the State Student Assessment Tests (SSAT) at the 3rd, 5th, 8th, and 10th grades using the three different NRTs used by the school districts in Florida. Three districts were chosen from those that used each of the three NRTs. These districts provided more than the requested data for 300 students each, so a random sampling technique was used to create a data bank of 1000 students per NRT. Racial composition varied by test, so comparisons among the three were not appropriate. Because the NRT total scores predicted the communications and mathematics subtest scores approximately as well as the NRT verbal and mathematics subscores, respectively, and because the NRT total scores predicted the SSAT total scores better than the part scores, it was decided to focus on the use of NRT total scores to predict SSAT total scores. The different NRTs were found to be similar in their ability to predict the SSAT scores. However, the NRT cut-off scores found to be optimal for predicting success on the SSAT varied among the different NRTs by as much as 10 normal curve equivalents at one grade level.

Another study (Demps & Onwuegbuzie, 2001) correlated the relationship between the ITBS taken in the eighth grade and all five subtests of the Georgia High School Graduation Tests (GHSGT). Students involved in the study were 102 members of the 1999 graduating class of one Georgia high school. ITBS scores had significant relationships with all subtests of the GHSGT. Moreover, students who failed a portion of the GHSGT typically scored in the lowest quartile of the distribution of eighth-grade ITBS reading scores.

Individual ITBS subtest scores were correlated with the ELA and mathematics portions of the LEAP 21 for 53 students attending two elementary schools in north Louisiana (Cole, & Watts, 2001). Scaled scores on the mathematics portion of LEAP 21 and national percentile rank on the reading comprehension subtest of ITBS were the best predictors of performance on the ELA component of LEAP 21. Performance on the ELA component of LEAP 21 and the mathematics problem-solving subtest of the ITBS were the best predictors of performance on the mathematics portion of the LEAP 21.

In Texas, a study (High, 1996) was undertaken to determine the predictive value of the Assessment of Student Skills for Entry Transfer (ASSET) college placement test, a NRT, on the Texas Academic Skills Program (TASP), a CRT. A total of 328 students from six colleges were selected who had taken the ASSET and completed the resulting remedial courses before the TASP. The ASSET scores were the best predictors among other variables in every comparison.

In the Atlanta Public Schools system, data from 83 elementary schools, 10 middle schools, and 20 high schools were used to determine which variables had the strongest relationship with scoring at or above the national norm in reading and mathematics on the ITBS (Brooks, 1988). In the elementary schools, 20 predictor variables were identified, while both middle and high schools had 16 predictor variables identified. Socioeconomic status (SES) was the best predictor for elementary students in 1986, but the Georgia Criterion-Referenced Test (GCRT) was the best predictor at this level for 1987. In the middle school, the GCRT score was the strongest predictor in 1986, but SES was the strongest in 1987. At the high school level, SES best predicted both reading and

mathematics scores for1987. However, the GCRT score was the best indicator for mathematics in 1986.

A recent attempt was made to correlate a performance-referenced test with a NRT (Meisels, Bickel, Nicholson, Xue, & Atkins-Burnett, 2001). Performance-referenced assessments, according to the U.S. Department of Education, Office of Civil Rights, are "product- and behavior-based measurements based on settings designed to emulate real-life contexts or conditions in which specific knowledge or skills are actually applied" (2001d, Appendix B: Glossary of Test Measurement Terms, ¶ 38). Thus, they are more similar to CRTs than to NRTs. In one study (Meisels et al., 2001), the Work Sampling System (WSS) scores of second- and third-grade students of 17 teachers who had demonstrated high inter-rater reliability were compared to the matching subtests of the Woodcock-Johnson Psychoeducational Battery-Revised (WJ-R). The majority of the correlations between the WSS and the comprehensive scores of children's achievement were similar to correlations between the WJ-R and other standardized tests, in the .50 to .80 range. Moreover, the probability of a student performing poorly or well on both the WJ-R and the WSS was 84% for language and literacy and 84% for mathematics, thus making it a reliable predictor of children at risk.

The Just for the Kids organization (n.d.) reported the linkage between the LEAP 21 and the NAEP as shown in Table 2. On the surface, the percent of students meeting the standard in each category of the two tests appear related. As is more fully explained in Chapter 3 of this study, the state of Louisiana took great care to align the LEAP 21 with the NAEP, so that the two would correlate highly with one another. Thus a SRT could be used to show a strong relationship with another SRT.

Table 2. *Comparison of LEAP 21 and NAEP Reading Scores*

Percent of Louisiana Students Meeting State and NAEP Standards

Fourth Grade Reading, 1998 and 1999

Standard	Percent Meeting the Standard
Louisiana *Basic* or above – fourth grade reading 1999	55%
Louisiana *Proficient* or above – fourth grade reading 1999	16%
Louisiana *Advanced* – fourth grade reading 1999	1%
NAEP *Basic* or above – fourth grade reading 1998	48%
NAEP *Proficient* or above – fourth grade reading 1998	19%
NAEP *Advanced* – fourth grade reading 1998	3%

Note: From *Just for the Kids: School Data Availability in Other States: Louisiana*, retrieved June 15, 2001, from http://www.just4kids.org/states/Louisiana.htm

Summary

This review of the literature on ten variables selected as being potentially related to students' scores on the LEAP 21 included a discussion of the varying strengths of the influence of students' SES at the national, state, district, school, and individual levels. Status as a Title I school was discussed as it relates to the aggregation of low SES at the school level. Race was noted as a potentially strong factor, particularly in Louisiana (Education Trust, 2001). Gender may have some bearing as it relates individually to the English Language Arts (ELA) and to the Mathematics subtests. Qualifications of teachers by their degree status, their years of teaching experience, and their participation in intense state-sponsored or district-sponsored professional development could potentially influence students' scores, according to the literature. Current impetus for a statewide preschool program was built on the evidence of studies that early childhood education makes a difference in academic achievement. Use of the statewide CRT, the Developmental Reading Assessment, as a predictor of a student's success on the LEAP 21 has some basis of merit in previous similar studies. The present model of identifying students for the LEAP 21 Tutoring Program when those students have initially failed either the ELA or Mathematics subtests was also justified in the literature.

CHAPTER 3

METHODOLOGY

Research Design

The researcher used an ex post facto design to determine if there was a strong likelihood that when particular predictor variables were present certain criterion variables were likely to be present as well (ELA, Mathematics, or combined ELA and Mathematics scores on the Louisiana Educational Assessment Program for the 21st Century [LEAP 21]). The predictor variables included student characteristics of grade retention in fourth grade, core total national percentile on the third grade Iowa Test of Basic Skills (ITBS) test, final third grade Developmental Reading Assessment (DRA) scores, socioeconomic status (SES), race, gender, and preschool attendance, and school characteristics of teachers' degree level, teachers' experience, teachers' participation in professional development, and the Title I status of the student's school. An ex post facto design is used frequently to identify possible causal relationships between variables. Crowl (1996) stated, "researchers use regression, not correlation, when they wish to predict values of one variable from values of another variable" (p. 159). Since multiple variables were identified in this study, a step-wise multiple regression was used. As Ferguson and Takane (1989) explained, the backward elimination technique for the step-wise multiple regression begins with all predictors and ascertains what degree of prediction is lost by

progressively dropping out one variable at a time. Thus the strongest predictors were identified, and the relative strengths of the two that the state of Louisiana currently uses as criteria for students' eligibility for the LEAP 21 Tutoring Program were determined.

Teacher observations during the gathering of the data for the preschool predictor variables were recorded along with their opinions about what contributions professional development made to their effectiveness in the classroom. This process served to add qualitative information to the research.

Sample

The sample of the study consisted of the entire population of 594 fourth grade students in one north Louisiana school district who received scores on both the English Language Arts and the Mathematics portions of the LEAP 21. This included students who had participated in the LEAP 21 Tutoring Program and those who had not. This school district had two rural K-12 schools, five schools in mid-sized urban areas, and three K-5 schools in consolidated rural districts for a total of 10 schools. A total of 40 teachers were included from both general education and special education classes.

The district was selected on the basis of convenience for the researcher. However, some attributes about the district were representative of the state of Louisiana as a whole. In Louisiana at the time of this study, 49.7% of elementary and secondary school students were White and 47.1% were African American (Louisiana Department of Education, 2000c). About 47.3% of the sample group was White (281), and about 51.9% was African American (309). The 1% remaining in the sample, three Hispanic and one Asian American, was not large enough to be considered. Approximately 55.2% were males and 44.8% were females. In Louisiana, 23.6% of all persons were living below the poverty

level (Louisiana Department of Education, 2000b), and 22.8% of all persons in this school district were living below the poverty level. ELA and Mathematics LEAP 21 scores in fourth grade for the year 2000 for this school district (Louisiana Department of Education, 2000a) were comparable with those of the state (see Table 3). Comparative information on additional measures (Louisiana Department of Education, 2001c) can be seen in Table 4.

Table 3. *State and District LEAP 21 Scores for 2000*

Achievement Level	State	District
	ELA Scores	
	% Attained	% Attained
Advanced	2	1
Proficient	14	16
Basic	39	41
Approaching Basic	25	26
Unsatisfactory	20	16
	Mathematics Scores	
	% Attained	% Attained
Advanced	2	1
Proficient	10	9
Basic	37	34
Approaching Basic	23	23
Unsatisfactory	28	33

Note: From *Louisiana Educational Assessment Program for the 21st Century: 1999-2000 annual report* by Louisiana Department of Education, 2000, p. 12-13, 28.

Table 4. *State and District Accountability Results for Elementary Schools in 2001*

Growth Labels		
	State %	District %
No label	6.0	7.7
Exemplary Academic Growth	45.2	53.8
Recognized Academic Growth	25.0	0
Minimal Academic Growth	17.7	15.4
No Growth	4.7	15.4
School in Decline	1.3	7.7
Performance Labels		
	State %	District %
School of Academic Excellence	0.1	0
School of Academic Distinction	1.3	0
School of Academic Achievement	15.6	15.4
Academically Above the State Average	30.8	30.8
Academically Below the State Average	50.0	53.8
Academically Unacceptable School	2.2	1.3
Other		
	State %	District %
Eligible for Rewards	69.4	53.8
Corrective Actions I	14.2	38.5
Corrective Actions II	2.3	0

Note: From *2001 Accountability Summary Results: State and District Table* by the Louisiana Department of Education, 2001. Retrieved August 8, 2002, from Author's on-line database: http://www.doe.state.la.us/DOE/asps/home.asp?1-REPORTC

Scores for this school district on the district component of the Louisiana Accountability Program were released for the first time in 2002. These scores were also compared to the average for the state in Table 5.

Table 5. *District Accountability Summary with Louisiana State Averages*

District Performance Score (DPS)		
	State Average	District Score
LEAP 21 Index	75.5	75.5
Iowa Index	79.8	77.8
Attendance Index	100.1	95.0
Dropout Index	132.5	112.5
District Performance Score (DPS)	80.8	79.0
District Responsibility Index (DRI)		
	State Average	District Score
District Responsibility Index (DRI)	115.6	111.5
District Responsibility Label	Very Good	Very Good

Note: From *2000-2001 District Accountability Summary Table* by the Louisiana Department of Education, 2002, p. 4. Retrieved August 8, 2002, from the Author's on-line database: http://www.doe.state.la.us/DOE/asps/home.asp?1-REPORTD

The school district selected for this study conducted the LEAP 21 Tutoring Program as a pullout program during the school day twice each week. During the first

half of the tutoring, the focus was on ELA; and during the second 5-weeks period, mathematics topics were addressed, for a total of 10 hours in each subject.

Instrumentation

LEAP 21. Development of the LEAP 21 test instrument was begun in 1996 with the formation of an assessment advisory committee of educators representing kindergarten through higher education, assessment specialists with the Louisiana Department of Education, and national consultants (Advanced Systems in Measurement & Evaluation, Inc., 1999). The committee first established specifications that included multiple-choice, short-constructed-response, and extended-constructed-response items. The ELA test also had a composition subtest to assess writing skills. Three unique operational forms were created for each content and grade level. Content validity was established for each test form across the scope and sequence of the content standards. To ensure that the standards of the LEAP 21 were in alignment with those of the National Assessment of Educational Progress (NAEP), a linkage between the two assessments was also analyzed during the field test.

An equity review committee, composed of a broad range of stakeholders, created sensitivity guidelines, performed a review of the test items prior to field testing, and established an on-going review process for the accountability program. Staff from the National Center on Educational Outcomes assisted the committee in selecting appropriate allowable testing accommodation and the administration of an alternative assessment for the 2% or less of the student population working to alternative standards. Large print and Braille formats of the test were produced.

After evaluators were trained and the field tests scored, several statistical analyses were performed to further refine the three test forms. An item analysis was conducted to isolate test items with gender or ethnic bias and to determine classic item difficulty and item discrimination. Cronbach's coefficient alpha was computed to show a range of .87 to .92 on the various field test forms. Content analysis ensures that the entire scope of the standards was covered. Released NAEP test items were included in part of the field test in a linking study of correlations to verify that standards for the LEAP 21 were in alignment with those of the NAEP. Performance standards of *Advanced*, *Proficient*, and *Basic* from the NAEP test were also adopted, but the NAEP's *Below Basic* category was divided into two: *Approaching Basic* and *Unsatisfactory*. Using four rating committees, the field tests were evaluated for cut points, the numerical scores that determined a change from one category to another. In most instances, the LEAP 21 cut points were established at a higher level than similar categories on the NAEP. Correlation with the ITBS was high with .84 in mathematics for both grades and .78 and .67 for grades four and eight, respectively, in reading. Phase II development of the science and social studies portions of the LEAP 21 and the development of the Graduate Exit Exam 21 followed a similar process. *Education Week* gave Louisiana's standards and accountability program a grade of A- in its *Quality Counts 2002* (Edwards, Chronister, and Olson, Eds., 2002).

Iowa Test of Basic Skills-Form M. Students' third grade composite national percentile scores was one of the student characteristics used as an predictor variable in the research model. The reliability for the Developmental Standard Score for the ITBS Form M Complete Composite with computation in the spring was .979 (Linda Machut, personal communication, April 2, 2002). National percentile ranks showed a student's

standing within the group of students in the same grade who were tested at the same time of year during the national standardization. The procedures used to generate empirical national norms for Form M were last conducted in 1995.

As part of the development of Form K/L/M of the ITBS (The Riverside Publishing Company, 1997), several national item tryout studies were conducted with members of target populations by gender and race over-represented to permit comprehensive analysis of differential item functioning. Several panels of ethnically and racially diverse educators were convened to review each test stimulus item and test question for possible inappropriateness or cultural or regional bias. Standard errors of measurement for the Composite with Computation for Level 10 (grade three Spring) were White 2.8, Black 2.6, female 2.7, and male 2.8.

Developmental Reading Assessment (DRA). The final third-grade score of the DRA was included as one of the student characteristics in the research model. A nationally representative sample of 306 students was included in the reliability study for this instrument (Williams, 1999). A total of 87 teachers from 10 states audio-taped assessments of three or more children. These were each sent to two other teachers who also assessed them. Rasch scale (facet) analysis revealed inter-rater agreement across the first two raters of 0.80. Adding the third rater dropped it to 0.74. The internal consistency was found to be quite strong for item separation reliability (K = 0.98), and for text separation reliability (K = 0.97).

To establish the construct validity of the DRA, 2470 students from one large, urban school district were assessed. Scores were correlated with the students' scores from fall of third grade on the ITBS subscales: Vocabulary, Reading Comprehension, and

Total Reading. All correlations were significant at the 0.01 level (2-tailed) using Spearman's Rho rank order correlation. However, the highest correlation was with Total Reading ($r = 0.71, p < .01$).

Survey of Pre-First Grade Experience. This researcher-created instrument has face validity because it received content validation from subject matter experts. It was used to determine whether or not students have preschool experience. A comparison of archival records with the first question of the survey was used as a verification technique for the accuracy of the information gathered. In one selected school, archival data for all students whose kindergarten enrollment was present were compared with responses to item 1 in this survey. Kindergarten attendance for all students for whom both archival data and survey data were present matched 100%.

Survey of Fourth Grade Teachers. This researcher-created survey instrument served as a record of fourth-grade teachers' observations and reflections only and therefore has face validity.

Procedural Details

The Survey of Pre-First Grade Experience (Appendix A) and the Participant Consent Form (Appendix B) were given to all fourth-grade students in the sample. Directions for Administering the Pre-First Grade Experience Survey (Appendix C) were given to the fourth-grade teachers. They read the items and possible responses, giving students as much time as necessary to respond to all items. Data recorded from the Survey of Pre-First Grade Experience consisted of a simple *yes* or *no* for participation in preschool. If all three last responses were marked in items three through five, the student

was considered as not having participated. If any other pattern of responses was marked for items three through five, the student was recorded as having attended preschool. Of the 594 students who took the LEAP 21, 507 (82.1%) completed the survey.

Third-grade percentile scores on the ITBS and final third-grade DRA scores for each student and demographic data such as race, gender, and participation in the federal free and reduced lunch program for each student were retrieved from school district archival data. Data on teachers of these students concerning their degree status, years of teaching experience, and participation in the state-sponsored Intech and Louisiana Systemic Initiative Program professional development experiences or in the Math Their Way, Math A Way of Thinking, or Balanced Literacy district-sponsored professional development opportunities were collected from the school district archives along with whether or not the student's school received Title I funds. The LEAP 21 scores were retrieved from the Louisiana Department of Education Web Reporting System.
The Survey of Fourth Grade Teachers (Appendix D) was sent in individually addressed envelopes with stamped return envelopes enclosed to each fourth-grade teacher at his or her school. Out of the 30 surveys sent, 17 (56.7%) were completed and returned.

Data Collection

Archival data on the Iowa Test of Basic Skills (ITBS) and the Developmental Reading Assessment (DRA) were missing for 46 students because they had transferred from out of state, had transferred from private schools, or had been home-schooled. An additional 26 DRA scores were missing due to problems with record keeping or with transfer of records within the school district. This made a total of 452 students who had complete records for the multiple regression model. Although only 36.4% of the students

were in the free- and reduced- federal lunch program, 57.4% of them were in Title I schools. Of the total population, 74 (12.5%) had repeated fourth grade and 416 (70%) had attended preschool. One teacher who taught 15 students had a temporary teaching certificate, 351 students were taught by teachers with bachelor's degrees, 105 students were taught by teachers with master's degrees, and 110 students had teachers with master's degrees plus 30 additional hours of credit. The mean for teacher experience was 17.23 years. Of the 593 students whose teachers' participation in professional development was recorded in the district archives, 233 (39.2%) of them were taught by teachers who had participated. Participation by the 40 teachers involved in the study in specific professional development experiences was accounted for as follows: Intech (N=3), Louisiana Systematic Initiative (N=3), Math Their Way (N=5), Math a Way of Thinking (N=4), and Balanced Literacy (N=2), with some teachers participating in more than one on the list. Twenty-seven teachers were nonparticipants in professional development experiences.

Since students rather than parents responded to the Pre-First Grade Experience Survey, concern over the accuracy of the responses was addressed by asking fourth-grade teachers to note comments made by the students that would help determine whether or not they remembered any preschool attendance. Of the Survey of Fourth Grade Teachers mailed to the 30 teachers involved in the study, 17 responses (57%) were returned. Only one teacher said that most of her students did not remember with comments like, "I don't know," and "That was too long ago." One other teacher said that she could not tell whether her students knew about their preschool attendance or not. Each of the 15 other teachers, 50% of the total teachers involved in the survey, said that most students knew

about their attendance although the students were not always clear about specific types of preschools, specific lengths of time spent in the preschool setting, or specific activities that were included in those experiences. Since these items were not used to determine the presence of this variable, these inaccuracies were irrelevant. Many of the students attended preschools associated with their elementary schools and knew their teachers names. One group of students discussed the fact that their preschool had relocated since they had attended it. Others made "I remember when..." comments while talking about their experiences.

Pilot Study

A pilot study of the Survey of Pre-First Grade Experience was conducted in all fourth-grade classes of one elementary school in north Louisiana outside of the school district involved in the study. Parents of the students completed this version of the Survey of Pre-First Grade Experience. Since the return rate was only 41%, subject matter experts agreed that the students should respond to the survey with their teachers reading each item aloud. The higher anticipated return rate was deemed to be more valuable than the potential loss of accuracy.

Data Analysis

The three research questions from Chapter 1 were considered for data analysis in the following null hypothesis: Both student characteristics (students' grade retention in fourth grade, core total national percentile on the third-grade ITBS, final third-grade DRA scores, SES, race, gender, and preschool attendance) and school characteristics (teachers' degree level, teachers' experience, teacher's professional development, and

students' school Title I status) have no relation to LEAP 21 scores. This hypothesis was applied to three outcome variables—ELA LEAP 21 scores only, Mathematics LEAP 21 scores only, and a combined total of the ELA and Mathematics LEAP 21 scores—because this gave more precise information for different weightings of ELA and Mathematics LEAP 21 scores.

The hypothesis was tested with a step-wise multiple regression. Regression analysis techniques allow the assessment of the relationship between a group of predictor variables and one criterion variable. The result of a regression is an equation that represents the best prediction of the criterion variable from several continuous, discrete, or dichotomous independent variables. All predictor variables were dichotomous, discrete, or normally distributed, except one, the national percentile score of the student's ITBS core total. The criterion variable was normally distributed around the prediction line.

Some predictor variables were entered as codes. Students who did not repeat fourth grade were coded as *0*, while students who did repeat fourth grade were recorded as *1*. Students who participated in the free- and reduced- federal lunch program were coded as *0* for SES, and those who did not were coded as *1*. White students were coded as *0*, and African American students were coded as *1* for race. Female students were coded as *0*, and males were coded with *1*. Students who did not participate in preschool were coded as *0* for that predictor variable, and those who did participate were coded as *1*. Teachers' degrees were recorded as follows: *0* for temporary certification, *1* for bachelor's degree, *2* for master's degree, and *3* for master's degree plus 30 additional hours of graduate credit. Teachers' years of experience were listed to the nearest 10th.

Students whose homeroom teacher participated in the state-sponsored Intech and Louisiana Systemic Initiative Program professional development experiences or in the Math Their Way, Math A Way of Thinking, or Balanced Literacy district-sponsored professional development opportunities were recorded as *1*, and students whose teachers did not participate in one of those professional development opportunities were coded as *0* for professional development. Students enrolled in Title I schools with either School-Wide programs or Targeted-Assistance programs were coded as *1* for the Title I predictor variable, and all other students were coded as *0*. The computer software used for analyzing the data was SPSS, developed by SPSS, Inc.

CHAPTER 4

RESULTS

Descriptive Results

Descriptive statistics for continuous variables and the national percentile scores of the core total of the Iowa Test of Basic Skills (ITBS) are displayed in Table 6. Scatterplots revealed a stronger linear relationship between ITBS and LEAP 21 scores than between teacher experience and LEAP 21 scores. The range of teacher experience may better be defined as 0-33 years with an outlier of 46 that affected 19 cases. Frequency distributions for dichotomous and discrete variables are listed in Table 7.

Table 6. *Descriptives for Selected Variables*

Variable	Minimum	Maximum	*M*	*SD*	*SE*
Teacher experience	0	46	17.23	9.35	.39
ITBS core total NPR[a]	1	99	46.46	26.06	1.11
LEAP 21 ELA score	100	492	309.79	58.73	2.42
LEAP 21 Mathematics score	100	490	310.47	49.67	2.04
Combined LEAP 21 ELA and Mathematics scores	238	884	620.25	102.62	4.22

[a]National Percentile Rank

Table 7. *Frequency Distributions for Dichotomous and Discrete Variables*

Variable	Frequency	Valid Percent
Repeating 4th grade		
Did not repeat	516	87.5
Did repeat	74	12.5
Developmental Reading Assessment (DRA)		
Below level	119	22.6
On level	183	34.8
Above level	224	42.6
Socioeconomic status		
Free- and reduced-lunch program	376	63.5
Non-federal lunch program	216	36.5
Race[a]		
White	308	51.9
African American	281	47.4
Gender		
Female	266	44.8
Male	328	55.2
Preschool		
Did not participate	91	17.9
Participated	416	82.1
Teacher degree		
Temporary certificate	16	2.7

Variable	Frequency	Valid Percent
Bachelor's degree	355	61.0
Master's degree	100	17.2
Master's degree + 30 graduate hours	110	18.5
Teacher professional development		
Did not participate	360	60.7
Participated	233	39.3
Title I School		
Not enrolled in Title I school	253	42.6
Enrolled in Title I school	341	57.4

[a]Three Hispanic students and one Asian American student were not included in the study because of small sample size.

Bivariate Relationships

Predating the prevalence of computers, point-biserial correlation was computed by hand as the normal correlation used when an interval variable is correlated with a dichotomous variable (Ferguson & Takane, 1989). However, SPSS calculates the exact correlation regardless of whether the variables are continuous or dichotomous (Garson, 2001). The correlations for all dichotomous predictor variables and the ELA LEAP 21 scores, the Mathematics LEAP 21 scores, and the combined ELA and Mathematics LEAP 21 scores are listed in Table 8. Only preschool attendance and teachers' participation in professional development had no correlation with any LEAP 21 scores. SES, race, and enrollment in Title I schools had the strongest relationship with LEAP 21 scores out of the dichotomous predictor variables.

Since students who did not repeat fourth grade were coded with a *0*, and students who did repeat fourth grade were coded with a *1*, the negative correlation for repeating fourth grade means that students who did not repeat had a stronger tendency toward higher scores than those who did repeat fourth grade. Participation in the federal free- and reduced-lunch program was coded as *0*, while nonparticipation was coded as *1*. Therefore, the positive correlation showed that students with a higher SES based on the federal lunch participation indicator had a strong relationship with higher LEAP 21 scores.

White students were coded as *0*, and African American students were coded as *1*, so the negative correlation for race means that White students were more strongly related to higher LEAP 21 scores than were African American students. This relationship was significant at the .01 level. Female students were coded as *0*, and male students were coded as *1*. The negative correlation means that females did better than males only on ELA LEAP 21 scores, $p < .01$. However, this relationship was strong enough to make the female student correlation with combined ELA and Mathematics LEAP 21 scores significant at the .01 level.

Students enrolled in a school that had either a Title I School-Wide program or a Title I Targeted-Assistance program were coded as *1* for the Title I predictor variable, and all other students were coded as *0*. The negative correlation therefore means that students who were not enrolled in Title I schools did better than students who were enrolled in Title I schools at the .01 level of significance. Two variables, preschool attendance and professional development, had no significant relationship to any of the LEAP 21 scores.

Table 8. *Pearson Correlations Between Dichotomous Variables and LEAP 21 Scores*

Variable	N	ELA	Mathematics	Combined ELA + Math
Repeating 4th grade	590	-.057	-.103*	-.082*
Socioeconomic status	590	.327**	.363**	.363**
Race	590	-.316**	-.341**	-.346**
Gender	590	-.162**	-.020	-.102**
Preschool attendance	507	.079	.025	.057
Professional development	589	.073	.031	.057
Title I school	590	-.220**	-.150**	-.199**

*$p < .05$. two-tailed.

**$p < .01$. two-tailed.

The bivariate relationships between ordered predictor variables and LEAP 21 scores were calculated using Kendall's tau-c. This was selected over Spearman's coefficient of rank correlation because the exact distributions are known for higher value of τ than of ρ, τ is a statistic generally more amenable to manipulation than ρ, and problems resulting from tied values are more readily solved (Ferguson & Takane, 1989). Both the ITBS and the DRA achievement assessments are strongly associated with LEAP 21 scores, while the two teacher-related variables show no more than random association. The approximate significance of both the ITBS and the DRA being at the .001 level indicate that these two predictor variables have the strongest of all the bivariate relationships with the criterion variables. The strength of this relationship overwhelmed the multiple regression model, obscuring other strong predictor variable/criterion variable

relationships. However, the strength of their bivariate relationships was used as the basis for including them in the discussion of the multiple regression model, in spite of the very small contribution they make to that model.

Table 9. *Kendall's Tau-c Associations Between Ordered Variables and LEAP 21 Scores*

Variable	*N*	Value	Asymp. *SE*[a]	Approx. τ^b	Approx. sig.
Teacher's degree					
with ELA LEAP 21 scores	579	.030	.030	1.028	.304
with Mathematics LEAP 21 scores	579	.018	.030	.584	.559
with combined LEAP 21 scores	579	.028	.030	.948	.343
Teacher's experience					
with ELA LEAP 21 scores	580	.061	.028	2.189	.029
with Mathematics LEAP 21 scores	580	.034	.029	1.156	.248
with combined LEAP 21 scores	580	.052	.028	1.821	.069
Iowa Test of Basic Skills					
with ELA LEAP 21 scores	548	.552	.021	26.923	.000
with Mathematics LEAP 21 scores	548	.546	.020	26.752	.000
with combined LEAP 21 scores	548	.587	.019	31.112	.000
Developmental Reading Assessment					
with ELA LEAP 21 scores	525	.624	.029	21.340	.000
with combined LEAP 21 scores	525	.618	.030	20.949	.000

[a]Not assuming the null hypothesis.

[b]Using the asymptotic *SE* assuming the null hypothesis.

Stepwise Linear Multiple Regression Analysis

A stepwise linear multiple regression analysis was first conducted with SPSS statistical analysis software, excluding cases listwise for missing data, using the model of interrelationships (see Figure 1) of the predictor variables: students' grade retention in fourth grade, core total national percentile rank on the third-grade ITBS, final third-grade Developmental Reading Assessment (DRA) scores, socioeconomic status (SES), race, gender, and preschool attendance, teachers' degree level, teachers' experience, and teachers' professional development, and school's Title I status and the criterion variable: LEAP 21 English Language Arts (ELA) score. A significant regression equation was found, $F(6, 446) = 104.617$, $p < .001$, with an R^2 value of .579, and the null hypothesis was rejected. A listing of the coefficients is shown in Table 10. The model summary is show in Table 11.

Table 10. *Regression Analysis of Predictor Variables and LEAP 21 ELA Scores*

Step	Variable	*B*	*SE B*	β
1				
	ITBS core total NPR[a]	1.375	.065	.703**
2				
	ITBS core total NPR[a]	1.053	.075	.539**
	Developmental Reading Assessment	19.326	2.556	.290**
3				
	ITBS core total NPR[a]	1.034	.075	.529**
	Developmental Reading Assessment	18.791	2.540	.282**

Step	Variable	*B*	*SE B*	β
	Title I	-9.567	3.184	-.095**
4				
	ITBS core total NPR[a]	1.040	.074	.532**
	Developmental Reading Assessment	18.583	2.511	.279**
	Title I	-11.350	3.191	-.113**
	Teacher professional development	10.750	3.165	.107**
5				
	ITBS core total NPR[a]	1.083	.075	.554**
	Developmental Reading Assessment	18.489	2.493	.277**
	Title I	-12.238	3.184	-.122**
	Teacher professional development	11.266	3.148	.112**
	Repeating 4th grade	13.648	4.966	.088**
6				
	ITBS core total NPR[a]	1.088	.074	.557**
	Developmental Reading Assessment	18.717	2.480	.281**
	Title I	-11.327	3.186	-.113**
	Teacher professional development	11.239	3.129	.111**
	Repeating 4th grade	13.878	4.937	.089**
	Preschool attendance	10.193	4.041	.078**

[a]National Percentile Rank.

**$p < .01$.

Table 11. *Model Summary for English Language Arts LEAP 21 Scores*

Model	*R*	R^2	Adjusted R^2	*SE*
1	.703[a]	.495	.494	35.32
2	.743[b]	.552	.550	33.31
3	.749[c]	.561	.558	33.02
4	.756[d]	.572	.568	32.64
5	.761[e]	.579	.574	32.40
6	.765 [f]	.585	.579	32.21

[a]Predictors: ITBS.

[b]Predictors: ITBS, DRA.

[c]Predictors: ITBS, DRA, Title I.

[d]Predictors: ITBS, DRA, Title I, professional development.

[e]Predictors: ITBS, DRA, Title I, professional development, repeating fourth grade.

[f]Predictors: ITBS, DRA, Title I, professional development, repeating fourth grade, preschool attendance.

A second stepwise multiple regression analysis was conducted using the same predictor variables and the criterion variable: LEAP 21 Mathematics score. A significant regression equation was found, $F(4,448) = 129.55$, $p < .001$, with an R^2 of .536, and the null hypothesis was again rejected. The model summary is show in Table 12. A listing of the coefficients is shown in Table 13. The student's core total national percentile ranking on the ITBS was once again the strongest predictor of LEAP 21 Mathematics scores. The DRA score added only .022 to the R^2 for Mathematics LEAP 21 scores. The total contribution of all of the other predictor variables combined added only .03 to the

Table 12. *Model Summary for Mathematics LEAP 21 Scores*

Model	*R*	R^2	Adjusted R^2	*SE*
1	.700[a]	.489	.488	30.83
2	.715[b]	.511	.509	30.19
3	.726[c]	.528	.524	29.72
4	.732[d]	.536	.532	29.48
5	.736[e]	.541	.536	29.35

[a]Predictors: ITBS.

[b]Predictors: ITBS, DRA.

[c]Predictors: ITBS, DRA, gender.

[d]Predictors: ITBS, DRA, gender, race.

[e]Predictors: ITBS, DRA, gender, race, teacher professional development.

value of R^2. The positive direction for the relationship with gender indicates that male students had a stronger relationship to higher LEAP 21 Mathematics scores than did female students. This positive direction is the opposite of the direction for the bivariate relationships between gender and LEAP 21 scores. The negative direction for the relationship with race indicates that White students performed better on the LEAP 21 Mathematics subtest than African American students did. The proportion of the variance in the Mathematics LEAP 21 model that could be explained by the predictor variables (.541) was less than the proportion of the variance in the ELA LEAP 21 model that could be explained by the predictor variables (.585).

Table 13. *Regression Analysis of Predictor Variables and LEAP 21 Mathematics Scores*

Step	Variable	*B*	*SE B*	β
1				
	ITBS core total NPR[a]	1.188	.057	.700**
2				
	ITBS core total NPR[a]	1.014	.068	.597**
	Developmental Reading Assessment	10.428	2.317	.180**
3				
	ITBS core total NPR[a]	1.017	.067	.599**
	Developmental Reading Assessment	11.008	2.286	.190**
	Gender	11.030	2.811	.128**
4				
	ITBS core total NPR[a]	.968	.069	.570**
	Developmental Reading Assessment	10.574	2.272	.183**
	Gender	10.943	2.789	.127**
	Race	-8.358	2.876	-.099**
5				
	ITBS core total NPR[a]	.969	.068	.571**
	Developmental Reading Assessment	10.460	2.263	.181**
	Gender	10.678	2.779	.124**
	Race	-9.063	2.881	-.108**
	Teacher professional development	6.295	2.828	.072*

[a]National Percentile Rank.

*$p < .05$.**$p < .01$.

A third stepwise multiple regression analysis used the same predictor variables with a combined total of the LEAP 21 ELA plus the LEAP 21 Mathematics score as the criterion variable. With combined ELA and Mathematics LEAP 21 scores, a significant regression equation was found, $F(6, 446) = 125.813$, $p < .001$, with an R^2 of .629, and the null hypothesis was rejected once again. The model summary is show in Table 14.

Table 14. *Model Summary for Combined ELA and Mathematics LEAP 21 Scores*

Model	*R*	R^2	Adjusted R^2	*SE*
1	.751[a]	.563	.563	57.35
2	.779[b]	.608	.606	54.43
3	.784[c]	.614	.612	54.01
4	.789[d]	.623	.619	53.49
5	.793[e]	.628	.624	53.16
6	.796 [f]	.634	.629	52.80

[a]Predictors: ITBS.

[b]Predictors: ITBS, DRA.

[c]Predictors: ITBS, DRA, teacher professional development.

[d]Predictors: ITBS, DRA, teacher professional development, race.

[e]Predictors: ITBS, DRA, teacher professional development, race, preschool attendance.

[f]Predictors: ITBS, DRA, teacher professional development, race, preschool attendance, repeating fourth grade.

A listing of the coefficients is shown in Table 15. The student's core total national percentile ranking on the ITBS was also the strongest predictor of combined ELA and Mathematics LEAP 21 scores. The DRA score added .045 to the R^2 for combined ELA and Mathematics LEAP 21 scores. The total contribution of all of the other predictor variables combined was less than half that of the DRA score. The fact that teacher professional development was a predictor for both ELA and Mathematics LEAP 21 scores separately gave it a stronger R^2 value than any of the other predictor variables besides ITBS and DRA in the combined ELA and Mathematics LEAP 21 scores model. The negative direction for the relationship with race indicates that White students performed better on the LEAP 21 Mathematics subtest than African American students did. The positive direction of the relationship for repeating fourth grade shows that students who did repeat fourth grade did better on the combined ELA and Mathematics LEAP 21 scores than students who did not repeat fourth grade.

Table 15. *Regression Analysis of Predictor Variables and Combined ELA and Mathematics LEAP 21 Scores*

Step	Variable	*B*	*SE B*	β
1				
	ITBS core total NPR[a]	2.563	.106	.751**
2				
	ITBS core total NPR[a]	2.067	.123	.605**
	Developmental Reading Assessment	29.754	4.177	.256**
3				
	ITBS core total NPR[a]	2.081	.122	.609**

Step	Variable	*B*	*SE B*	β
	Developmental Reading Assessment	29.605	4.145	.254**
	Teacher Professional Development	14.668	5.167	.083**
4				
	ITBS core total NPR[a]	1.985	.124	.581**
	Developmental Reading Assessment	28.747	4.115	.247**
	Teacher professional development	16.425	5.149	.093**
	Race	-16.360	5.250	-.097**
5				
	ITBS core total NPR[a]	1.983	.124	.581**
	Developmental Reading Assessment	28.983	4.090	.249**
	Teacher professional development	16.741	5.118	.095**
	Race	-17.498	5.237	-.104**
	Preschool attendance	17.080	6.648	.074**
6				
	ITBS core total NPR[a]	2.045	.125	.599**
	Developmental Reading Assessment	28.844	4.063	.248**
	Teacher professional development	17.511	5.091	.099**
	Race	-19.094	5.234	-.113**
	Preschool attendance	17.785	6.607	.078**
	Repeating 4th grade	21.775	8.106	.080**

[a]National Percentile Rank.

**$p < .01$.

The fact that a student's core total national percentile ranking on the Iowa Test of Basic Skills (ITBS) was the strongest predictor of Louisiana Educational Assessment Program for the 21st Century (LEAP 21) English Language Arts (ELA) and Mathematics scores validates the choice of the LEAP 21 Tutoring Program using these ITBS scores as a criterion for participation. In all cases, the ITBS score accounts for about half of the predictive value of the models, ranging from $R^2 = .489$ to $R^2 = .563$. The DRA score added .057 to the R^2 for ELA LEAP 21 scores, .022 to the R^2 for Mathematics LEAP 21 scores, and .045 to the R^2 for the combined ELA and Mathematics LEAP 21 scores. Although this was not very much, it was more than all of the other predictive variables combined added to the total predictive models. In other words, all of the predictor variables together added less than .1 to the value of R^2, and therefore added very little, indeed, to any of the models. The inclusion of Title I, along with the exclusion of SES, as a predictor of ELA LEAP 21 scores established aggregate poverty as a stronger influence on academic achievement than individual poverty. The second criterion for inclusion in the LEAP 21 Tutoring Program, having repeated fourth grade, showed a positive relationship. In other words, students who repeated fourth grade were more likely than students who did not repeat fourth grade to predict higher LEAP 21 scores. Preschool was a predictor of ELA LEAP 21 scores and of combined ELA and Mathematics LEAP 21 scores. The final predictor variable that added any value to the predictive model for combined scores was teachers' professional development. It was only significant at the .05 level, and added less than .01 unique variance.

Race was not a predictor of ELA LEAP 21 scores, yet it was just behind ITBS and DRA scores as a predictor for combined ELA and Mathematics LEAP 21 scores. This meant that White students showed a relative strength when both areas were considered together, while African American students showed a weakness in this combination. This may have been a result of the strength of race as a predictor of Mathematics LEAP 21 scores alone. However, it added less than .01 to the unique variance of any part of either model, so it had a great deal less influence than ITBS and DRA scores.

Summary

The multiple regression analysis revealed that the two cognitive predictor variables, ITBS scores and DRA scores, were the most predictive of both ELA and Mathematics LEAP 21 scores and combined ELA and Mathematics LEAP 21 scores. Other predictor variables in the models contributed such a slight amount to the total amount of unique variance that they should almost not be mentioned. However, the strength of the bivariate relationships with most of the predictor variables justifies discussion of their contribution to the predictor models. In addition to the predictive value of the ITBS and DRA scores, the ELA LEAP 21 scores were further predicted by two student characteristics: repeating fourth grade (positive relationship) and attending preschool (positive relationship) and two school characteristics: Title I (negative relationship) and professional development (positive relationship). Mathematics scores were further predicted by two student characteristics: gender (positive for male) and race (positive for White). Combining the ELA and Mathematics LEAP 21 scores caused the Title I and gender variables to drop out leaving ITBS, DRA, race, preschool, and

repeating fourth grade from the student characteristics and professional development from the school characteristics as the final predictive model.

CHAPTER 5

CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

Conclusions

The first research question addressed was: Are the current criteria for being categorized as at-risk for participation in the LEAP 21 Tutoring Program (grade retention in fourth grade or 30th percentile or below on the third grade Iowa Test of Basic Skills [ITBS]) actually related to Louisiana Educational Assessment Program for the 21st Century (LEAP 21) scores? The fact that a student's core total national percentile ranking on the ITBS was the strongest predictor of English Language Arts (ELA) and Mathematics LEAP 21 scores validates the choice of the LEAP 21 Tutoring Program to use these ITBS scores as a criterion for participation. Also, the strong bivariate relationship between the ITBS scores and the LEAP 21 scores supports the use of ITBS scores as a viable indicator of potential risk on the LEAP 21. It must be remembered that most students who scored at the 30th percentile or below on the third grade ITBS also participated in the LEAP 21 Tutoring Program. Therefore, this strong relationship between ITBS scores and LEAP 21 scores may also reflect to some degree on the effectiveness of the tutoring program as well. This study did not compare students who were in the tutoring program and on or below the 30th percentile on the ITBS with those who were not in the tutoring program and on or below the 30th percentile on the ITBS

because of the small numbers of students who did not participate. Students who participated in the LEAP 21 Tutoring Program may have done even more poorly on the LEAP 21 without the tutoring, resulting in an even stronger relationship between LEAP 21 and ITBS scores. Another possibility is that the LEAP 21 Tutoring Program may not be very effective, making little impact on the relationship between LEAP 21 and ITBS scores, and the Matthew effect, which posits that students who do poorly continue to do poorly, may have exhibited itself in this relationship between the two scores. One cannot say from this study whether the size of the predictive value for this variable would be different without the LEAP 21 Tutoring Program.

The second criterion for inclusion in the LEAP 21 Tutoring Program, having repeated fourth grade, showed a positive relationship in the multiple regression analysis for ELA LEAP 21 scores and combined ELA and Mathematics LEAP 21 scores. In other words, all students who repeated fourth grade were more likely than all of the rest of the students who did not repeat fourth grade to predict these two higher LEAP 21 scores. At first glance, this seems to go against the conclusions of most of the literature on this relationship. However, Roderick, et al. (1999) found some immediate positive consequences of retention in their Chicago study. It was in the longer term that students who were retained made less of a learning gain than those students who were promoted regardless of having failed the grade. This may reflect to some degree on the effectiveness of the LEAP 21 Tutoring Program, which may have assisted a large portion of repeating students to do better on the LEAP 21 than they may otherwise have done. It may also reflect a change in the attitude or motivation of students who have experienced the consequences of not passing the LEAP 21 the first time they took it. This change in

attitude or motivation may have enabled them to make more gains in their regular learning experiences rather than the tutoring program being the cause of this positive relationship between repeating fourth grade and LEAP 21 scores. Although different forms are used each year, repeating students have a familiarity with the test, which may have influenced their test-taking strategies or their anxiety levels. Students who are repeating fourth grade may also have passed through the fourth grade slump phenomenon developmentally and were therefore better prepared to use reading to learn rather learning to read.

Examination of the Pearson Product Moment correlation reveals a negative relationship ($p < .05$) between repeating fourth grade and Mathematics LEAP 21 scores. This would indicate that students who did not repeat fourth grade did better on the Mathematics portion of the LEAP 21 than those students who did repeat fourth grade. The research question about whether or not a relationship exists between repeating fourth grade and LEAP 21 scores must be answered in the affirmative, with a positive direction for ELA LEAP 21 scores and a negative direction for Mathematics LEAP 21 scores. In any case, students who repeated fourth grade and participated in the LEAP 21 Tutoring Program would need to be compared in larger numbers than were available in this study with students who repeated fourth grade and did not participate in the LEAP 21 Tutoring Program to say definitively whether the LEAP 21 Tutoring Program was what caused this positive relationship.

The second research question to be addressed was: To what extent are other student characteristics (Developmental Reading Assessment [DRA] score, socioeconomic status [SES], race, gender, and preschool attendance) and school characteristic (teacher

degree level, teacher experience, teacher participation in professional development, and school Title I status) also related to student scores on the LEAP 21? Although it contributed only about a tenth as much as the ITBS score to the predictive model of the multiple regression analysis for the ELA and the combined ELA and Mathematics LEAP 21 scores, the DRA score added more value for R^2 than all of the other predictor variables combined. It was also the most dominant of all of the other variables in relationship to the predictive model for the Mathematics LEAP 21 score, although adding less total value than for the ELA and combined LEAP 21 scores. This means that students' previous academic performances as measured by the ITBS and the DRA, influenced by changeable school and home experiences, were more indicative of future academic performance than the student demographics of gender, race, and poverty, which cannot necessarily be changed. This aligns with Jeanne Chall's Stages of Reading Development, which outlines a progressive reading competency based on one's previous stage of development regardless of personal demographics, and with the National Council of Teachers of Mathematics (NCTM) principle of equity for all students.

For the second predictor variable in this research question, SES, this study supports the findings of Darling-Hammond (1999), Klein et al. (2000), and the U.S. Department of Education (2001a) that poverty is more influential at collective levels than at the individual level. The Title I status of a school reflected the fact that a sufficient percentage of students in that school were in the free- and reduced-federal lunch program to qualify under the Title I guidelines for federal funds. The fact that this variable, along with the preschool variable, was only predictive of the ELA and combined scores and not the Mathematics score makes for interesting speculation about causes. Could community

financial poverty with its consequential lack of available, affordable preschools and community poverty of language development with its use of substandard English, lack of reading materials in the home, and scarcity of community language resources such as libraries and transportation to them be more strongly related to the English language arts curriculum than to the mathematics curriculum? Could it be a cultural value that mathematics is more important to survival in an impoverished world than is reading? The Matthew effect continues to be supported with poor communities being associated with poorer academic results among students of those communities.

Preschool attendance was a predictor of ELA LEAP 21 scores and of combined ELA and Mathematics LEAP 21 scores in the multiple regression predictive model, albeit very small (R^2 = .004). However, the bivariate relationship was not significant. The Pearson Longitudinal studies (Campbell & Pungello, 2000; Marcon, 2000; Ramey & Ramey, 1994; Schweinhart & Weikart, 1997) have previously demonstrated the benefits of preschool attendance, but the results of this study add very little to the preponderance of evidence.

The fact that gender was a predictor of Mathematics LEAP 21 scores only, and not the ELA or combined ELA and Mathematics LEAP 21 scores, is the opposite of the trends in the National Assessment of Educational Progress (NAEP). In that assessment, females consistently outperformed males in reading on all three levels. Average scaled scores on the NAEP 2000 Mathematics Assessment showed no significance by gender (U.S. Department of Education, 2001b). However, more males scored at or above *Proficient* than females. Perhaps the use of more abstract strategies by males, as opposed to the more frequent use of modeling and counting strategies by females, observed by

Fennema et al. (1998), served the males better in the LEAP 21 due to its emphasis on problem-solving and higher-level thinking skills.

Race was not a predictor of ELA LEAP 21 scores in the predictive model of the multiple regression analysis, yet it was just behind ITBS and DRA scores as a predictor for combined ELA and Mathematics LEAP 21 scores. This meant that White students tended to score higher than African American students when both areas were considered together. This may have been a result of the strength of race as a predictor of Mathematics LEAP 21 scores. However, it added less than .01 to the R^2 of any part of either model, so it had a great deal less influence than ITBS and DRA scores.

The final independent variable that added any predictive value to any model was professional development. It was not significant in a bivariate relationship with LEAP 21 scores, and added less than .01 to the value of R^2 in the multiple regression analysis. It did influence ELA, Mathematics, and combined LEAP 21 scores, however. Perhaps this could be explained by noting that all of the professional development selected for this study promotes cross-curriculum integration. Techniques that may primarily contribute to more effective learning in one curriculum area may therefore carry over to other areas.

Two other independent variables, teacher degrees (variance = 1.4) and years of teacher experience (variance = 7.42), did not become part of the predictive model. The literature was not conclusive about the relationship between these two variables and academic performance. A situation in the studied school district may have also contributed to the lack of relationship. Only one teacher, the homeroom teacher or the special education teacher in the case of special education students, was identified for each student in the study. This matched the students with the correct teacher, and consequently

his or her degrees and years of experience, in schools where the fourth grade students were taught in self-contained classrooms. Some of the schools used departmentalization in the fourth grade, so that some teachers' degrees and years of experience were matched properly for a student's ELA LEAP 21 scores, his or her Mathematics LEAP 21 scores, or neither, in the case of the homeroom teacher being the science or social studies instructor. For example, one school had three fourth grade teachers, each of whom had a third of the students in her homeroom class. One of the teachers taught all students mathematics and spelling, another taught reading and language, and the third taught science and social studies as the students rotated from classroom to classroom during the school day. Consequently, one third of those students had correctly identified teachers for ELA LEAP 21 scores, another third had correctly identified teachers for Mathematics LEAP 21 scores, and the last third had no correctly identified teacher. A different data gathering technique would have more adequately assessed these two variables. The question would still remain about which teacher's demographic data should be listed for the combined ELA and Mathematics LEAP 21 scores. This lack of refinement may also bring into question the predictive value of the professional development variable as well.

In addition to the problem inherent with proper teacher identification, the variety of comments on the Fourth Grade Teachers Survey by the teachers who took professional training indicated another problem. They described an uneven transfer from learning the theories presented in the professional development to applying them consistently in the classroom. One teacher said that she did not use the theories at all because her current teaching assignment was not in the curriculum area of her training. Another said that it "just depends on the group of students you have as to what types of activities you can do

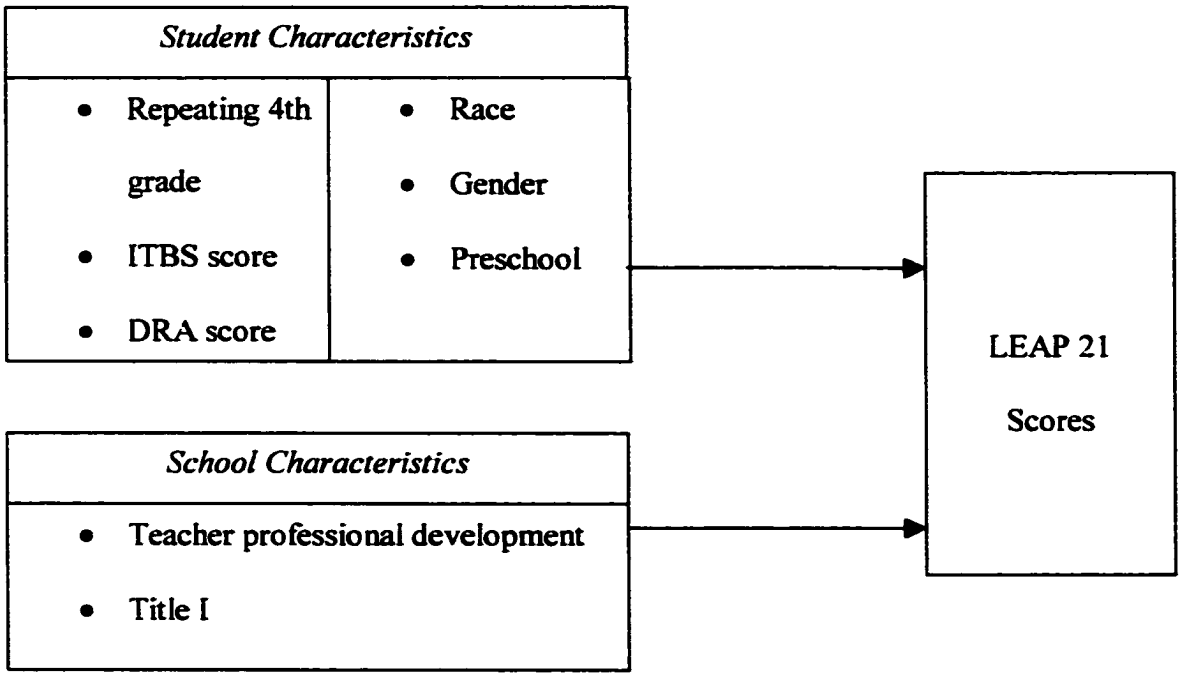
with them." A third stated that she incorporated ideas, methods, activities, or philosophies from her training weekly, that she was better prepared, and that she included more hands-on activities in classroom instruction.

Concerning the second research question about other student and school characteristics that may be related to LEAP 21 scores, data from this study support the conclusion that only DRA scores may be worthy of consideration in addition to the two criteria already selected for the LEAP 21 Tutoring Program. Since this predictive variable contributed more than the others to the predictive model of the multiple regression analysis, and the bivariate relationship between the DRA and LEAP 21 scores was very strong ($p < .001$), students who received a *below level* score on the DRA, but who ranked higher than the 30th percentile on the ITBS, should be considered for inclusion in at-risk programs such as the LEAP 21 Tutoring Program.

The third research question to be considered was: What combination of variables has the strongest predictive value? This study found that the student's national percentile ranking on the ITBS and the student's DRA score was the combination of variables with the strongest predictive value, accounting for .552 R^2 value for ELA LEAP 21 scores, .511 R^2 value for Mathematics LEAP 21 scores, and .608 R^2 value for combined ELA and Mathematics LEAP 21 scores.

A revised model of the predictive variables that this study validated as being predictive of LEAP 21 scores is shown in Figure 2. This contains all predictive variables that were included in the model summaries for ELA LEAP 21 scores, Mathematics LEAP 21 scores, and combined ELA and Mathematics LEAP 21 scores.

Figure 2. Revised interrelationship of variables



Implications

The finding that a student's prior academic achievement was his or her best predictor of future academic achievement rather than race or poverty level showed that education in the South has come a long way since Coleman's landmark study in 1966. No longer were students' demographics the strongest predictors of their academic ability as measured by achievement tests. Rather, students' previously measured academic progress, which has more potential to be affected by variables under the school's control, was the best predictor of performance on the criterion-referenced LEAP 21. However, the students who performed poorly in previous assessments continue to perform poorly, so much more needs to be done in terms of educational equity for all students.

Recommendations Based on Findings

The state of Louisiana should continue using ITBS national percentile scores and repetition of fourth grade as criteria for inclusion in the LEAP 21 Tutoring Program, unless additional research proves specific ineffectiveness by analyzing data from both qualifying participants and qualifying nonparticipants under each criterion for inclusion in the program. Additional efforts should be made to offer professional development opportunities such as the LaSIP training for teachers that promote quality learning experiences for students without any racial or gender bias. Preschool opportunities for all should receive full support regardless of the poverty level of the communities served.

This study revealed a relationship between ITBS scores and LEAP 21 scores in spite of the inclusion of students who participated in the LEAP 21 Tutoring Program. Therefore, the effect of the tutoring program is unclear. Students who participated in the program may have done worse if they had not participated, or they may have done equally as well. Students who repeated fourth grade showed a favorable relationship with LEAP 21 scores. Again, this study does not show whether the relationship is a result of the LEAP 21 Tutoring Program because the number of students in this study who qualified for the LEAP 21 Tutoring Program and did not participate was too small for statistical comparison. One recommendation for further study is to compare scores of students who chose to participate in the LEAP 21 Tutoring Program with those who did not participate, separately by criteria of inclusion. This would need to be done by stratified sampling throughout the state or by random sampling of several school districts to have a large enough sample. A study of this nature would help determine whether the program only helps students who are repeaters, only helps students with low ITBS

national percentile scores, helps both, or helps neither and is a waste of educational resources.

Recommendations for Further Study

Since the relationships between gender and Mathematics LEAP 21 scores were the opposite of those found in the NAEP Assessment, a qualitative study of a few strategically selected mathematics classrooms might reveal whether teachers were transmitting gender stereotypes in mathematics instruction. A quantitative study of types and amounts of feedback to mathematics students by gender might also give insight into this problem. Another approach might be to try to determine what influence community values have on students' performance in mathematics by gender.

Another study could more precisely correlate student LEAP 21 subject matter results with the subject matter teachers. Achievement scores of students whose teachers were participants in intense professional development programs could be analyzed. Matching subjects taught with professional development targeted to that specific curricular area might also lead to understanding relationships between specific types of professional development and their effects on student achievement. Perhaps educational funding directed to the LEAP 21 Tutoring Program could better be spent on professional development programs.

Replicating this study, or conducting a similar one, using a randomly selected sample of students from the whole state, would allow for greater generalizability of the conclusions. Randomly selecting multiple school districts to replicate this study would also permit conclusions to be generalized to a larger population.

REFERENCES

Advanced Systems in Measurement & Evaluation, Inc. (1999). *Louisiana Educational Assessment Program for the 21st Century (LEAP 21): Technical report, phase 1.* Baton Rouge, LA: Louisiana Department of Education.

Baharudin, R., & Luster, T. (1998). Factors related to the quality of the home environment and children's achievement. *Journal of Family Issues, 19,* 375-404.

Baker, J. M. (1995). *IRA's dissemination of past and present knowledge and direct action could be a reading literacy lifeline for many of the "at risk" urban poor.* (ERIC Document Reproduction Service No. ED384865)

Bankston, C. L., III, & Caldas, S. J. (1998). Family structure, schoolmates, and racial inequalities in school achievement. *Journal of Marriage & the Family, 60,* 715-724.

Barron, S. I., & Koretz, D. M. (1994). *An evaluation of the robustness of the NAEP trend lines for racial/ethnical subgroups.* Washington, DC: National Center for Research on Evaluation.

Battista, M. T. (1999). The mathematical miseducation of America's youth. *Phi Delta Kappan.* Retrieved December 13, 2001, from the *Phi Delta Kappan* Web site: http://www.pdkintl.org/kappan/kbat9902.htm

Beard, J. G., Kevan, R., Posrie, R., & Sheridan, V. B. (1990*). The prediction of state student assessment test scores from scores on the districts' standardized norm-referenced tests.* Tallahassee: Florida State University, College of Education.

Bennett, N. G., & Lu, H. (2000). *Child poverty in the states: Levels and trends from 1979 to 1998* (Childhood Poverty Research Brief 2). New York: National Center for Children in Poverty.

Biddle, B. J. (1997). Foolishness, dangerous nonsense, and real correlates of state differences in achievement [Electronic version]. *Phi Delta Kappan, 79*(1). Retrieved December 13, 2001, from http://www.pdkintl.org/kappan/kbid9709.htm

Binkley, M., & Williams, T. (1996). *Reading literacy in the United States* (NCES 96-258). Washington, DC: Office of Educational Research and Improvement.

Boe, E. E., & Barkaric, G. (2000, April). *Critical factors in developing a highly qualified national teaching force.* Paper presented at the meeting of the American Education Research Association, New Orleans, LA.

Borman, G. D., & D'Agostino, J. V. (1996). Title I and student achievement: A meta-analysis of federal evaluation results. *Educational Evaluation and Policy Analysis, 18,* 309-326.

Brookhart, S. M. (1997). Effects of the classroom assessment environment on mathematics and science achievement. *Journal of Educational Research, 90,* 323-331.

Brooks, M. G. (1988, April). *An analysis of variables for predicting the student achievement of the Atlanta public schools.* Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.

Caldas, S. J. (1999). Multilevel examination of student, school, and district-level effects on academic achievement. *Journal of Educational Research, 93,* 91-101.

Caldas, S. J., & Bankston, C., III. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research, 90,* 269-277.

California Budget Project. (2001). *Poor, majority non-White schools less likely to benefit from school performance awards.* Sacramento, CA: Author. Retrieved December 20, 2001, from http://www.cbp.org/press/pr010313.html

Campbell, F. A., & Pungello, E. (2000, June). *High quality child care has long-term educational benefits for poor children.* Paper presented at the Head Start National Research Conference, Los Altos, CA.

Campbell, J. R., & Beaudry, J. S. (1998). Gender gap linked to differential socialization for high-achieving senior mathematics students. *Journal of Educational Research, 91,* 140-148.

Capraro, M. M., Capraro, R. M., & Wiggins, B. B. (2000). *An investigation of the effect of gender, socioeconomic status, race and grades on standardized test scores.* Paper presented at the meeting of the Southwest Educational Research Association, Dallas, TX.

Catsambis, S. (1994). The path to math: Gender and racial-ethnic differences in mathematics participation from middle school to high school. *Sociology of Education, 67,* 199-215.

Center for Business & Economic Research. (n.d.). *Total & per capita personal income and population (CA13) – Louisiana: County – Louisiana.* Retrieved January 8, 2002, from http://leap.nlu.edu/BEA/CASeries.asp

Chall, J. S. (1983). *Stages of reading development.* New York: McGraw-Hill Book Co.

Chall, J. S. (1996). American reading achievement: Should we worry? *Research in the Teaching of English, 30,* 303-310.

Choy, S. P., & Chen, X. (1998). *Toward better teaching: Professional development in 1993-1994* (NECS 98-230). Washington, DC: National Center for Education Statistics.

Cizek, G. J. (1998). Filling in the blanks: Putting standardized tests to the test. *Fordham Report, 2* (11), 1-44. Retrieved July 16, 2001, from Thomas B. Fordham Foundation on-line database: http://www.edexcellence.net

Cizek, G. J., Trent, E. R., Crandell, J., Hirsch, T., & Keene, J. (2000, April). *Research to inform policy: An investigation of pupil proficiency testing requirements and state education reform initiatives.* Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.

Clifford, D., Peisner-Feinber, E., Culking, M., Howes, C., & Kagan, S. L. (1998). Quality care does mean better child outcomes. *NCEDL Spotlights, 2.* Retrieved February 21, 2002, from the National Center for Early Development & Learning Web site at: http://www.fpg.unc.edu/~ncedl/pages/spotlt2.htm

Cole, S., & Watts, S. (2001, March). *Using the ITBS to predict success on LEAP 21.* Paper presented at the meeting of the Louisiana Educational Research Association, Baton Rouge, LA.

Coleman, J. S. (1966). *Equality of educational opportunity* (Report No. OE-3800). Washington, DC: National Center for Educational Statistics.

Council of Chief State School Officers (1996). *The status of state student assessment programs in the United States.* (State student assessment programs database school year 1994-1995). Oak Brook, IL: North Central Regional Educational Lab.

Council of the Great City Schools. (2001). *Striving for excellence: A report on Stanford Achievement Test results in the Great City Schools.* Washington, DC: Harcourt Educational Measurement.

Crowl, T. K. (1996). *Student study guide to accompany Fundamentals of Educational Research: Second edition.* Dubuque, IA: Times Mirror Higher Education Group.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33,* 934-945.

Darling-Hammond, L. (1999). *Teacher quality and student achievement: A review of state policy evidence* (Document R-99-1). Retrieved December 26, 2001, from the Center for the Study of Teaching and Policy Web site: http://www.ctpweb.org

Demps, D. L., & Onwuegbuzie, A. J. (2001). The relationship between eighth-grade reading scores and achievement on the Georgia high school graduation test. *Research in the Schools, 8*(2), 1-10.

Diamond, P. J., & Onwuegbuzie, A. J. (2001). Factors associated with reading achievement and attitudes among elementary school-aged students. *Research in the Schools, 8*(1), 1-11.

Dimitrov, D. M. (1999, April). *Mathematics and science achievement profiles by gender, race, ability, and type of item response.* Paper presented the meeting of the American Educational Research Association, Montreal, Canada.

Doherty, K. M. (2002, January 10). Early learning: The state of the states. Quality Counts 2002 [Special Issue]. *Education Week, 21*(17), 54-67.

Donahue, P. L., Voelkl, K. E., Campbell, J. R., & Mazzeo, J. (1999). *NAEP 1998 reading report card for the nation* (NCES Report No. 1999-459). Washington, DC: National Center for Educational Statistics, Office of Educational Research and Improvement.

Duke, N. K. (2000). For the rich it's richer: Print experiences and environments offered to children in very low- and very high-socioeconomic status first-grade classrooms. *American Educational Research Journal, 17,* 441-478.

Early, D., Clifford, R., & Howes, C., (1999, April). *Quality practices and barriers in early childhood settings: A national survey.* Paper presented at the meeting of the American Research Association, Montreal, Canada.

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfield, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development, 64,* 830-847.

Education Trust (2001). *State summary of Louisiana.* Retrieved November 9, 2001, from the Education Watch on-line database on the World Wide Web: www.edtrust.org

Edwards, V. B., Chronister, G., & Olson, L. (Eds.). (2002, January 10). The state of the states. Quality counts 2002 [Special Issue]. *Education Week,21*(17), 120-121.

Entwisle, D. R., & Alexander, K. L. (1996). Family type and children's growth in reading and math over the primary grades. *Journal of Marriage and the Family, 58,* 341-355.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2001). Keep the faucet flowing: Summer learning and home environment. *American Educator,25*(3), 10-15, 47.

Erickson, R. N., Brunninks, R. H., Elliott, J. I., Graffstrom, P. J., McGrew, K. S., Scott, D. I., & Seppanen, P. S. (1995). *State special education outcomes: A report on how states are assessing educational outcomes for students with disabilities, 1994.* Minneapolis, MN: National Center on Educational Outcomes.

Farkas, G., & Hall, L. S. (2000). Can Title I attain its goal? In D. Ravitch (Ed.), *Brookings papers on education policy 2000.* Washington, DC: Brookings Institution Press.

Fennema, E., Carpenter, T. P., Jacobs, V. R., Franke, M. L., & Levi, L. W. (1998). A longitudinal study of gender differences in young children's mathematical thinking. *Educational Researcher, 27* (5), 6-11.

Ferguson, G. A., & Takane, Y. (1989). *Statistical analysis in psychology and education* (6th ed.). Boston, MA: McGraw-Hill.

Fetler, M. (1999). High school staff characteristics and mathematics test results. *Education Policy Analysis Archives, 7*(9). Retrieved December 23, 2001, from http://epaa.asu.edu/epaa/v7n9.html

Finn, Jr., C. E., & Petrilli, M. J., (Eds.). (1998). *The state of state standards: 2000.* Retrieved July 18, 2001, from the Thomas B. Fordham Foundation Web site: http://www.excellence.net/library/s0ss2000/2000soss.html

Garrison, G. D. (2001). Correlation. *PA 765 Statnotes: An Online Textbook.* Retrieved August 11, 2002, from the North Carolina State University Web site: http://www2.chass.ncsu.edu//garson/pa765/correl.htm

Grissmer, D. W., Kirby, S. N., Berends, M., & Williamson, S. (1994). *Student achievement and the changing American family: An executive summary.* Retrieved November 12, 2000, from the Rand on-line database on the World Wide Web: http://www.rand.org/publications/MR/MR535/MR535.html

Grissmer, D. W., Flanagan, A., Kawata, J., & Williamson, S. (2000). *Improving student achievement: What NAEP state test scores tell us.* Retrieved November 12, 2000, from the Rand Web site: http://www.rand.org/publications/MR/MR924

Haney, W. (2000). The myth of the Texas miracle in education. *Education Policy Analysis Archives, 8*(41). Retrieved December 20, 2001, from http://epaa.asu.edu/epaa/v8n41

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1998, August). *Teachers, schools, and academic achievement* (NBER Working Paper Series No. 6691). Cambridge, MA: National Bureau of Economic Research.

Harkreader, S., & Weathersby, J. (1998). *Staff development and student achievement: Making the connection in Georgia schools.* Atlanta, GA: The Council for School Performance.

Hauser, R. M. (1999, April 7). What if we ended social promotion? [Electronic version] *Education Week, 18*(30), 34-36. Retrieved January 1, 2002, from http://www.edweek.org/ew/ewstory.cfm?slug=30hauser.h18&keywords=Hauser

High, C. F. (1996, October). *A regression analysis of selected factors that influence the scores of students on the TASP test.* Paper presented at the Annual Conference of the Texas Association of College Testing Personnel, Houston, TX.

Hoff, D. J. (2001, December 12). U.S. students rank among world's best and worst readers. *Education Week, 21*(15), 7.

Hoff, D. (2002, January 10). State of the states: Louisiana. *Education Week: Quality Counts 2002, 21*(17), 120.

Howley, C. (1995, November). The Matthew Principle: A West Virginia replication? *Education Policy Analysis Archives, 3*(18). Retrieved December 12, 2001, from http://epaa.asu.edu/epaa/v3n18.html

Howley, C., & Bickel, R. (1999). *The Matthew project: National report.* Randolph, VT: Rural Challenge Policy Program.

Ingersoll, R. M. (2002). *Out-of-field teaching, educational inequality, and the organization of schools: An exploratory analysis.* University of Washington, Center for the Study of Teaching and Policy.

International Reading Association. (2000). *Making a difference means making it different: Honoring children's rights to excellent reading instruction.* Retrieved December 21, 2001, from the International Reading Association Web site: http://www.reading.org/positions/MADMMID.html

Jacobson, J., Olsen, C., King Rice, J., & Sweetland, S. (2001). *Educational achievement and Black-White inequality* (NCES 2001-061). Washington, DC: National Center for Education Statistics, Office of Research and Improvement.

Janiak, R. (1999). Identifying third grade students who are likely to score at Level 1 on the fourth grade FCAT reading test. (Doctoral dissertation, University of South Florida, 1999) *Dissertation Abstracts International, 61,* 124.

Jerald, C. D. (2001). *Dispelling the myth revisited: Preliminary findings from a nationwide analysis of "high-flying" schools.* Retrieved January 20, 2002, from The Education Trust Web site: http//www.edtrust.org/documents/DTMreport.pdf

Johnson, J., Duffett, A., Foleno, T., Foley, P., & Farkas, S. (2001, February 20). Public Agenda reality check 2001. *Education Week, 20* (23), S1-S8.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80,* 437-447.

Just for the Kids. (n.d.). *Just for the Kids: School data availability in other states: Louisiana.* Retrieved June 15, 2001 from the Author's Web site: http://www.just4kids.org/states/Louisiana.htm

Kahlenberg, R. D. (1999, March 31). Economic school desegregation. *Education Week, 18*(29). Retrieved December 18, 2001.from the Education Week Web site: http://www.edweek.org/ew/vol-18/29kahlen.h18

Kasely, L. A., Kilburn, M. R., Bigelow, J. H., Caulkins, J. P., Cannon, J. S., & Chiesa, J. R. (2001). *Assessing costs and benefits of early childhood intervention programs: Overview and application to the Starting Early Starting Smart program.* Seattle,WA: Casey Family Programs; Santa Monica, CA: RAND. Retrieved February 12, 2002, from http://www.rand.org/publications/MR/MR1336.1

Kim, J. J., Crasco, L. M., Smith, R. B., Johnson, G., Karantonis, A., & Leavitt, D. J. (2001). *Academic excellence for all urban students: Their accomplishment in science and mathematics.* Washington, DC: National Science Foundation.

Klein, S. P., Hamilton, L. S., McCaffrey, D. F., & Stecher, B. M. (2000). *What do test scores in Texas tell us?* Retrieved November 12, 2000, from the Rand on-line database (#IP-202) on the World Wide Web: http://www.rand.org/publications/IP/IP202

Lee, J. Y. (1998, December). *Using high-stakes test scores to provide outstanding responsive teachers to disadvantaged kids.* Paper presented at The Harvard Invitational Conference on Civil Rights and High-Stakes K-12 Testing, New York.

Lewis, L., Parsad, B., Carey, N., Bartfai, N., & Farris, E. (1999). *Teacher quality: A report on the preparation and qualifications of public school teachers* (NCES 1999-080). U.S. Department of Education. Washington, DC: National Center for Education Statistics, Office of Educational Research and Improvement.

Lopez, O. S. (1995). *The effect of the relationship between classroom student diversity and teacher capacity on student performance: Conclusions and recommendations for educational policy and practice.* Austin, TX: The Strategic Management of the Classroom Learning Enterprise Research Series.

Louisiana Department of Education. (1999). *Louisiana Educational Assessment Program for the 21st Century: 1998-1999 annual report.* Baton Rouge, LA: Louisiana Department of Education, Division of Student Standards and Assessments.

Louisiana Department of Education. (2000a). *Louisiana Educational Assessment Program for the 21st Century: 1999-2000 annual report.* Baton Rouge, LA: Louisiana State Board of Elementary and Secondary Education, Division of Student Standards and Assessments.

Louisiana Department of Education. (2000b). *1998-99 District Composite Report: Webster Parish.* Baton Rouge, LA: Louisiana State Board of Elementary and Secondary Education.

Louisiana Department of Education. (2000c). *150th Annual Financial and Statistical Report: 1998-99.* (Bulletin 1472). Baton Rouge, LA: Louisiana State Board of Elementary and Secondary Education.

Louisiana Department of Education (2001a). *Grade level retention in Louisiana public schools 1997-98 to 2000-01.* Retrieved November 11, 2001, from the Louisiana Department of Education on-line database: http://www.doe.state.la.us

Louisiana Department of Education. (2001b). *LEAP 21 Tutoring Program: Guidelines and application, 2001-2002.* Baton Rouge, LA: Author.

Louisiana Department of Education. (2001c). *2001 Accountability Summary Results: State and District Table.* Retrieved August 8, 2002, from the Louisiana Department of Education on-line database: http://www.doe.state.la.us/DOE/asps/home.asp?1-REPORTC

Louisiana Department of Education. (2002). *2000-2001 District Accountability Summary Table.* Retrieved August 8, 2002, from the Author's on-line database: http://www.doe.state.la.us/DOE/asps/home.asp?1-REPORTD

Marcon, R. A. (1995). Fourth-grade slump: The cause and cure. *Principal, 74*(5), 16-17, 19-20.

Marcon, R. A. (2000, June). *Educational transitions in early childhood, middle childhood, and early adolescence: Head Start vs. public school pre-kindergarten graduates.* Paper presented at the National Head Start Conference, Washington, DC.

Marcon, R. A., Randall, T., & Brooks, C. (1997, April). *Differential impact of preschool models on achievement of inner-city children.* Paper presented at the meeting of the Southeastern Psychological Association, Atlanta, GA.

Mayer, D. P., Mullens, J. E., & Moore, M. T. (2001). *Monitoring school quality: An indicators report* (NCES 2001-030R). Washington, DC: National Center for Education Statistics, Office of Educational Research and Improvement.

McCoy, A. R., & Reynolds, A. J. (1998, August). *Grade retentions and school performance: An extended investigation* (Institute for Research on Poverty Discussion Paper No. 1167-98). Retrieved January 1, 2002, from http://www.ssc.wisc.edu/irp/pubs/dp116798.pdf

McKenna, M. C., Kear, D. J., & Ellsworth, R. A. (1995). Children's attitudes toward reading: A national survey. *Reading Research Quarterly, 30,* 934-956.

Meisels, S. J., Bickel, D. D., Nicholson, J., Zue, Y., & Atkins-Burnett, S. (2001). Trusting teachers' judgments: A validity study of a curriculum-embedded performance assessment in kindergarten to grade 3. *American Educational Research Journal, 38,* 73-95.

Merton, R. K. (1968). The Matthew effect in science. *Science, 159,* 56-63.

Miller-Whitehead, M. (2001, April). *Alabama's education report card 2000: Significant predictors of student achievement at the district and school level.* Paper presented at the meeting of the American Educational Research Association, Seattle, WA.

Moats, L. C., & Smith, C. (1992). Derivational morphology: Why it should be included in language assessment and instruction. *Language, Speech and Hearing Services in Schools, 23,* 312-319.

Mullin, S. P., & Summers, A. A. (1983). Is more better? The effectiveness of spending on compensatory education. *Phi Delta Kappan, 64*(5), 339-347.

National Council of the Teachers of Mathematics, (2000). *Principles for School Mathematics.* Retrieved October 20, 2001, from Author on-line database CD-ROM ISBN 0-87353-480-8 on the Web site: http://standards.nctm.org/document/prepost/copyr.htm

National Education Goals Panel. (2001). *Eighth graders show improvement but achievement gap persists.* Retrieved December, 22, 2001, from the National Education Goals Panel Interactive Data Center via NEGP Access: http://www.negp.gov

National Research Council, (1998). *Preventing reading difficulties in young children.* Commission on Behavioral and Social Sciences and Education. Washington, D.C.: National Academy Press.

National Research Council, (1999a). *How people learn: Brain, mind, experience, and school.* Committee on Developments in the Science of Learning. Washington, D.C.: National Academy Press.

National Research Council. (1999b). Promotion and retention. In Author's *High stakes: Testing for tracking, promotion, and graduation.* Washington, DC: National Academy Press.

National Research Council. (1999c). Toward a theory of action. In Author's *Testing, teaching, and learning.* Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press. Retrieved January 1, 2002, from http://www.nap.edu/openbook/0309065348/html/15.html

North Carolina State Department of Public Instruction. (2000). *The North Carolina state testing results: 1998-1999.* Washington, DC: Division of Accountability/Testing.

Olson, L. (2001, January 11). Finding the right mix. *Education Week, 20* (17), 12-20.

Orfield, G. (2001). *Schools more separate: Consequences of a decade of resegregation.* Retrieved October 20, 2001, from the Harvard University on-line database on the World Wide Web: http://www.law.harvard.edu/civilrights/publications/pressegexs.html

Organization for Economic Co-operation and Development. (2001). OECD country note: Early childhood education and care policy in the United States of America. In Author's *Early childhood education and care.* Retrieved February 19, 2002, from the OECD Web site: http://www.oecd.org/oecd/pages/home/displaygeneral/0,3380,EN-document-176-no-27-20308-0,FF.html

Palmaffy, T. (1999). Title I: Despite the best of intentions. In M. Kanstoroom & C. E. Finn, Jr. (Eds.), *New directions: Federal education policy in the twenty-first century.* Retrieved December 12, 2001, from the Thomas B. Fordham Foundation Web site: http://www.edexcellence.net/library/newdrct.htm

Payne, K. J., & Biddle, B. J. (1999). Poor school funding, child poverty, and mathematics achievement. *Educational Researcher, 28*(6), 4-13.

Phillips, M., Crouse, J., & Ralph, J. (1998). Does the Black-White test score gap widen after children enter school? In C. Jencks and M. Phillips (Eds.), *The Black-White test score gap* (pp. 229-272). Washington, DC: The Brookings Institution.

Pogrow, S. (1999). Title I: Wrong help at the wrong time. In M. Kanstoroom & C. Finn, Jr. (Eds.), *New directions: Federal education policy in the twenty-first century.* Washington, DC: Thomas B. Fordham Foundation.

Popham, W. J. (2001, September 19). Standardized achievement tests: Misnamed and misleading. *Education Week, 21* (3), 46.

Public Agenda, (2002a). *Child care.* Retrieved February 14, 2002, from http://www.publicagenda.org/issues/factfiles_detail.cfm?issue.type=childcare&list=11

Public Agenda. (2002b). *Education: People's chief concerns.* Retrieved January 6, 2002, from http://www.publicagenda.org/issues/pcc.cfm?issue_type=education

Puma, M. J. (1999, April). *The* Prospects *study of educational growth and opportunity: Implications for policy and practice.* Paper presented at the meeting of the American Educational Research Association, Montreal, Canada.

Puma, M. J. (2000). *Exploring new directions: Title I in the year 2000.* Alexandria, VA: Nation School Boards Association.

Rainwater, L., & Smeeding, T. M. (1995). *Doing poorly: The real income of American children in a comparative perspective* (Luxembourg Income Study Working Paper Series No. 127). Washington, DC: National Science Foundation.

Ramey, C. T., & Ramey, S. L. (1994). Which children benefit the most from early intervention? *Pediatrics, 94,* 1064-1067.

Randle, L. (1997). *What is the effect of pre-school attendance on reading achievement at the third grade?* (ERIC Document Reproduction Service No. ED404630)

Rector, R., Johnson, K. A., & Fagan, P. F. (2001). *Understanding differences in Black and White child poverty rates* (Report No. CDA01-04 of The Heritage Center for Data Analysis). Washington, DC: The Heritage Foundation.

Richard, A. (2002, January 9). Cambridge becomes latest district to integrate by income. *Education Week, 21*(16), 11.

Riverside Publishing Company. (1997). *Validity information for the* Iowa Test of Basic Skills (ITBS) *and* Iowa Tests of Educational Development (ITED): *Forms K, L, M.* Itasca, IL: Author.

Robelen, E. W. (2001, November 28). States sluggish on execution of 1994 ESEA. *Education Week, 21*(13), 1, 26-27.

Roderick, M., Bryk, A. S., Jacob, B. A., Easton, J. Q., & Allensworth, E. (1999). *Ending social promotion: Results from the first two years.* Retrieved January 1, 2002, from http://ww.consortium-Chicago.org/Html_web_store_3.0/Html/endsocialpromo.html

Roderick, M., Nagaoka, J., Bacon, J., & Easton, J. Q. (2000). *Update: Ending social promotion: Passing, retention, and achievement among promoted and retained students 1995-1999.* (Data Brief).Chicago, IL: Consortium on Chicago School Research. (ERIC Document Reproduction Service No. ED451300)

Roth, J., Carter, R., Ariet, M., Resnick, M. B., & Crans, G. (2000, April). *Comparing fourth-grade math and reading achievement of children who did and did not participate in Florida's statewide prekindergarten early intervention program.* Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.

Salisbury, J., & Rees, G. (November, 1999). Accounting for the differential attainment of boys and girls at school. *School Leadership & Management, 19,* 403-427.

Sander, W. (2001). Chicago public schools and student achievement. *Urban Education, 36*(1), 27-38.

Schoen, H. L., Finn, K. F., Griffin, S. F., & Fi, C. (2001, April). *Teacher variables that relate to student achievement in a standards-oriented curriculum.* Paper presented at the meeting of the American Educational Research Association, Seattle, WA.

Schweinhart, L. J., & Weikart, D. P. (1997). The High/Scope preschool curriculum comparison study through age 23. *Early Childhood Research Quarterly, 12*(2).117-143.

Smrekar, C., Guthrie, J. W., Owens, D. E., & Sims, P. G. (2001, September). *March toward excellence: School success and minority student achievement in Department of Defense schools.* Washington, D.C.: National Education Goals Panel, Peabody Center for Education Policy.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360-407.

Steen, L. A. (1997, April). *Equalizing expectations, achieving equity.* Paper presented at a workshop sponsored by the American Association for the Advancement of Science, Seattle, WA.

Trusty, J. (2000). High educational expectations and low achievement: Stability of educational goals across adolescence. *Journal of Educational Research, 93,* 356-366.

U.S. Census Bureau. (2001). *Poverty in the United States: 2000* (Report No. P60-214). Retrieved January 4, 2002, from http://www.census.gov/hhes/poverty/poverty00/pov00src.pdf

U.S. Department of Education. (1996a). *Mapping out the National Assessment of Title I: The interim report.* Retrieved March 7, 2002, from the Author's Web site: http://www.ed.gov/pubs/NatAssess/cover.html

U.S. Department of Education. (1996b). *Pursuing excellence: A study of U.S. eighth-grade mathematics and science teaching, learning, curriculum, and achievement* (NCES 97-198). Washington, DC: National Center for Education Statistics.

U.S. Department of Education. (2000a). *The condition of education 2000* (NCES 2000-062). Washington, DC: National Center for Education Statistics, Office of Educational Research and Improvement.

U.S. Department of Education. (2000b). *Participation in education: Preprimary education.* Retrieved February 18, 2002, from the National Center for Education Statistics Web site: http://nces.ed.gov/programs/coe/2001/section1/tables/t01_1.html

U.S. Department of Education. (2000c). *Promising results, continuing challenges: Final report of the National Assessment of Title I.* Retrieved March 7, 2002, from the Author's Web site: http://www.ed.gov/offices/OUS/PES/exsum.html

U.S. Department of Education. (2000d). *The use of tests as part of high-stakes decision-making for students: A resource guide for educators and policy-makers.* Retrieved July 28, 2001, from the Office for Civil Rights Web site: http://www.ed.gov/offices/OCR/testing/chapter1.html

U.S. Department of Education. (2001a). *The Longitudinal Evaluation of School Change and Performance (LESCP) in Title I schools, Volume 1: Executive summary* (Doc. No. 2001-20). Washington, DC: Office of the Deputy Secretary.

U.S. Department of Education. (2001b). *The nation's report card: Mathematics highlights 2000* (NCES 2001-518). Washington, DC: National Center for Education Statistics, Office of Educational Research and Improvement.

U.S. Department of Education. (2001c). *School improvement report: Executive order on actions for turning around low-performing schools.* Washington, DC: Office of the Under Secretary and Office of Elementary and Secondary Education.

U.S. Department of Education (2001d). Appendix B: Glossary of test measurement terms. In Author's *The use of tests as part of high-stakes decision making for*

students: A resource guide for educators and policy-makers. Washington, DC: Office for Civil Rights.

Webster, B.J., Young, D. J., & Fisher, D. L. (1999, April). *Gender and socioeconomic equity in mathematics and science education: A comparative study.* Paper presented at the meeting of the American Association of Educational Research, Montreal, Canada.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*(3) 461-481.

Wiley, D., & Yoon, B. (1995). Teacher reports on opportunity to learn: Analyses of the 1993 California Learning Assessment System (CLAS). *Educational Evaluation and Policy Analysis, 17,* 355-370.

Williams, E. J. (1999). *Developmental Reading Assessment: Reliability study.* Retrieved March 1, 2002, from: http://www.pearsonlearning.com/correlation/rsp/DRA.doc

Wong, K. K., & Meyer, S. J. (1998). Title I schoolwide programs: A synthesis of findings from recent evaluation. *Education Evaluation and Policy Analysis, 20,* 115-136.

Yopp, R. H., & Yopp, H. K. (2000). Sharing informational text with young children. *Reading Teacher, 53,* 410-444.

Zemelman, S., Daniels, H., & Bizar, M. (1999). Sixty years of reading research—But who's listening? *Phi Delta Kappan.* Retrieved December 13, 2001, from the *Phi Delta Kappan* Web site: http://www.pdkintl.org/kappan/kzem9903.htm

Zhang, L., & Manon, J. (2000, April). *Gender and achievement—Understanding gender differences and similarities in mathematics assessment.* Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.

Appendix A

Survey of Pre-First Grade Experience

Name of student _____

1. **I would describe my kindergarten experience as:**
   - o A full school year
   - o More than half, but less than a full school year
   - o Less than half a full school year
   - o No kindergarten experience

2. **My kindergarten school day was:**
   - o A full school day
   - o A half school day
   - o Three days a week
   - o No kindergarten experience

3. **I would describe most of my preschool experience as:**
   - o Attending Head Start
   - o Attending a public preschool in an elementary school building
   - o Attending a public preschool in another public school building
   - o Attending a church affiliated preschool
   - o Attending a community preschool
   - o Living at home

4. **My preschool experience at Head Start or a public preschool was:**
   - o Less than one year
   - o More than one year, but less than two years
   - o More than two years
   - o Does not apply

5. **My preschool experience at a church affiliated or community preschool was:**
   - o Less than one year
   - o More than one year, but less than two years
   - o More than two years
   - o Does not apply

Check all that apply on the following item.

6. **My preschool experience (at home or in another setting):**
   - o Involved a regular formal lesson time
   - o Included pencil and paper activities
   - o Provided art activities
   - o Included musical activities such as singing and rhythmic movement
   - o Provided outdoor recreational equipment such as swings, slides, etc.
   - o Included a regular story time at least once a day
   - o Had a regularly scheduled nap time
   - o Included supervised play with other children

Appendix B

Participant Consent Form

The following is a brief summary of the project in which you have been asked to participate. Please read it before you sign the statement below.

TITLE: Predictor Variables of Performance on the Louisiana Educational Assessment of Progress for the 21st Century

PURPOSE OF THE STUDY: To identify students at risk of failing the LEAP 21 test.

PROCEDURE AND INSTRUMENTS: Students will take the Survey of Pre-First Grade Experience. These data will be analyzed along with scores on the third grade Iowa Test of Basic Skills, the Developmental Reading Assessment, other student records on file, teachers' degrees, experience, and participation in LaSIP workshops, and school Title I participation to help identify students at risk of failing the LEAP 21 test.

RISKS/ALTERNATIVE TREATMENTS: There are no risks associated with participation in this study. Participation is voluntary.

BENEFITS/COMPENSATION: None

I, _____, show by my signature that I have read and understood the description of the study, "Predictor Variables of Performance on the Louisiana Educational Assessment of Progress for the 21st Century," and its purpose and methods. I understand that my participation in this research is strictly my choice, and my participation or refusal to participate in this study will not affect my grades. Further, I understand that I may withdraw at any time or refuse to answer any questions without penalty. I understand that I may request the results of this study when it is completed. I understand that my name will not be revealed in any way and my answers on the survey will be confidential. These are my rights related to participation in this study, and no one has asked me to give them up.

_____ _____

Signature of Participant **Date**

CONTACT INFORMATION: The researchers listed below may be reached to answer questions about the research, your rights, or related matters.

June Thomas
Doctoral Student, LEC Consortium
212 Garrett Loop
Dubberly, Louisiana 71024
(318) 371-4458
(318) 377-5657

Dr. Cathy Stockton
Major Professor
College of Education
Louisiana Tech University
Ruston, Louisiana
(318) 257-3229

Members of the Human Use Committee of Louisiana Tech University may also be contacted if a problem cannot be discussed with the researchers:
Dr. Terry McConathy (318) 257-2924, Dr. Mary M. Livingston (318) 257-2292,
Mrs. Deby Hamm (318) 257-2924

Appendix C

Directions for Administering the Survey of Pre-First Grade Experience

1. Have all students who took the LEAP 21 test look at the Participant Consent Form while it is read aloud.

2. Have students print their name in the first blank, sign their name in the second blank, and write the date in the third blank.

3. Have students print their name at the top of the Survey of Pre-First Grade Experience.

4. Read each item aloud and all possible responses. Help may be given to ensure student understanding of each item. Have students mark only one answer in questions 1-5.

5. Students may mark as many items as they believe answer the question in number 6.

6. Return all completed items in the pre-addressed envelope.

7. Thank you very much for your assistance with this project.

Appendix D

Survey of Fourth Grade Teachers

Name _____

School _____

1. Do you feel that your students really knew whether or not they attended preschool? What comments did you hear that made you feel the way you do?

2. Check any of the following training that you have completed:

☐ LaSIP ☐ Intech ☐ Math Their Way ☐ Math a Way of Thinking

☐ Balanced Literacy ☐ Another 1-2 week workshop _____

Do you incorporate ideas, methods, activities, or philosophies from any of these workshops in your lessons?

☐ Daily ☐ Weekly ☐ Monthly ☐ At least once each six weeks ☐ Seldom ☐ Never

Do you feel that participation in these workshops has helped you to be better prepared to help students pass the LEAP 21 test? Please explain.