

University of Louisville

## ThinkIR: The University of Louisville's Institutional Repository

---

Electronic Theses and Dissertations

---

12-2003

### **An exploration of the influence of demographic factors on individual and aggregate student achievement measurements in the Kentucky accountability system.**

Beverly Lee Derrington Moore 1938-  
*University of Louisville*

Follow this and additional works at: <https://ir.library.louisville.edu/etd>

---

#### **Recommended Citation**

Moore, Beverly Lee Derrington 1938-, "An exploration of the influence of demographic factors on individual and aggregate student achievement measurements in the Kentucky accountability system." (2003). *Electronic Theses and Dissertations*. Paper 1003.  
<https://doi.org/10.18297/etd/1003>

This Doctoral Dissertation is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact [thinkir@louisville.edu](mailto:thinkir@louisville.edu).

AN EXPLORATION OF THE INFLUENCE OF DEMOGRAPHIC FACTORS ON  
INDIVIDUAL AND AGGREGATE STUDENT ACHIEVEMENT MEASUREMENTS  
IN THE KENTUCKY ACCOUNTABILITY SYSTEM

By

Beverly Lee Derington Moore  
A.B., Emory University, 1960  
M.S., Emory University, 1964  
M.Ed., University of Louisville, 1985

A Dissertation  
Submitted to the Faculty of the  
Graduate School of the University of Louisville  
in Partial Fulfillment of the Requirements  
for the Degree of

Doctor of Philosophy

Department of Teaching and Learning  
College of Education and Human Development  
University of Louisville  
Louisville, KY

December 2003

Copyright 2003 by Beverly Derington Moore

All rights reserved

AN EXPLORATION OF THE INFLUENCE OF DEMOGRAPHIC FACTORS ON  
INDIVIDUAL AND AGGREGATE STUDENT ACHIEVEMENT MEASUREMENTS  
IN THE KENTUCKY ACCOUNTABILITY SYSTEM

By

Beverly Lee Derrington Moore  
A.B., Emory University, 1960  
M.S., Emory University, 1964  
M.Ed., University of Louisville, 1985

A Dissertation Approved on

November 13, 2003

by the following Dissertation Committee:

---

Stephen K. Miller, Dissertation Director

---

Allan E. Dittmer

---

Karen S. Karp

---

Marco A. Muñoz

---

Joseph M. Petrosko



## DEDICATION

This dissertation is dedicated to the children and youth in the Jefferson County Public Schools, whose families send them off to school with trust and hope, who begin school with enthusiasm and great potential, and who carry away knowledge of themselves and of the world molded by their experiences during their most vulnerable years. May we educators better fulfill our most sacred mission: guiding the development of all the young minds entrusted to our stewardship.

## ACKNOWLEDGEMENTS

I have been most fortunate to have Dr. Stephen K. Miller as a dissertation advisor. His guidance and endless patience have been invaluable. Our many discussions about the results and issues involved were stimulating intellectual experiences that have made this study both more interesting and more valid. I have benefited over and over from his attention to detail and professional approach to both teaching and writing, which enabled me to produce a final draft that passed muster with my committee. I would like to express my great thanks to him and hope to follow his good example as a mentor of developing researchers.

I could not have done this study without the special help and advice from Dr. Joseph Petrosko on the forms, processes, and interpretations of multiple regression, in addition to the background he provided in classes on research design and multivariate analysis. His corrections and clarifications for my statistical analyses were confirming and reassuring. I appreciate his efforts to make sure this work was correct and appropriate.

I would like to thank the other members of my committee, Dr. Allan Dittmer, Dr. Karen Karp, and Dr. Marco Muñoz. They each provided unique insights and required improvements in the design and text of the study that helped widen my understanding of the issues and produced a better document.

Special thanks to Dr. Stephen Daeschner who suggested a topic and provided direction for the work; to the numerous College of Education and Human Development Faculty members who have helped me through the years in various ways and given me warm and enthusiastic encouragement; to my program advisor Dr. Betty Lou Whitford, who helped prepare me to do well-balanced research; to Kelly Ising, who with great kindness would help out with any problem; and to Dr. Wayne Ross, Dr. Bill Bush and the Science Team for their support.

I would also like to thank my colleagues and family members who read earlier drafts and provided critiques of the substance and style from their unique perspectives. Patricia Kannapel brought the advantages of her extensive experience in research on Kentucky educational reform to her careful analysis of my work. She helped me think through the problems and focus on critical issues. Rev. Richard Beal, with expertise in composition and perspectives uncontaminated with jargon, asked critical questions that forced the clarification of my ideas and language. My husband, Dr. Gerald L. Moore, provided careful stylistic editing for the final drafts, as well as taking over even more family and home responsibilities.

I am grateful to other friends and family members who in a variety of ways encouraged me and enabled me to complete this work, especially by forgiving my absence from their lives. My sons Kevin and Colin and daughter Jocelyn have supported my work and cheered me on in this project as well as in other educational endeavors. Their belief in the importance of the work is sustaining.

## ABSTRACT

### AN EXPLORATION OF THE INFLUENCE OF DEMOGRAPHIC FACTORS ON INDIVIDUAL AND AGGREGATE STUDENT ACHIEVEMENT MEASUREMENTS IN THE KENTUCKY ACCOUNTABILITY SYSTEM

Beverly Lee Derington Moore

November 13, 2003

This study investigates the progress of standards-based accountability in eliminating the effects of student background and school composition factors on student achievement and school performance in a large, urban district in Kentucky. The factors included gender, socioeconomic status (SES), ethnicity, mobility, family structure, giftedness, and disability. The school composition variables were the percentages of those factors in the school populations. Each grade and level--elementary, middle, and high school--was analyzed by multiple regression.

At the student level, SES, giftedness, and disability predicted 15 to 36% of the variance in scores. Black was an influential factor on norm-referenced tests but not on criterion-referenced tests.

At the school level, SES, family structure, and mobility rate accounted for 56 to 91% of the variance in aggregate scores. The effects were greater in middle and high schools than in elementary. These findings have implications for creating more equitable and effective schools and accountability systems.

## TABLE OF CONTENTS

	PAGE
ACKNOWLEDGMENTS.....	iv
ABSTRACT.....	vi
LIST OF TABLES.....	ix
CHAPTER	
I. STATEMENT OF THE PROBLEM.....	1
Standards-based Reform .....	1
Description of the Problem.....	6
Purpose of the Study.....	27
Research Questions.....	29
Significance of the Study.....	30
Limitations of the Study.....	33
Definition of Terms.....	36
Summary.....	37
II. REVIEW OF THE LITERATURE.....	39
Introduction.....	39
The Development of Standards-Based Educational Reform.....	41
Equity and Standards-based Reform.....	45
The Effects of Demographic Factors.....	55
Demographic Factors and Accountability.....	88

Summary.....	99
<b>III. METHODOLOGY.....</b>	<b>101</b>
Overview of the Study.....	101
Sources of Data.....	102
Variables in the Study.....	117
Empirical Research Questions.....	123
Statistical Analysis.....	124
Summary.....	132
<b>IV. RESULTS.....</b>	<b>134</b>
Introduction.....	134
Descriptive Context.....	135
Multiple Regression Analyses.....	175
Summary.....	239
<b>V. DISCUSSION AND CONCLUSIONS.....</b>	<b>243</b>
Overview of the Study.....	243
Study Methods.....	246
Discussion.....	248
Recommendations.....	282
Conclusions.....	291
<b>REFERENCES.....</b>	<b>305</b>
<b>APPENDICES.....</b>	<b>327</b>
<b>CURRICULUM VITAE.....</b>	<b>351</b>

## LIST OF TABLES

TABLE	PAGE
1. Number ( <i>N</i> ) of JCPS Students Tested per Grade Level.....	103
2. Academic Index Formula for Reading, Mathematics, Science, and Social Studies.....	114
3. Formula for National Norm Reference Index.....	117
4. Correlations for Independent Variables and Third Grade Student CTBS Reading and Mathematics.....	145
5. Correlations Between Independent Variables and Fourth Grade Student KCCT Reading Scores.....	146
6. Correlations for Independent Variables and Fifth Grade Student KCCT Mathematics Scores.....	148
7. Correlations for Independent Variables and Sixth Grade Student CTBS Reading and Mathematics Tests Scores.....	149
8. Correlations for Independent Variables and Seventh Grade Student KCCT Reading Scores.....	150
9. Correlations for Independent Variables and Eighth Grade Student KCCT Mathematics Scores.....	152
10. Correlations for Independent Variables and Ninth Grade Student	

CTBS Reading and Mathematics Scores.....	153
11. Correlations for Independent Variables and Tenth Grade Student	
KCCT Reading Scores.....	154
12. Correlations for Independent Variables and Eleventh Grade	
Student KCCT Mathematics Scores.....	155
13. Correlations of Students' Background Variables with CATS Achievement Test	
Results for all Grades.....	157
14. Correlation Matrix for Elementary School Independent Variables.....	165
15. Correlations Between Independent Variables and Elementary School	
CATS Achievement Tests and Performance Indices.....	167
16. Correlation Matrix for Middle School Independent Variables.....	168
17. Correlations Between Independent Variables and Middle School	
CATS Achievement Tests and Performance Indices Scores.....	170
18. Correlation Matrix for High School Independent Variables.....	171
19. Correlations Between Independent Variables and High School	
CATS Achievement Tests and Performance Indices Scores.....	173
20. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Third Grade CTBS Reading Student Scores on Student Background Variables.....	178
21. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Third Grade CTBS Mathematics Student Scores on Student Background Variables.....	180



22. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Fourth Grade KCCT Reading Student Scores on Student Background Variables.....	182
23. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Fifth Grade KCCT Mathematics Student Scores on Student Background Variables.....	183
24. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Sixth Grade CTBS Reading Student Scores on Student Background Variables .....	185
25. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Sixth Grade CTBS Mathematics Student Scores on Student Background Variables.....	186
26. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Seventh Grade KCCT Reading Student Scores on Student Background Variables.....	187
27. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Eighth Grade KCCT Mathematics Student Scores on Student Background Variables.....	189
28. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Ninth Grade CTBS Reading Student Scores on Student Background Variables.....	190

29. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Ninth Grade CTBS Mathematics Student Scores on Student Background Variables.....	191
30. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Tenth Grade KCCT Reading Student Scores on Student Background Variables.....	192
31. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Eleventh Grade KCCT Student Mathematics Scores on Student Background Variables.....	194
32. Effect Size for Optimum Regression Models of Student Background Variables on Student-Level CATS Tests at Elementary, Middle, and High School.....	195
33. Unstandardized <i>B</i> Coefficients for Optimal Equation Variables for the Regression of Student Background Variables on CATS Test Scores.....	196
34. Standardized ( $\beta$ ) Coefficients of Optimal Equation Variables from the Regression of Student Background Variables on CATS Test Scores .....	199
35. Standard Deviations of the Student-Level CTBS and KCCT Test Comparing Reading and Mathematics Results.....	201
36. Effect Size for Optimum Regression Models of School Composition Variables on Aggregate CATS Tests and Performance Indices at Elementary, Middle, and High School.....	204
37. Tolerance Values for School Composition Variables.....	206

38. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Third Grade CTBS Reading School Scores on School Composition Variables.....	208
39. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Third Grade CTBS Mathematics School Scores on School Composition Variables.....	209
40. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Fourth Grade KCCT Reading School Scores on School Composition Variables.....	211
41. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Fifth Grade KCCT Mathematics School Scores on School Composition Variables.....	212
42. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Elementary School NRT Index on School Composition Variables.....	213
43. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Elementary School Academic Index on School Composition Variables.....	214
44. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Sixth Grade CTBS Reading School Scores on School Composition Variables.....	216
45. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Sixth Grade CTBS Mathematics School	

Scores on School Composition Variables.....	217
46. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Seventh Grade KCCT Reading School Scores on School Composition Variables.....	219
47. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Eighth Grade KCCT Mathematics School Scores on School Composition Variables.....	220
48. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Middle School NRT Index on School Composition Variables.....	222
49. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Middle School Academic Index on School Composition Variables.....	223
50. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Ninth Grade CTBS Reading School Scores on School Composition Variables.....	225
51. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Ninth Grade CTBS Mathematics School Scores on School Composition Variables.....	226
52. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Tenth Grade KCCT Reading School Scores on School Composition Variables.....	227

53. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of Eleventh Grade KCCT Mathematics School Scores on School Composition Variables.....	228
54. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of High School NRT Index on School Composition Variables.....	230
55. Standardized ( $\beta$ ) and Unstandardized <i>B</i> Coefficients from Forward Linear Regression of High School Academic Index on School Composition Variables.....	231
56. Standardized $\beta$ Regression Coefficients for Optimum Equations on School CATS Test .....	233
57. Standardized $\beta$ Regression Coefficients for Optimum Equations for the School NRT and Academic Indices.....	225

# CHAPTER I

## STATEMENT OF THE PROBLEM

### Standards-based Reform

The history of education in the United States since 1900 has been characterized by recurrent, overlapping cycles of reform. These reform movements have waxed and waned with varying degrees of lasting effect and with a plethora of interwoven social and political purposes. The recent standards-based reform movement in American education can be connected back to the publication of *A Nation at Risk* (National Commission on Excellence in Education, 1983), which called for increased excellence in American education by requiring more rigorous curriculum and higher standards for student graduation and for credentialing teachers. That push was followed by more complex and comprehensive reform proposals designed to improve education by approaches such as restructuring schools and empowering teachers. Concomitantly, through the 1980's and 1990's, the demand for governmental initiation of reform evolved and most recently was incarnated as the No Child Left Behind Act of 2001.

In the current wave of reform, the generally accepted solution to the purported inadequacy of American education is the adoption of an educational system that establishes high academic standards and holds educators accountable for students meeting those standards (Fuhrman, 2001). M. S. Smith and O'Day (1991) advocated states taking the responsibility and initiative to establish systemic reform based on

principles that include establishing high academic standards for all students, government (federal, state, local) support for standards-based curricular materials and professional development, local responsibility for professional practices, involvement of teachers in developing standards and curriculum, and accountability based on assessments that are aligned with those standards. Several state governments have embraced this approach and enacted legislation establishing standards-based systems. The incorporation of standards-based reform principles was evident in the Goals 2000 Educate America Act and the reauthorization of the Elementary and Secondary Education Act of 1965 adopted by the U.S. Congress in 1994 (M. S. Smith, 1995). The No Child Left Behind Act of 2001 (*NCLB*, 2001) creates federal mandates for assessment and accountability.

#### *Kentucky Education Reform*

Kentucky's educational reform legislation incorporates the principles of standards-based reform. In 1990 as a culmination of reform efforts by civic groups and several governors and in response to the Kentucky Supreme Court decision that declared the entire Kentucky educational system inequitable and unconstitutional under the Kentucky Constitution, the Kentucky General Assembly passed the Kentucky Education Reform Act of 1990 (KERA). KERA changed the laws and requirements of Kentucky schools extensively in the areas of finance, governance, and curriculum. The changes in the area of curriculum centered on creating a standards-based system with explicit educational goals and with provisions for holding schools accountable for reaching those goals (Alston et al., 1999).

#### *Assessment and Accountability in Kentucky*

For over 13 years, school districts in Kentucky have been operating under the KERA statutes. As in other states, a major focus of reform has been on holding local school educators responsible for student achievement. Kentucky's accountability system includes the five essential elements for current accountability systems identified by the Southern Regional Education Board: (a) rigorous content standards apply to all students; (b) student achievement is assessed; (c) professional development is aligned with standards and assessment; (d) results are reported publicly; and (e) results lead to rewards, sanctions, and targeted assistance (Watts, Gaines, & Creech, 1998).

The assessment and accountability system used in Kentucky from 1992 until 1998 was the Kentucky Instructional Results Information System (KIRIS). That system evolved and changed during the time it was used, but because of persistent problems and widespread discontent, the legislature replaced it with the Commonwealth Accountability Testing System (CATS) in 1998. CATS required the Kentucky Board of Education to make changes in the assessment and accountability system. The new system was put in place in spring of 1999 (Kentucky Department of Education [KDE], 2002a).

In CATS, academic achievement in reading, mathematics, science, social studies, writing, arts and humanities, and practical living is assessed primarily by the Kentucky Core Content Tests (KCCT). The new system retains some of the major characteristics of the KIRIS system: the KCCT are criterion-referenced tests aligned with the educational standards delineated in the Kentucky Academic Expectations: different subjects are assessed at different grades (for example, reading and science are tested in the fourth grade, mathematics, and social studies in the fifth grade); writing



portfolios are part of the assessment; school improvement is measured by comparing different cohorts of students from year to year rather than by longitudinal comparisons of individual student academic growth from year to year (KDE, 2002a).

### *The Accountability Index*

Although the students take the tests and prepare the portfolios, the teachers and principal of a school are held accountable for the results. To that end, an Accountability Index is computed by combining the results of the KCCT assessments, a national norm referenced test, and non-cognitive (non-academic) factors. It is used to rate schools.

To derive this Accountability Index, the KCCT assessment results are used to create an Academic Index in each subject. Then, those subject indices are combined to create a composite Academic Index for each school. (Note: more detailed information on the construction of the Academic Indices appears in Chapter III.) The school Academic Index is then combined with non-cognitive factors and the results from the norm referenced test, the Comprehensive Test of Basic Skills (CTBS), to create an Accountability Index for each school (KDE, 2001a).

The non-cognitive factors included in the formulas for Kentucky's Accountability Index are attendance, drop out, and retention rates, as well as data on successful transition to adult life. However, those non-cognitive factors and the CTBS results represent only a small portion of the Accountability Index. The greatest weight is given to the KCCT-based Academic Index, which represents 90.25% of the total Accountability Index for elementary schools and 85.5% for middle and high schools (KDE, 2001b).

The Kentucky Board of Education set the goal for all schools in Kentucky to

attain an Accountability Index of 100 on a scale of 0 to 140 points by the year 2014. That scale, which is used for rating test scores and the accountability measurements in Kentucky, arbitrarily designates a score of 100 as *Proficient*. Schools must show a specified level of progress toward this goal every two years. This level of progress is determined by drawing a straight line from its 1998-2000 biennium base line score (minus one standard error of measurement) to the statewide goal of 100 for the year 2014 (minus one standard error of measurement). In addition, at the end of the first biennium in 2002 an assistance line was drawn from the school's first biennium score to 80 on the scale for the year 2014 (KDE, 2001a, see Figure 1). This method established fixed goals for growth for each year and was a significant change from the method used in KIRIS, which set no single long-range score targets.

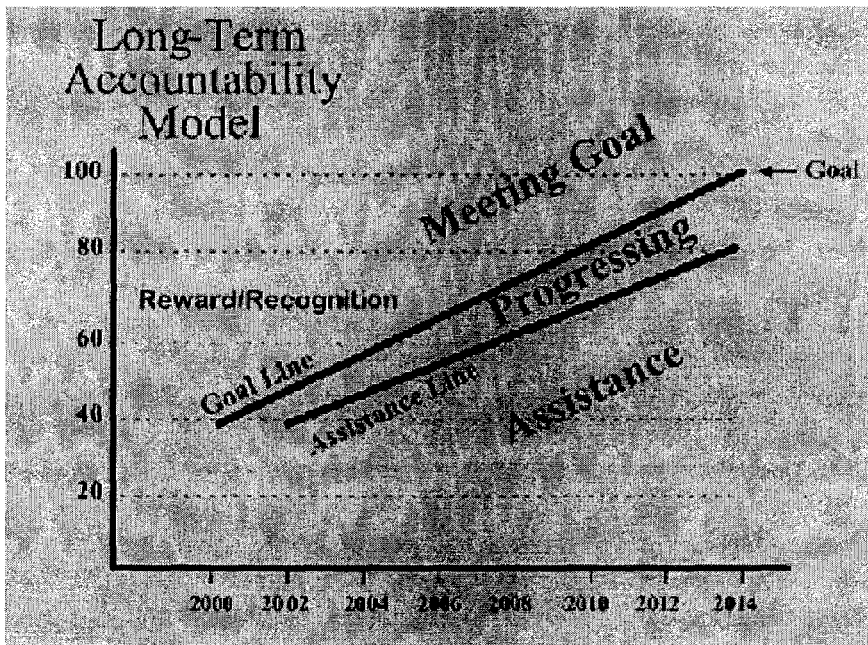


Figure 1. Model showing school improvement goal and assistance lines.

Note: From *CATS 2002 Interpretive Guide*, Version 1.02 [Electronic version], by the Kentucky Department of Education, 2002, p. 20.

Under KIRIS, a school received a new goal each biennium based on improving its previous score by 10% of the difference between its current score and 100. Consequently, if a school did not improve, its goal would remain the same, or if it improved less than its goal, the new goal would still only be 10% of the difference between the most recent score and 100--not a cumulative 20% (Foster, 1999).

CATS is considered a high-stakes system because the state allocates rewards and punishments to the schools (not directly to the students) according to how well schools perform on the assessments. Under CATS, schools are given monetary rewards if their Accountability Index exceeds their growth goal and meets other requirements for lowering the number of dropouts and the percentage of low scoring students. They are given a smaller reward if they exceed their assistance line. Schools with Accountability Indices below their assistance line are subject to sanctions, including state intervention. The indices and extensive disaggregated data are provided to the schools and available for the public. The major Kentucky newspapers publish the results statewide on a yearly basis. The publicity about school scores intensifies the high-stakes nature of the tests for schools and teachers.

The passage of the NCLB Act (2001) added additional high stakes for public school systems that puts further pressure on schools and school districts to raise test scores. The consequences for unsuccessful schools mandated by the NCLB Act include allowing students to transfer to successful schools, requiring use of school funds for public or private supplemental services for low-achieving students, and reconstituting “failing” (p. 2) schools.

#### Description of the Problem

## *The Achievement Gap in Kentucky*

The term *achievement gap* is typically used to denote different levels of achievement between groups of people, especially between ethnicities or between rich and poor. As in other states, a demonstrable achievement gap exists among students and among schools and districts in Kentucky (Guskey, 1997; Muñoz, 2000; Roeder, 1999).

This persisting difference in achievement is antithetical to two of the basic purposes of KERA as dictated by the Kentucky Supreme Court when it declared the entire system of schools in Kentucky unconstitutional and prompted reform legislation: (a) to equalize funding between the rich and poor school districts and (b) to provide all Kentucky students an adequate education (Alston et al., 1999).

Kentucky's accountability system is predicated on the popular credo "All children can learn and nearly all at high levels" (Foster, 1999, p. 20). The assumption underlying the accountability system is that applying high standards to all schools and using standards-based measurements to hold schools accountable will motivate school staff to be more effective and hence enable all students to learn at high levels. It also assumes that the schools have the capacity and power to overcome all previous barriers to students' learning at high levels.

Furthermore, holding schools accountable by the direct measurement of outcomes--student achievement--is a change in the fundamental assumptions of the educational system. It reverses the previous "uncoupling" (Meyer & Rowan, 1978, p. 374) between the activities and structure of schools in which schools were judged by inputs (teacher qualifications, facilities, etc.) rather than from inspection of whether

they had accomplished their purported output, i.e., created the desired outcomes in student achievement (Swanson & Stevenson, 2002). Previous assumptions about the effectiveness of schools were based on a “logic of confidence” paradigm (Meyer & Rowan, p. 357) that schools are properly functioning based on inputs such as the professionalism of the staff, that student achievement is determined by the students, and that teachers have little effect on a school’s achievement level. These assumptions institutionalize lower performance for at-risk students and relieve teachers from accountability. It is precisely this system of assumptions that KERA accountability policies were designed to replace.

A corollary to the belief that all children can learn at high levels is the assumption that any failure of students to reach high achievement levels can be ameliorated by the proper schooling. This is in contrast to the assumption under the logic of confidence that there are different expectations of children according to their presumed ability.

Extensive research supports the position that schools and teachers can make a difference in the achievement of the most disadvantaged students (Brookover, Beady, Flood, Schweitzer, & Wisenbaker, 1979; S. K. Miller, 1985). Studies have shown that several factors under teachers’ control enable previously underachieving students to improve their performance. For example, a beneficial school climate has been identified as a mediating variable in high-achieving disadvantaged schools (Solomon, Battistich, & Hom, 1996). Also, culturally sensitive instructional methods enable previously underachieving minorities to excel (Delpit, 1995; Tharp & Gallimore, 1988), and teachers’ perceptions and attitudes affect students’ learning (Purkey & Novak,

1984; Solomon et al., 1996).

Although there have been demonstrable effects of the standards-based reform on education in Kentucky in test score improvement, the improvement has not been evenly distributed among students or schools (David, 1999; Kannapel, Aagaard, Coe, & Reeves, 2001; Poggio, 2000; Roeder, 1999). At the student level, achievement has improved overall since KERA was implemented, but the achievement gaps that existed before KERA persist. On the KIRIS assessments (which preceded the CATS), females consistently scored higher than males, Whites scored higher than African-Americans, and non-poor scored higher than poor. Also the difference between African-Americans and Whites increased slightly over the six years (D. C. Smith, Neff, & Nemes, 1999). A similar disproportionate improvement shows up on national comparisons. NAEP (National Assessment of Educational Progress) data on Kentucky indicate that overall scores of Kentucky students improved from 1992 to 1998; however, the achievement gap between poor and non-poor students continues, and the gap between African Americans and Whites has increased (Education Trust, 2001).

Comparisons among schools reveal that in general schools with high levels of poverty continue to achieve below schools that serve more economically advantaged students (Poggio, 2000; Roeder, 1999). Moreover, Roeder found that the achievement gap has increased between those schools and districts that were most advantaged and those that were most disadvantaged (based on child poverty rates, state and local revenue, and academic success rates prior to reform) since the implementation of KERA. The persistence of achievement gaps raises questions about the effectiveness of Kentucky's standards-based accountability system in providing the means to ensure

high levels of learning for all children regardless of their social and demographic background.

### *Individual Background Characteristics and Student Achievement*

The correlation of individual background variables with individual achievement test results is extensively documented in the research literature, including such attributes as poverty, ethnicity, mother's education, time students spend working or watching TV, family structure, student mobility (rate entering and exiting school), number of books in the home, and English proficiency (Caldas & Bankston, 1997; Coleman et al., 1966; L. S. Miller, 1995; White, 1982). Several researchers have reported findings that indicate that poverty itself is not the proximal cause of low achievement among individual students, but that variables such as home atmosphere and language experiences (Molfese, DiLalla, & Bunce, 1997) or school climate and student body composition affect individual achievement more directly (cf. Brookover et al., 1979; Caldas & Bankston, 1997; Tharp & Gallimore, 1988). Teacher perceptions and expectations of students are influenced by gender, race, and ethnicity and, consequently, affect instruction and curriculum (Delpit, 1995; McDermott, 1983; Ogbu & Simons, 1998).

Characteristics associated with poverty may serve as mediating factors between poverty and student achievement. Student mobility (changing schools within the same district) can be a function of poverty and family crisis. Also, poor families frequently have only a single parent who cannot provide the support of an intact, two-parent family. There is a higher rate of poverty among African Americans and Hispanics than among European Americans, so that the interplay between ethnicity and poverty is complicated (Huston, McLoyd, & Coll, 1994). And the opposite effect is seen among

children identified as advanced or gifted. They are disproportionately from advantaged families and have access to high quality childcare and other resources.

The association of student achievement with such background variables as poverty, ethnicity, culture, or home environment provides evidence that some children have greater or different needs in order to succeed academically. It is generally acknowledged that for optimum learning, some students will require more help than others and that teachers of children with special needs may require more resources. In addressing the differences in achievement among students, educators have consistently emphasized that children learn neither at the same rate nor in the same ways (Bloom, 1974; Clinchy, 1997; Gardner, 1985; Stipek, 2002).

The Kentucky Board of Education (2001) acknowledged the special needs of some students in its Strategic Plan. The first item listed under “Core Values and Beliefs” is “We believe all children can achieve at high levels, given adequate opportunity and support” (p. 5). However, actual resource allocation requires policy-relevant information far more specific than that imputed in the simple credo that all children can learn. Thus, policy decisions about the allocation of resources and the targeting of support services depend on having accurate measures of the correlation of student background factors with achievement.

The difficulty of separating the effects of individual background factors from school demographics confounds the issue. For example, the Coleman Report (1966) found that personal poverty was the overwhelming predictor of student low achievement and schools make only a small difference in student achievement. However, after the student’s background was taken into account, the social composition



of the school--its racial and socioeconomic (SES) composition--was more highly related to student achievement than any other school factors. A reanalysis of the Coleman Report data by Mayeske and Beaton (1975) found that very little of students' social background could be separated from the school's influence and vice versa. More recent studies (Caldas & Bankston, 1997; Lee, 2002; Sutton & Soderstrom, 1999) have demonstrated strong correlations among the respective individual level and school composite level measures for African American ethnicity, poverty, and social class.

### *School Composition and School Performance*

Although reports of the degree of correlation of poverty or SES with individual student achievement are contradictory--some researchers (White, 1982) not finding a high correlation and others (L. S. Miller, 1995) finding it to be a strong predictor--the substantial effect of school-wide poverty on aggregate student achievement is broadly supported in the research literature (Brookover et al., 1979; Caldas & Bankston, 1997; Mayeske & Beaton, 1975; L. S. Miller; Sutton & Soderstrom, 1999; White). Because of the disproportionate representation of African Americans and Hispanics among those who are poor, the relationships of ethnicity, poverty, and student achievement are commingled. Student mobility rates and proportion of single parent families are other variables confounded by their relationship with poverty (Jennings, Kovalski, & Behrens, 2000; Mao, Whitsett, & Mellor, 1997).

Tharp and Gallimore (1988) note that although individual poor families may read and talk to their children, low-income parents as a group are less likely to do so. Thus, at the individual level, poverty by itself does not seem to determine one's degree of literacy and subsequent academic achievement. However, at the school level, the

proportion of poverty in a school serves as an indirect measure of the proportion of children who do not have home environments conducive to the development of literacy compatible with the dominant culture (in government, universities, commerce, and media).

A common critique of middle class teachers in low-income schools is that they do not understand and appreciate the dialect and cognition of students from different cultures (Delpit, 1995). Tharp and Gallimore (1988) suggest another mechanism for the association of school poverty with student achievement in their description of the development of literacy and thinking. Many poor children have not experienced the language development conversations that, based on Vygotsky's theories, support cognitive development and enable them to be successful in school and other societal settings dominated by middle class values. Instead of high-poverty schools providing interactive conversational experiences that will promote cognitive development, "the schools themselves have adopted the interactional patterns so often attributed to disadvantaged homes," (p. 100) such as controlling and giving orders to the children rather than engaging the children in dialogue and problem solving. This position is supported by the work of Anyon (1981) who found that interaction patterns differed in middle-class versus working-class and under-class schools. Unlike high-poverty schools, middle-class schools regularly provided the type of high-level cognitive discourse that Vygotsky's theories endorse.

Just as gender, poverty, and ethnicity affect the way individual students are perceived and treated by teachers and other students, so student body characteristics such as the prevalence of poverty or students' ethnicity affect the way teachers and

students interact and affect total school climate. For example, the proportion of special education students or the proportion of gifted students can be expected to affect school climate due both to teacher attitudes about those programs and to effects students themselves have on school climate because of their perceptions, attitudes, and behaviors (L. S. Miller, 1995).

### *Addressing School Improvement*

Standards-based accountability defines school effectiveness not by evaluation of inputs into the school program, such as teacher qualifications or instructional methods, but by the school's product, the outcomes of teaching, i.e., student achievement. Valid and accurate measurement of school performance (aggregate student achievement) is an essential component of holding schools accountable. But assessment and accountability are ultimately effective only as the measurements inform and thereby improve instruction and student success.

D. K. Cohen (1995) pointed out that policymakers have relied on standards setting and assessments with rewards and sanctions to change instruction, but the alignment of curriculum, instruction, assessment, teacher education, and professional development has not been coherent. For educational reform to materialize in practice as well as in policy, the nature of teaching must change. However, provisions for discerning and implementing the needed changes are not in place. D. K. Cohen noted that only Vermont and Kentucky included significant professional development in their reform, and that even those states underestimated how much would be needed.

In order for schools to increase student achievement significantly, teachers must be able to identify the barriers to learning, or to effective teaching, and make the needed

changes. For example, most students are assumed to be able to learn at high levels regardless of their backgrounds, yet students who qualify for free and reduced lunches continue to be disproportionately low achieving. Apparently, instructional approaches that are successful with many middle-income students do not produce the same degree of achievement in low-income students (or perhaps the students are not being taught with equivalent methods and expectations). Since the schools and teachers cannot change the students' backgrounds, the teachers must find the instructional approaches that will enable all students to learn at the same high levels.

Logically those schools that are initially low-achieving and have been most ineffective in producing improved student achievement will have to make the greatest changes. In some respects this fact is not given adequate consideration in policy decisions. Whatever the causes of low achievement, the schools have not been able to overcome the barriers to successful learning for their students. The educators in those schools will have to find improvements for instruction, curriculum, and school climate that are more extensive than in high achieving schools in order to be successful with the student populations they serve.

The identification of high standards and assessments are first steps in changing schools by changing what is expected of students and teachers. But establishing standards and revising assessment are not sufficient to change all schools; systemic reform also requires that teachers and school communities be provided the training and support they need to help students meet those high standards. It follows that teachers and principals in low scoring schools with students from disadvantaged backgrounds would need specialized training and help with special methods tailored to the learning

styles and special needs of their students because the traditional methods previously employed have not been effective enough. Unfortunately, according to some critics, the Kentucky assessment system was put into place without the support that teachers needed to learn the instructional approaches that would prepare all students to meet the standards (Appalachia Educational Laboratory, 1994; Clements, 2000). An understanding of the effects of background factors on student and school achievement is preliminary to choosing and implementing more effective instructional approaches.

### *The Steeper Growth Line*

The CATS accountability is based on the assumptions that if the schools are required to produce higher achieving students (a) all schools will be able to reach a CATS score of 100, and (b) they will all be able to do it in the prescribed 14 years between 2000 and 2014. Unavoidably, those Kentucky schools that have the lowest scores on the baseline tests will have a steeper proposed growth line than those that initially score higher. The lower the prior achievement of a school, the greater must be the yearly improvement in score in order for the school to progress toward the ultimate goal of 100 by the year 2014. For example, a school with a baseline of 30 would need to improve 70 points to reach 100 in 14 years or 10 points each biennium, whereas a school with a baseline score of 70 would only need to improve 30 points by 2014 or 4.3 points per biennium.

### *Changing Low-performing Schools*

Defenders of the expectation that all schools achieve a score of 100 by 2014 point to those schools with high proportions of poor students that have achieved above expectations. Case studies and surveys of high-achieving schools with high percentages

of poverty report the importance of principal leadership in promoting a determination to raise expectations and achievement for all children (David, 1999; Kannapel et al., 2001). As the examples of high-scoring high-poverty schools demonstrate, it may be that committed teachers and principals with high expectations and extraordinary effort can reverse those influences. However, analysis of successful high-poverty, high-achieving schools (Kannapel et al.; Solomon et al., 1996) indicates that changing teacher attitudes and instructional methods, which have been shaped by adults' responses to disadvantaged children, requires extraordinary leadership.

But from whence will come the extraordinary school leadership to counter the social and psychological forces that affect the achievement of high-poverty children? Either the preparation of a new generation of principals and teachers, or the reeducation of existing educators, will require an investment of time, money, and expertise that may not be available.

It is possible that reaching an index of 100 by 2014 may be an almost impossible goal for many lower achieving schools. An examination of school success in Kentucky since the advent of KERA by Roeder (1999) indicates that most previously disadvantaged, low-achieving schools continue to be low achieving. Roeder found the percent of poor and minority students was significantly and substantially correlated with low accountability scores. Although there were some high-poverty schools that performed very well, overall schools with higher proportions of poor and minority children consistently had lower levels of performance in the years 1993-97. In this respect, the traditional findings of research on the effects of demographic factors (cf. Coleman Report, 1966; White, 1982) have not yet been substantially altered by the

advent of KERA.

The institution of high standards and accountability was designed to spur higher and more equitable achievement in all schools. Policy makers need to ask why it appears that progress toward that goal is not being achieved in all schools? Discerning causes of the low performance and reasons for the lack of adequate improvement requires careful examination of the characteristics of low-performing and high-performing schools. It is likely that the same factors that were associated with low performance before educational reform still exist and continue to affect the ability of a school to change and improve. M. S. Smith (1995) predicted that the move to standards-based assessments would likely result in an increase in the achievement gap because advantaged children are more likely to have access to the well-trained teachers and other resources that are able to provide the level of curriculum and instruction needed to achieve at high standards. And, of course, in general, students from advantaged families have more home support and resources to meet the increased demands of higher standards.

There are provisions in the reformed Kentucky system aimed at providing the extra help and resources that disadvantaged students need to be more successful in school. These include the Family Resource/Youth Service Centers (FRYSC) that provide social services and eliminate barriers to education at high-poverty schools, Extended School Services (ESS) that provide additional tutoring and teaching for at-risk students, and free preschool for at-risk four year olds and developmentally delayed three year olds. All of these aforementioned programs are directed at remedying barriers attributed to the child's deficits or deficit background.

There are also provisions in the Kentucky law directed at changing schools and upgrading the skills of local teachers and principals. Additional funds were provided by the state legislation for professional development at local schools, and schools are required to develop comprehensive plans to address student achievement. Low-performing schools are provided additional help by highly skilled educators (HSE). These are carefully selected and trained teachers and principals who have been identified as outstanding educators. They are assigned to low-achieving schools (Kannapel & Coe, 2000) to guide the local school personnel in school improvement. The provision for HSE is acknowledgment on the part of the state that some schools need greater resources and guidance to improve instruction and that the simple imposition of standards and assessment is not sufficient.

At issue is whether the HSE and other support for low-performing schools are adequate to effect the radical changes that may be needed for disadvantaged schools to become successful on the same level as more advantaged schools. The demographic conditions that contributed to the lower scores in the first place--such as high concentrations of poverty or high mobility--can be expected to continue to represent barriers to student achievement at a low scoring school. Ecological factors, such as school climate, will continue to affect school achievement unless intentionally changed. The influences of school composition on teacher effectiveness or on school climate are still extant. In a study of reform across the United States in many districts and schools, Elmore and Fuhrman (2001) found that different types of schools responded differently to high-stakes assessment: "High capacity, high-performing schools respond more quickly and more imaginatively than lower-capacity, lower-performing schools" (p. 70).



For many reasons, which might include inadequate resources and preparation as well as student demographics, Kentucky results show generally that those schools that have initially low base line scores, and hence a steeper expected growth line, have not yet been able to attain their growth goals. Although the scores in such schools may have improved as a result of real changes in the school, if the improvement does not reach the established goals, the school is still labeled *unsuccessful*. That label can be disheartening to teachers and students in the school. It is discouraging to further effort and growth to see one's efforts and accomplishments downgraded in such a way (see Calvert, Gaus, & Ruscoe, 2000).

Thomas and Bainbridge (2001) point out that the consequence of the "simplistic 'all children can learn' approach" (p. 661) is that it downplays the need for early intervention for disadvantaged children and provides ineffective motivation.

...There is a widespread attitude that, if students and teachers cannot overcome the obstacles created by poverty and poor nutrition in the short amount of time available in the average school year, they have "failed." The pressure is especially strong when children and their teachers are expected to achieve some arbitrary standard established by a state-mandated proficiency test.

The results of this attitude are that students rarely catch up, and teachers become demoralized. (p. 662)

Test results and accountability indices are published in the newspapers. Unavoidably, schools are compared to each other on their absolute scores and on their progress toward the goal of 100 ("Top and Bottom," 2001). Hence, the poor picture

created by comparison of such struggling schools with other more advantaged schools can reinforce the cycle of failure rather than break it.

### *Alternative Accountability Systems*

Other states have accountability systems that make provisions for demographic differences such as poverty or for previous success when using assessment for accountability. Linn (2001) describes several states and school systems that have developed procedures that take poverty into account in their respective reporting systems. California and Pennsylvania report the absolute rankings on school tests results, but they also report “similar school scores” (p. 12). North Carolina uses a quasi-longitudinal system wherein each school’s growth is compared with the average statewide growth in a benchmark year. Tennessee has a sophisticated system in which matched student-level longitudinal data from several previous years is used to estimate expected gains each year.

There are arguments against making allowances for background or school composition differences. Making allowances for high-poverty schools builds in lower expectations for students in those schools. This basic objection to making allowances is based on the first essential element of educational reform, which mandates high standards for all children. If scores were adjusted to take poverty into account, then the poorer students would not be subject to the same high standards, which would hinder them from ever achieving on a comparable level with advantaged children.

Making adjustments for background factors in assessments or accountability formulas can also be misleading. Linn (2001) reports that in California and Pennsylvania where schools are ranked with similarly scoring schools, the practice may

disguise their ranking on absolute scores.

There is an argument on the opposite side of the absolute scale, one that is rarely articulated. Given the theoretical perspective, a complementary argument for using an absolute scale is this: schools that start off with high initial baselines are held to a *lower* standard of improvement. To reprise the earlier example, the school with an initial baseline of 30 (with the ultimate goal of 100) is required to effect a value-added of 70 points. In contrast, the school with a baseline of 70 is required a value-added effort of only 30. Apparently, faculty from the higher performing school will not have done as much with their students as faculty in the lower performing school. Critics of this position will note that the higher performing school can continue to improve (and be rewarded for it) up to 140 (the maximum on the scale), but no formal expectations exist for moving beyond 100 for advantaged schools. The fact remains that the state has reserved negative sanctions and the label of *failing school* for those who do not reach 100, inevitably the schools with higher concentrations of at-risk children.

This is a true dilemma for which a politically and educationally acceptable solution is not yet apparent. If an *absolute scale* is employed, the higher poverty schools will have to work harder and improve more. If an *equal-improvement* model is used, then the lower achieving, disadvantaged schools will never catch up. What is most disturbing about this dilemma, however, is first, that it is seldom recognized explicitly, and second, that it is accepted as an unavoidable condition so that exploration of alternative approaches to accountability is neglected. A philosophical statement that *all children can learn at high levels* should not be treated as a maxim that stifles informed theoretical and empirically based debate and exploration. Unfortunately, most

proponents of the *all can learn* maxim have done just that by stating that any contrary position is the equivalent of not believing in the potential of all children.

### *Jefferson County Public Schools*

The Jefferson County Public Schools (JCPS) system is unique in several ways. With more than 91,000 students in grades k-12 and 152 school sites, it is by far the largest school district in Kentucky (*Jefferson County District Report Card, 2000-2001, 2001*). The next largest school district, Fayette County, has 32,000 students (*Fayette County District Report Card, 2000-2001, 2000*). Most Kentucky districts include small towns or cities and are predominantly White. (Note: JCPS uses both the terms *Black* and *African-American* to indicate ethnicity. For simplicity, the terms *Black* and *White* are used henceforth in this study.) JCPS serves students from urban Louisville and suburban neighborhoods and has a minority school enrollment (predominantly Black) of about 37%. Like many of the rural districts, it has a high proportion of students (49%) who qualify for free or reduced lunches, but unlike rural counties in the state where most students on free or reduced lunches are White, in JCPS the majority (53%) of free and reduced lunch students are Black (JCPS, 2001a; 2001b; 2001c). JCPS also differs from most of the other districts in having a student assignment plan based on magnet schools and *managed-choice* (JCPS, 2002a, p. 4) that produces racial desegregation of its schools by transportation of students from their home neighborhoods to other parts of the district

### *Differences Among Schools on Student Achievement*

While the managed-choice plan has successfully integrated JCPS, it has not ameliorated achievement differences among schools. Jefferson County Public Schools

are very divergent in their achievement as demonstrated by results on the KCCT, SAT, ACT, and CTBS (JCPS, 2001a, 2001b, 2001c). In 2001, of 20 high schools in Jefferson County, three were in the top ten high schools in the state, while six JCPS schools were in the bottom ten on the KCCT. Of 23 middle schools, one JCPS magnet middle school was in the top ten scoring middle schools in the state, but seven JCPS schools were among the bottom ten middle schools. Of 87 JCPS elementary schools, none were in the top ten, but six were in the bottom ten in the state ("Top and Bottom," 2001). Most recent test results show Jefferson County continues to differ from the rest of the state in having wider achievement gaps between Black and White students, between students eligible for free and reduced lunch and students not eligible, between students with a disability and those without, and between gifted and talented (G&T) students and those not so identified (KDE, 2002b).

In spite of the larger county-wide gaps, however, a recent study of 2002 KCCT results by the Kentucky Association of School Councils (KASC, 2002) showed that some JCPS schools demonstrate high achievement by students on free and reduced lunches and by Black students. At the elementary level four magnet and three regular schools were in the top 10 elementary schools for Black-student achievement on the KCCT (KASC). Except in one of those elementary schools, however, the achievement gaps persisted.

At middle school level, four JCPS magnet schools were among the top four in KCCT results for Black students. Among high schools, five JCPS magnet schools were in the top ten in KCCT results for Black students. One JCPS magnet middle school and five JCPS magnet high schools were in the top ten results for free and reduced lunch

students (KASC, 2002).

### *The JCPS Response to the Achievement Gaps*

The many factors that might affect the disproportionate achievement among JCPS schools include the large size of the district, the large size of its schools, the diversity of its students, the high percentage of urban poverty in contrast to rural poverty, and the managed-choice/magnet system. The district has made various efforts to identify and ameliorate causes of low achievement among students.

During the late 1990s, the JCPS administration focused on low-achieving/underachieving students, first to see if existing compensatory and remedial programs were serving those students and later to target those students for special programs (personal communication, S. O'Daniels, January 27, 2002). In 1999, the district initiated "The Individual Success Plan" (JCPS, 1999c, p. 5), which required teachers to develop specific individualized learning plans for low-achieving students. In 1999, the district made narrowing the achievement gap a specific goal in its consolidated plan (JCPS, 1999a, p. 122): "Reduce the achievement gap in reading, writing, and mathematics by 25% between students of different races and SES groups as measured by KIRIS."

Implementation of these plans included charging and preparing principals to lead professional development in their schools, specifically disaggregating student data, identifying gaps, and devising strategies to improve achievement and eliminate gaps. The administration created a rubric-type document entitled "The Criteria for Academic Improvement" that defined specific expectations about using data to monitor student achievement and address achievement gaps. This document was shared with the

principals at their summer professional development retreat (JCPS, 1999b).

The emphasis on narrowing the achievement gaps increased as the topic received public scrutiny, and JCPS joined the KDE Minority Student Achievement Task Force in March 2000 as one of seven pilot implementation districts to address the issue of minority achievement (JCPS, 2002b). The JCPS superintendent enlisted the Louisville Urban League to join the effort and appointed a Minority Student Achievement Team. The Team adopted a multi-strategy approach to closing the gap, including research to identify successful practices and expansion of the existing safety net (e.g., Title I, tutoring, preschool). They created an *Equity Audit* (p. 15) for evaluating school practices that contribute to minority achievement or gaps. This audit was added to the yearly school *Dialogues* (p. 15). These Dialogues occur once a year at each school in the district. Teams of central office staff and other administrators spend two days at each school observing classrooms; interviewing students, parents, and teachers; and reviewing documentation. At the end of the day, the dialogue team meets with the school representatives (principal, teachers, and parents) to discuss the school's plans for improvement. The team writes a report of its visit for the administration and for the local school.

One aspect of the equity audit was to identify attributes of high-scoring schools, especially those with low gaps. Unfortunately, the district found that the initial efforts to identify distinguishing characteristics of high and low-gap schools were not productive. Therefore, the district continues to explore other forms of data and strategies for eliminating the gaps (JCPS, 2002b).

JCPS researchers Muñoz and Dossett (2001) explored the effects of factors such

as student demographic and social variables, financial variables, teacher characteristics, and student-teacher ratio on school scores. They found that two variables, the proportion of students on free and reduced lunch and the proportion of Black students, accounted for 58% of the variance in achievement test scores over a four-year period.

In another approach, Muñoz and Dossett (2000) created a weighted input index, a *need* factor to include in the school accountability formula, which would compensate for the effects of school composition variables. The index included percentage of students receiving free and reduced lunch price, mobility rate, percent of special education students, and percent of households without two parents. The weighted outcome index included results of the Kentucky tests (KIRIS or KCCT), the percentage of students not academically at-risk, writing portfolio scores, attendance percentage, and the CTBS results. A simple regression analysis indicated that the weighted input index contributed appreciably to the variance in the weighted outcome index. The degree to which each of the separate independent variables might contribute to the variability of the academic index was not addressed nor was the interaction of such variables. Further research is required to ascertain the degree to which each of those variables contributes to student achievement independently.

### Purpose of the Study

In light of the circumstances outlined above and JCPS's commitment to an empirical search for solutions to the problem of achievement gaps, an investigation of the relationship of social and demographic factors with student achievement in this urban district is especially salient. Research is needed to distinguish between the effects of individual background variables and school composition variables on student



achievement to identify more clearly how those effects might be altered.

Kentucky has adopted standards-based reform with school accountability as the strategy for upgrading K-12 education. One of the policy issues that require examination is whether the accountability system itself is accurate and effective in producing the desired changes or whether its goals and timeline are unrealistic and counter productive. Among the problems with the current system are the following:

1. Although there has been some success, the high performance promised by this reform has been inconsistently distributed among children and among schools and characterized by persistent achievement gaps between socio-demographic groups, such as those based on ethnicity or poverty.

2. The extent of the effects of social and demographic background variables on student achievement is not considered in the accountability measures.

3. The effects of student body composition on aggregate achievement, whether by its effects on instruction or on student attitudes, are not addressed.

4. Low-performing schools have a steeper growth line (expected rate of improvement) than high-performing schools. In high-poverty schools, that means large numbers of disadvantaged students must make more rapid progress than advantaged students in other schools. In addition, high-performing schools are not challenged as much as lower performing schools.

5. Educators in low-performing schools, who have not previously been able to produce high-achieving students, are expected to be willing and able to adjust to the greater demands of the steeper growth line on low-performing schools and students. They are expected to do this so quickly that low-performing schools will make faster

improvement than high-performing schools--an expectation that, with a few exceptions, has not materialized.

6. There has been lack of adequate attention to the dilemma of holding all students to high standards of achievement while at the same time making demands that are unattainable and ensure failure.

These problems are all dimensions of the overall policy issue of how best to improve education through accountability. JCPS has addressed this issue in its search for data on the factors affecting student achievement, and it has attempted to create a more equitable accountability formula by the exploration of a need index. However, that index may not be a good reality fit since the included factors were chosen theoretically rather than empirically, and the extent of their effect was only estimated.

The overarching purpose of this study is to provide data that will contribute to an understanding of the interaction of demographic factors and student achievement and that will inform evaluation of accountability formulas and processes. The specific purpose of this study is to ascertain and compare the effects of individual student background factors on individual student achievement and the effects of demographic school composition factors on composite school performance measures.

### Research Questions

The research questions in this study address the relationship of individual background variables with individual student achievement test results and school demographic composition variables with school level achievement in Jefferson County Public Schools in 2002. Analyses that address both individual student achievement and composite school performance are conducted. For both individual and composite

analyses, the respective demographic factors are used as predictor variables.

Each level of schooling--elementary, middle, and high--is analyzed separately, since schools at different levels are subject to different tests, have different characteristics, and serve different ages, all of which can be expected to affect differentially the relationships of background and school demographic variables with achievement.

### *Core Research Questions*

At each level: Elementary, Middle, and High School,

1. To what extent do individual student background factors--gender, SES, ethnicity, mobility, family structure, gifted and talented status, and disability--predict individual student achievement?
2. To what extent do school composition variables--proportions representing gender, SES, ethnicity, mobility, family structure, gifted and talented status, and disability--predict school performance?

The specifics of the methodology addressing the research questions are explained in Chapter III, including more specific empirical research questions, description of the variables, and methods of analysis.

### *Significance of the Study*

The factors affecting student and school success are many, complex, and interrelated. This study addresses the extent of the influence of selected demographic factors on student achievement as a preliminary to further exploration of causes of the correlation of demographic factors with school achievement, the dynamics that diminish educational equity, and the processes by which learning barriers are put in place.

Within the context of school accountability and high standards for all students, there are several potential contributions of this research to the fields of school reform, contextual school effects, and accountability:

1. This study is unique in that it examines the extent to which individual student background factors correlate with student achievement at the individual level compared with the correlation of student demographic composition variables with school performance at the aggregated level. Previous studies on Kentucky reform have addressed only individual or school-level achievement; the unstated but implied presumption is that the effects of background factors are the same at each level. For example, this study demonstrates whether or not individual poverty predicts individual student achievement to the same or a different extent than school-level poverty affects school performance. Policies or practices to ameliorate those different effects may be quite different.

2. Teachers' expanded understanding of the extent of the effects of student background factors on student achievement could help them to focus on possible causes and remedies for specific barriers to learning, thereby also informing their decisions about needs for professional development.

3. This study informs future evaluation of other agencies that might affect individual student achievement. For example, Family Resource and Youth Service Centers are another aspect of education reform in Kentucky. They were especially created to counteract social or environmental factors acting outside the classrooms that constitute background barriers to student achievement, e.g., poverty or instability of residence. Therefore, evaluation of the effectiveness of those extra-classroom programs

or activities requires examination of the extent to which they have removed any negative influences of background variables on individual student achievement.

4. This study expands the work begun by Muñoz and Dossett (2000), which incorporated an estimated need factor into the accountability formula. This study provides empirical statistical measurement of the contribution of school demographic variables to school performance so that a more accurate need factor may be calculated.

5. Consideration of the association of school-level demographic composition variables with the school-level scores contributes to a more complete understanding of the dynamics affecting whole school improvement for practitioners. Identification of the extent of the contribution of school composition variables provides empirical evidence for future research to differentiate the contributions of those variables from other school experiences, such as instruction and curriculum.

6. Assessment of school composition effects on student achievement can provide direction for future research on the identification of mediating influences on student achievement such as interaction with school peers or the influences of teacher attitudes and expectations.

7. Knowledge of the effects of school composition allows a more accurate and equitable comparison of school success among schools for purposes of rewards and sanctions. The findings provide empirical evidence for discussions of alternative approaches to accountability systems and their effects.

8. Knowing the extent of the effects of school population characteristics would predict changes in a school's scores due to changing demographics in the school. For example, if ECE classes are transferred from one school to another or if two schools are

merged, the expected changes in school scores can be anticipated and schools judged accordingly.

9. In contrast to individual student background factors, the school population is somewhat under the control of a school system because it is responsible for the assignment of students to schools. In a managed-choice/magnet student assignment system, the policies and practices of the school district determine the parameters of choice and to varying degrees, the composition of individual schools. Findings regarding effects on student achievement of school composition factors that are within the control of the district can prompt consideration of school assignment policy changes.

10. Concepts and assumptions about school accountability in Kentucky are impacted by this study. The uniform goal of 100 by 2014 for all Kentucky schools puts extraordinary demands on the initially low-achieving schools. It is important to address the question of whether expecting the same results in the same amount of time from all students perpetuates injustice. This study provides evidence for addressing the dilemma of maintaining an accountability system that promotes high standards for all students yet incorporates consideration of background and demographic factors that are beyond the control of the school staff. For example, the quantitative findings could be applied in considering alternative accountability systems.

#### Limitations of the Study

This study is preliminary and addresses only a small portion of the ramifications of the effects of demographic factors on assessment within the context of an accountability system. The limitations of the study include:

1. The study does not address the basic issue of whether accountability is the most beneficial means of effecting school improvement.

2. Students are tested on different subjects in different years, e.g., reading in the fourth grade and mathematics in the fifth grade. Therefore, no student takes all the subject area tests at each grade level. Thus a composite Academic Index is not available for individual students; and the effects on an Academic Index at the individual level cannot be compared with the Academic Index at the school level, which is a school-level composite of the scores of students who take the tests in different subjects at different grades.

3. The study only includes variables available from KDE and JCPS data management systems. Consequently, there are many other aspects of student's lives at home and at school not in this study that could affect student achievement. The neglected areas include (a) variables under some control by the school personnel such as school climate, teacher competence, instructional practices, and teacher expectations; (b) mediating student characteristics such as locus of control, self-concept, self-efficacy, or ambition; and (c) unique contextual circumstances such as local school or community traditions.

4. While this study raises the issue of the problems related to basing comparisons on different cohorts of students instead of on longitudinal studies of student achievement, it does not provide data that pertains to that issue.

5. Free and reduced lunch rate, although frequently used and easily accessible, is a gross measure of poverty based on reported family income. Inaccuracies are commonplace, especially at the high school level where eligible students may not

participate due to stigma. Neither is this measure sensitive to the broad range of indicators of socioeconomic status such as home characteristics, mother's education, or father's occupation that are dimensions of the social stratification system in the United States. Also, the cutoff point that defines eligibility (poverty vs. non-poverty) is at the lower end of the SES spectrum. The dichotomous nature of the variable does not capture the range or diversity of incomes present in the student population. For example, no distinction is made between two *paying* students, one from a family with a \$25,000 income versus another from an affluent family with a six-figure (or higher) income.

6. The designation of students as ECE (Exceptional Child Education, special education) includes a broad and diverse range of student learning disabilities from mild speech impediments to severe mental retardation and is also a gross measure. At the school level, the type of ECE students in a school, such as whether learning disabled or physically handicapped, is not identified and could confound the results.

7. The two variables ECE and G&T identification can both be considered proxies for academic aptitude or previous achievement, yet are analyzed separately which means there are two variables that are correlated with the same construct, aptitude.

8. The measure of family structure--an intact family with two parents versus other arrangements--is based on guardian and student reporting. It is, therefore, subject to diverse interpretations and misrepresentation. As with the free and reduced eligible measurement, it does not reflect the range or variety of home circumstances that might exist within any of those family structures.

9. The study does not eliminate the problems caused by multicollinearity among the variables and does not provide a principal components analysis of the independent



variables, one strategy for dealing with multicollinearity.

10. The study does not address causative or mediating factors. While multiple regression methods can help sort out relationships, longitudinal studies are necessary to identify probable causal chains.

11. The schools in this study are in Kentucky's largest district only. Because much of Kentucky is rural and many counties have few or no minorities, the JCPS data are not representative of schools beyond that district.

12. This study does not address the regression of the overall school Accountability Index directly, nor the additional components of the Accountability Index--attendance, retention, and transition to adult life. Implications with regard to the relationship of the predictor variables to school performance and accountability are focused on the Academic Index and the Norm-Referenced Test Index.

13. Although the CATS has been in place for four years, this study analyzes the test scores only for the year 2002 and does not measure school improvement. Analysis of change scores over a span of multiple years would give a more accurate picture of the value added by the schools.

14. For the school-level analyses, the number of schools in the data set is fewer than the recommended number of predictor variables for reliable multiple regression analysis, especially at the middle and high school levels (see explanation in Chapter III).

15. The regression equations from this study are not validated by comparison with regressions of comparable data from the previous three years that CATS has been in place.

#### Definition of Terms

In the context of policy discussion about education in this study, the following definitions are used:

*Alterable variables*--These are characteristics of students, teachers, or the learning situation that may be changed--altered--either before or during the teaching and learning processes by actions of the professional school staff (Bloom, 1980).

*Highly Skilled Educators*--These teachers, principals, and other educators are specially selected and trained, released from their regular duties, and assigned to help low performing schools make the changes needed to improve in the Kentucky education system.

*Non-alterable givens (characteristics)*--These are the static, stable variables that cannot be changed by the teaching-learning process.

Additional, more technical definitions are provided in Chapter III.

### Summary

In the context of national standards-based educational reform, Kentucky adopted KERA and instituted an accountability system designed to raise student achievement by setting high standards for all children and schools and by assessing the results. This policy is based on the assumptions that being held accountable will motivate practitioners to make the instructional changes necessary for all children to succeed and that schools have the power and capacity to make those changes. The current accountability system, CATS, requires schools to make constant progress toward a score of 100 in the year 2014.

Problems are evident in this accountability system that demands a diverse population of students reach the same high standards at the same time. In spite of

demonstrated improvements in student achievement, achievement gaps (e.g., between rich and poor, Black and White, male and female) persist both at the level of the individual student and among schools. These effects of demographic factors are not taken into account either in comparing individual student achievement or aggregate school-wide performance. Low-achieving students have not been able to catch up with the high-achieving students, and most disadvantaged, low-performing schools are still low performing because educators have not been able to change the effectiveness of those schools substantially. The dilemma of maintaining high standards and compensating for demographic influences on student achievement has not received adequate attention or debate.

This study addresses the extent of the relationship between demographic factors and student achievement both at the level of the individual student and at the level of the school in a large urban/suburban district. The findings contribute to better understanding of the dynamics affecting both underachieving students and low-performing schools. They have implications for policy changes in regard to accountability systems and attempts to improve education.

## CHAPTER II

### REVIEW OF THE LITERATURE

#### Introduction

The question of the effects of background factors and demographics on student achievement has ramifications that extend in many directions and ultimately include all aspects of pedagogy, curriculum, and the socio-cultural role of education. The background factors chosen for this study--gender, race, SES, mobility, family structure, academic giftedness, and disability (Exceptional Child Education, ECE, special education in Kentucky)--are all indicators of one of the three dimensions of the larger economic stratification system in the United States: class, race, and gender (for a definitive explanation, see Dika & Singh, 2002; Persell, 1977, Chapters 1-3). Groups located in lower strata typically do less well on various measures of societal success or performance, including school achievement tests, the target of the current study.

According to critical theorists (Apple, 1985; Bowles & Gintis, 1976; Oakes, 1985; Oakes & Wells, 1998; Persell, 1977; Tyack, 1974), the entire system of education and the larger society, in which education is situated, function to maintain the place of lower SES groups in the lower rungs of society. It legitimates that inequality to members of the society by creating unequal outcomes that justify unequal employment and status. In schools, inequality is perpetuated by the way students are treated and socialized by experiences including (but not limited to) tracking, differential access to

knowledge, and disparate expectations for students of different social and economic status. KERA (as explained in Chapter I) was an attempt to counter the forces that resulted in unequal educational attainment based on economics or school locations.

The essence of KERA's accountability system is a *value-added* model (S. K. Miller, 1992) in which an initial baseline of achievement is established. This baseline presumably represents what the students bring to the school in terms of their composite human capital (L. S. Miller, 1995). All schools are then held accountable for improving student achievement relative to the baseline marker. The schools' faculty/staff are responsible for developing instructional capacity that will result in improved student achievement. Thereby Kentucky schools have been assigned the ambitious goal of reversing, rather than perpetuating, the unequal educational outcomes attributed to the inequities of the economic and social system.

This study is focused on the effects of demographic and background factors as they pertain to policy. Examination of the effectiveness of the Kentucky reform in reversing the confluence of academic achievement with socioeconomic stratification has implications for changing policies regarding accountability. Consequently, this review is limited to the most pertinent aspects of the relationship of those background indicators of economic stratification to student achievement, accountability, and educational reform within Kentucky's Commonwealth Accountability Testing System (CATS). This analysis is intended to help assess the feasibility of eliminating the inequalities among schools by adoption of an accountability system that expects all schools to reach the same standards in the same amount of time under current conditions.

To establish the context for understanding the assumptions underlying Kentucky's reform, this chapter first summarizes the development of standards-based reform with its emphasis on assessment and accountability, nationally and in Kentucky. Next, the chapter focuses on studies of the effects of individual and school demographic factors on student achievement in the United States, Kentucky and the Jefferson County Public Schools. Lastly, the chapter presents existing and proposed approaches to accountability that (a) include accommodation for demographic factors and (b) may provide models for improvements in the Kentucky system.

### The Development of Standards-Based Educational Reform

It is a cliché to point out that Americans expect schools to solve the country's social problems. With that expectation, as society has changed so have the demands upon the schools. Educators' responses to the changing demands in the form of new instructional approaches or revised curriculum have created waves of reforms. In the twentieth century, reform efforts were constant, but the end result seems to be that, overall, schools changed very little (cf. Cuban, 1990; Goodlad, 1984; Perkinson, 1977; Sarason, 1996). Nevertheless, as American education enters the twenty first century, another reform based on standards, assessments, and accountability is widely touted as the solution that finally will result in better schools and higher achieving students.

The current reform movement is based on establishing high academic standards and holding districts, schools, teachers, and/or students accountable for students' attaining those standards. The beginning of this reform movement was signaled by the publication of *A Nation at Risk* in 1983 (National Commission on Excellence in Education) and has manifested most recently in the No Child Left Behind Act of 2001.

*A Nation at Risk* emphasized the lack of--and the need for--excellence in American education in elementary and secondary schools as well as colleges. The report aimed at instituting higher standards in content and in student outcomes. The recommendations included more stringent requirements for high school graduation, upgrading of curriculum and instructional materials, and more demanding testing. Raising standards for teachers consisted of increased requirements for certification and teacher education as well as establishing financial incentives and career ladders to enhance teaching as a career. Initiatives in line with these recommendations constituted the first wave of reform in the mid-1980s. These changes were focused on inputs typically mandated from state governments. Although the reform movement has evolved since the publication of *A Nation at Risk*, the basic tenets of that report--that American schools are ineffective and that the low achievement of American students portends dire consequences for the nation and its economy--continues to be the *raison d'être* of current reform.

The second wave of reforms attempted to change schools by delegating more authority to teachers and schools (and sometimes parents) over curriculum and instruction (Lunenburg, 1992). Fuhrman (2001) notes that many educators recognized that top-down directives would not be effective without changes at the local school level and that teachers are the key to improved achievement. Decentralizing control of schools by creating local school-based decision-making structures was part of this second wave. Other reforms directed at schools included restructuring schools to promote collaboration among teachers and block scheduling to provide more sustained teacher-student contact (Fuhrman).

When increased curricular requirements and restructuring apparently had little success in improving overall student achievement, the focus of reform turned from prescribing inputs to monitoring the outputs by assessments of student achievement (Fuhrman, 2001). This third wave of reform, which has continued to gather momentum into the Twenty-first Century, is distinguished by the clarion call for standards and accountability. The accountability movement rests on holding educators responsible for outputs.

According to Fuhrman (2001), several influences were responsible for the emergence of standards-based reform as the proposed solution to the perceived inadequacy of American education. She pointed out that other countries with higher achievement-test scores had systems that included a clear, national curriculum. Achievement tests were aligned with the curriculum and teachers were prepared to teach it. Also, in this country, the Advanced Placement program and the Title I program were well known models for the alignment of curriculum, professional development, and assessment. Fuhrman reasoned the increased course requirements of earlier reform efforts had not led to greater learning because there was nothing to insure that the content had been learned. The common requirement for minimum competency graduation tests only encouraged teaching for minimum skills (Fuhrman).

The late Al Shanker (1995), the longtime executive director of the American Federation of Teachers, supported the standards movement. He advocated clear, high-level national standards coupled with student accountability. He pointed out that students in other countries must pass high-level tests, which are more demanding than the Advanced Placement tests in the United States, in order to be admitted to college.



He concluded that because all teachers and students know the standards and all teachers teach to those standards, educational systems in those countries are more equitable.

The conflict between the first and second waves of reform was evident. At the same time that there was a growing movement to ensure high quality by centralized control of standards and assessment at the state level, educators were insisting that local control of schools by teachers and principals was necessary for effective change. Standards-based reforms seemed a way to reconcile the earlier efforts at reform by identifying different but complementary roles for the local schools and for states or districts. States or districts would set goals and monitor achievement while responsibility for providing the ways and means to reach those goals would rest with local teachers and schools (Fuhrman, 2001).

In a seminal essay, M. S. Smith and O'Day (1991) presented a thorough, inclusive approach to school reform based on their analysis of research on the effectiveness of educational policy. They proposed that successful reform must be systemic, incorporating both top-down directives and local initiative with roles for policy makers as well as for local district and school practitioners. In their approach, the state has the unique position, authority, and responsibility to provide the coherent leadership, resources, and support necessary for systemic K-12 educational reform. Therefore, the state must establish a unifying vision, system-wide goals and content, curriculum frameworks, and a governance system that ensures educational quality and teacher empowerment. Curriculum frameworks guide the other components necessary for the improvement of instruction, including teacher pre-service education, professional development, and student assessment. Local teachers and schools are

responsible for developing and implementing curriculum and instruction for their schools within the context of broad curriculum frameworks established by the state. The state then holds the schools and school districts accountable by assessments that monitor student achievement and, thereby, stimulate superior instruction.

By the middle of the 1990s, these elements of systemic reform were accepted generally by educators and politicians. M. S. Smith (1995) explains that the Goals 2000: Educate America Act (1994) was based on principles that included establishing high academic standards for all students; government (federal, state, local) support for standards-based curricular materials and professional development; local responsibility for professional practices; involvement of teachers in all processes of developing standards and curriculum; accountability based on assessments that are aligned with the standards; and rewards, corrective measures, and sanctions based on the assessments. In this legislation, states were encouraged to participate voluntarily in testing and in the adoption of national standards.

Under the No Child Left Behind Act (NCLB, 2001), no longer is testing voluntary. To qualify for federal funding, NCLB requires that all states implement yearly reading and mathematics standards-based assessments for all students in grades 3-8. It also requires that the state provide rewards and sanctions based on the assessment results according to a prescribed rating system and that students who attend failing schools must have the opportunity to transfer to successful public schools. With these new federal mandates, it becomes even more crucial that the effects of assessment on students and schools be better understood.

#### Equity and Standards-based Reform

When the first wave of educational reform began in the eighties, one major criticism was that the imposition of uniform standards on children who were neither uniform nor equally prepared for school success was inherently inequitable. Reforms of the second wave, which were directed at increased local decision-making and responsibility, attempted to address that inequity. Local decision-making was expected to encourage teachers to find the best instructional methods to meet the special needs of their students. Involving parents and community members in decision-making would support those efforts because they are close to the students and understand them better than outsiders. Yet, these approaches had no effect on the inequity found in educational outcomes, especially on closing the achievement gaps between ethnic groups and between students of different socioeconomic circumstances (Fuhrman, 2001). However, as Meyer and Rowan (1978) note, local decision making and performance standards pegged to those locales serve to provide flexibility in definitions of what is high achievement, which protects localities from inspection and unflattering comparisons.

In contrast, standards-based reform promises to provide equal educational opportunity by establishing the same high standards for all students and holding schools accountable for all students fulfilling those high standards (Darling-Hammond, 1994; Hornbeck, 1990; M. S. Smith & O'Day, 1991). The principles of this systemic reform assume that schools will be able to overcome background and social influences on student achievement. Theoretically, holding schools responsible will result in schools making the changes in attitudes and approaches that will enable all children to achieve at high levels. Accountability reinforced with rewards for good performance and

punishments for lack of improvement is seen as the key to changing schools. Because accountability no longer allows schools to write off difficult or low-achieving students but instead requires high expectations for all children, it is an engine of equity that will eliminate achievement gaps based on social identifications (Hornbeck, 1990). These assertions are founded on a belief that the schools have the power and capability to enable all children to reach high standards.

The position that schools can make a difference in student achievement is supported by extensive research on exemplary and “effective schools” (Boysen, 1992; Brookover et al., 1979; Levine & Lezotte, 1990), which attributes much of the variance in student achievement to differences in expectations and associated differences in instruction. At question is whether schools can completely eliminate the effects of non-school influences and experiences or whether the school’s power to effect such changes is limited, consistent with the original conclusions from the Coleman Report (1966).

Many educators, including proponents of reform, have warned that standards-based reform may exacerbate inequities if not properly designed and implemented. M. S. Smith and O’Day (1991) warned that unless the standards-based curriculum is supported by common curriculum and common expectations for all children with local resources and flexibility to meet those expectations, the achievement gaps between rich and poor and between minorities and majorities would surely increase. M. S. Smith (1995) explained that the move to standards-based assessments would likely result in an increase in the achievement gap because advantaged children are more likely to have access to the well-trained teachers and other resources that are able to provide the level of curriculum and instruction needed to achieve at high standards.

Linn (2000) makes several points regarding the inequity of the dual emphasis on setting high standards and including all students. He states that requiring all students to reach the same standards in the same amount of time inevitably results in some students not reaching the standards and failing. He points out the type of accountability model used makes a large difference in the scores. SES and prior achievement adjustments can produce non-trivial differences in scores. He notes, however, that adjusting for student background variables can result in lowering expectations for low-income students.

Linn's (2000) work points out the key policy issue with regard to accountability and equity. First, not adjusting for SES, ethnicity, and other demographic factors is unfair; some children come to school less well prepared to succeed in schools than others. These differences in early childhood cognitive and social environment are both real and vast, with resulting gaps in achievement. Further, the conditions that produce these gaps before kindergarten and first grade continue to operate across the span of school years. Second, if the effects of these demographic factors are adjusted through statistical procedures, then expectations for children with those characteristics are concomitantly lowered (whether all of those children are necessarily low-scoring or not), which removes the expectation in the accountability system that they as a group will overcome those initial handicaps and *catch up* to their more advantaged peers. There is an *inherent conflict* between the two positions. Resolving this inherent conflict represents a true dilemma as discussed in Chapter I.

Thomas and Bainbridge (2001) consider it a fallacy to expect all children to learn at the same level and in the same amount of time because some factors are beyond

the control of the school. For example, not all children receive equal nourishment or stimulation in early childhood during the critical period of brain development. They suggest that improving the lives of children before school age would be more effective than proclaiming “all children can learn” without providing “the economic opportunity for families, health care for children, and parenting education for young mothers” (p. 661). Thomas and Bainbridge contend that the simplistic *all children can learn* approach downplays the need for early intervention for disadvantaged children, uses punishment as a motivator (which is not effective), establishes accountability based on arbitrary standards in state-developed tests, and, thereby, perpetuates injustice. They quote Edmund Burke: “The equal treatment of unequals is the greatest injustice of all” (p. 661).

#### Standards-based Reform in Kentucky

Essentially, the question of achieving equity in standards-based assessment systems hinges on alleviating the achievement gaps associated with demographic and social factors such as socioeconomic status and ethnicity. Those effects of demographic and social factors may be direct or indirect functions of individual background, of the composition of school populations, of the testing system, of school practices, or of the greater society. Whatever the causes, the Kentucky system assumes that schools are capable of eliminating the effects of those causes. At issue is whether inequity associated with socioeconomic status and other background factors continues to affect student achievement in Kentucky or whether standards-based reform has mitigated the effects of those factors.

#### *Kentucky Education Reform Act*

In the 1980s civic groups such as the Prichard Committee as well as several governors promoted improvements in Kentucky education, but there were only superficial changes until the adoption of the Kentucky Education Reform Act (KERA) in 1990. In 1985 the Council for Better Education, a coalition of 66 high-poverty Kentucky school districts, filed a complaint against the governor, the Kentucky superintendent of public instruction, the state treasurer, the Senate president pro tempore, the Speaker of the House, and the State Board for Elementary and Secondary Education. The suit alleged that funding was “inadequate and inequitable” (cited in Pankratz, 2000, p. 14) and unconstitutional.

In 1988 the Kentucky Circuit Court ruled in favor of the plaintiffs, finding the “finance system was unconstitutional and discriminatory” (cited in Pankratz, 2000, p. 16). Upon appeal of that decision, the Kentucky Supreme Court went even further saying the state was failing in its duty to provide an equal education for all children and ruled the “entire public school system unconstitutional” (cited in Pankratz, , p. 16).

In their response to the ruling, the governor and legislative leaders seized the opportunity to overhaul the entire Kentucky school system. In 1990, the Kentucky General Assembly passed the Kentucky Education Reform Act (KERA) and a tax package necessary to fund the extensive reforms it created.

KERA extensively changed the laws and requirements of Kentucky schools in the areas of finance, governance, and curriculum, placing Kentucky in the forefront of educational reform. Nationally, it was the most comprehensive and sweeping of the state reform actions--primarily because of its adoption of an outcomes-based accountability system (S. K. Miller, 1992) as the core of the curriculum changes. The

curriculum section includes the standards, assessment, and accountability provisions that are central to this study.

In line with the principles put forth by M. S. Smith and O'Day (1991), KERA requires a standards-based system beginning with goals for schools and student learning and includes an accountability system that measures school success and provides for rewards and sanctions. KERA addresses the five essential elements in current accountability systems described by the Southern Regional Education Board (Watts et al., 1998): rigorous content standards apply to all students; student achievement is assessed; professional development is aligned with standards and assessment; results are reported publicly; and results lead to rewards, sanctions, and targeted assistance.

KERA centered on creating a standards-based system, but it also provides school support for helping teachers and students reach those standards. It expanded professional development opportunities and provided expert help for low scoring schools. To avoid the stigma and long term consequences of early failure by students, KERA prescribed that kindergarten and the first three grades of elementary school be coalesced into a non-graded primary program that would foster continuous progress. That program has not worked out as planned (for the demise of the primary program, see Gnadinger, McIntyre, Chitwood-Smith, & Kyle, 2000).

KERA included several programs designed to counteract or remove barriers to education that are commonly associated with poverty and other background factors. Free preschool is offered for all four year olds who qualify for free and reduced lunches and for children with disabilities three, four, and five years old. The Extended School Services program provides extra help for low-achieving students from teachers during



non-school hours. Family Resource Centers and Youth Services Centers (FRYSC) are available for schools that have at least 20% of students who qualify for free and reduced lunches. The purpose of the centers is to reduce barriers to education that are related to conditions outside the classrooms. Depending on local needs, they provide an array of social and health care services (Alston et al., 1999).

### *Standards, Assessments, and Accountability in KERA*

The content standards for Kentucky schools are codified in Academic Expectations published by the Kentucky Department of Education (Kentucky Department of Education [KDE], 1998). Guidelines that define more specifically what students are expected to know, grade by grade and subject by subject, are available online in the Core Content for Assessment (KDE, 1999). These specific core content standards provide the basis for assessment and accountability. They also provide guidance for teachers in preparing their students for assessment.

KERA required that a performance-based student assessment system be created by 1995-1996 to measure student attainment of the learning goals authentically. The first assessment system developed was the Kentucky Instructional Results Information System (KIRIS), which included both the assessment and the accountability specifications. The components of the assessment were writing portfolios and a KIRIS test that included multiple-choice and open-ended content questions and some performance events. Students were rated as Novice, Apprentice, Proficient, and Distinguished according to their performance on KIRIS tests (Foster, 1999).

KIRIS evolved as educators and the public demanded changes. The original plan was that as performance events and portfolios were developed, they would

gradually replace the more traditional multiple-choice type questions. However, complaints from teachers and the public and evaluations by the Office of Educational Opportunity and Western Michigan University prompted the legislature to commission a full review of the accountability and assessment program (Foster, 1999). The legislative changes that resulted from that review made the system more traditional, rather than more performance-based (Whitford & Jones, 2000).

The criticisms continued, and newly elected Governor Patton appointed another task force to review every aspect of KERA and present recommendations for the 1998 legislative session (Foster, 1999). As a result of that review, the legislature replaced KIRIS with the Commonwealth Accountability Testing System ([CATS] 703 KAR 5:020), which is described in more detail in Chapters I and III.

The major changes in assessment from KIRIS to CATS were the addition of more multiple-choice questions in the tests and the administration of a national-normed achievement test in third, sixth, and ninth grades. The increased number of multiple-choice questions allows broader coverage of the Core Content than was possible in KIRIS. Also, the number of pieces required for the writing portfolios was decreased (Petrosko, 2000).

The Accountability Index in KIRIS included only test results and non-cognitive factors (attendance, retentions, drop-outs, and transition to adult life). The Accountability Index was used to determine each school's progress toward reaching an arbitrary goal of 100 out of a possible 140 points on the accountability scale within 20 years. Each biennium, the accountability index attained by a school became the new threshold and a new goal was calculated (Foster, 1999).

The KIRIS accountability formula became problematic for schools that exceeded their goals in one biennium, which pushed their next growth goal up even higher than it would have been if they had just reached their goals, whereas schools that did not improve continued with an unchanged growth goal. The problems with KIRIS resulted in a revised goal-setting process under CATS, in which each school's growth goal is determined by drawing a straight line from its 1998-99 and 1999-2000 biennium base line score to 100, creating a fixed progression of growth goals (Foster, 1999).

Also in CATS, the Novice and Apprentice scoring levels were further broken down into low and high subcategories. Apparently due to the addition of subcategories, the CATS scores turned out to be generally higher than the KIRIS scores, but the arbitrary goal of 100 points was retained (Petrosko, 2000).

Although accountability in both KIRIS and CATS is based on school improvement, CATS represents a change of emphasis. In the KIRIS accountability system, the emphasis was on change scores (the difference between yearly scores) because expected change scores were used to determine biennial growth goals. Under KIRIS, therefore, the emphasis was on what students learned (value-added by schooling) not on how high were the absolute outcomes (S. K. Miller, 1992).

Under CATS, the expectations continue for schools to raise their scores progressively, and schools are held accountable for yearly improvement. Although CATS still represents value-added improvement, the emphasis has changed because an inflexible line of growth is predetermined, and schools are measured by whether or not they reach the absolute goals that were established in 1999. Schools are evaluated not by how much they improve year to year but by where their yearly absolute scores are

located on their goal line, which is determined by connecting their 2000 score to a score of 100 in 2014 (see Figure 1, p. 5). Also, the assignment of highly skilled educators is based on schools' having low absolute scores, e.g., schools with high absolute scores that do not improve are not assigned highly skilled educators.

### The Effects of Demographic Factors

Much of the research on educational reform in Kentucky was conducted during the time KIRIS was used for accountability and assessment. A brief look at the effects of KERA on student achievement overall is in order before reviewing the influence of demographic factors on student achievement.

#### *Student Achievement in Kentucky*

A review of KERA research (Petrosko, 2000) shows the scores of Kentucky students on the KIRIS assessments increased during the three accountability cycles from 1993 through 1998. High school and elementary students showed substantial increases. Middle school scores increased substantially in mathematics, but not as much in other subjects. In reading, high schools increased 31.2 points, elementary schools increased 26.0 points, but middle schools increased only 8.6 points. In mathematics, high schools increased 24.8 points, elementary schools 22.1 points, and middle schools by 28.6 points.

Poggio (2000) found improvement was different for different grades and different subjects. He reported the greatest improvement was at the elementary level, and the least improvement at the middle school level.

Likewise, examination of the Spring 2002, Kentucky Performance Reports (KDE, 2002b) reveals CATS average Academic Index scores (identified in Chapter I

and described in detail in Chapter III) have increased each of the four years (1999-2002) CATS has been in place. The size of the increases was, however, proportionately much decreased from the changes under KIRIS although the two systems use the same 0 to 140 point scale. Over the four years, reading scores improved 3.0 points in elementary, 3.3 points in middle school, and 4.2 points in high school; mathematics scores improved 8.4 points in elementary, 4.4 points in middle school, and 6.3 points in high school. Under CATS, students also take the norm-referenced CTBS. The average scores on the CTBS Index (a calculated index comparable to the Academic Index, see Chapter III for details) improved each year also. For the four years the improvement was 10.5 points at third grade, 3.6 points at sixth grade, and 4.7 points at ninth grade.

Several studies demonstrated the improvement in the KIRIS scores was not reflected in the results of other assessments: neither in the National Assessment of Educational Progress (NAEP) reading test nor in the ACT college entrance test given during the 1991-1994 period (Petrosko, 2000). However, more recent NAEP reading test results for Kentucky's fourth grade students showed a statistically significant increase in 1998 (National Assessment of Educational Progress, 2003). Poggio (2000) refutes criticism that KIRIS and CATS scores are not substantiated by other measurements because recent NAEP results show increases similar to those in the state assessments; furthermore, the state assessment tests are different from the other tests. ACT is a college entrance exam, measures higher level thinking skills, and is taken only by those students contemplating college, whereas the Kentucky tests measure content from the Kentucky Core Content and are given to all students except for a small percentage of the special needs students who are exempted.

### *Socioeconomic Status and Race/Ethnicity*

Socioeconomic status, race, and ethnicity are so intertwined that usually research studies addressing one also address the other, and they are reviewed together in this section. Socioeconomic status (SES) is a term that includes both the economic conditions of a person or family and their social status based on other factors such as education, occupation, and dialect. There are many kinds of measurements that may be used to represent the various dimensions of socioeconomic status of an individual student or a school. Frequently, parents' occupations or educational level are used as operational definitions of SES, which can be measured by questionnaire (Babbie, 1986). The most frequently used operational measure of economic status (degree of poverty or wealth) in research in Kentucky schools is the eligibility of the student for free or reduced meals as an indication of family income (Guskey, 1997; Kentucky Association of School Councils, 2002; Muñoz, 2000a; Petrosko, 2000; Pitts, 1999; Roeder, 2000)

Race and ethnicity are frequently used interchangeably in the literature although they have different definitions (L. S. Miller, 1995). Ethnicity is generally used in reference to people who share distinctive cultural characteristics such as customs, geographic origin, language, or religion. Race is used to refer to people that share some particular physical traits, such as skin color or physiognomy. However, the definitions of race and ethnicity are controversial and vary in meaning depending on the context and purposes of their use. Frequently, the same term may be used for both race and ethnicity, such as African American.

Rather than attempt to resolve the dilemma of race/ethnicity definitions and meanings, the definitions used by various authors are taken at face value in this review,

and the authors' terms for race/ethnicity are maintained. This study is primarily concerned with the racial definition, the distinction between groups of students identified as *Black* or *White* (or *other*) because that racial distinction is most representative of historical inequities and social stratification in Kentucky. Also, although the numbers of immigrants and immigrants' children are increasing in Kentucky schools, specific ethnic groups still represent very small proportions of students in Jefferson County Schools (see Demographic Factors, Chapter III).

The Coleman Report (1966), based on the Equality of Educational Opportunity Survey in 1965, brought the impact of socioeconomic factors on student achievement to national attention. Coleman et al. reported that personal poverty was the overwhelming predictor of student low achievement and that schools make only a small difference in student achievement. They found, however, that after the student's background was taken into account, the social composition of the school--its racial and socioeconomic (SES) composition--was more highly related to student achievement than any other school factors. Negro students, whose schools generally lacked resources, were more affected by the quality of the school than White students. They also found that a student's attitudes "such as a sense of control of the environment, or a belief in the responsiveness of the environment, are extremely highly related to achievement" (p. 325) and that teacher characteristics influenced achievement but facilities or curriculum did not.

In a later analysis of that survey Mayeske and Beaton (1975) addressed the ways school characteristics affect the achievement of individual students. Using regression analysis and partitioning of the variance, they were able to distinguish the percentage of

school outcomes associated with influences from the school's characteristics (defined by 31 factors including teaching conditions, teacher verbal skills and attitudes, students on free and reduced lunch), with the students' social background, and with both. Social class and type of school explained student achievement nearly equally. They found that very little of the school's influence could be separated from social background and vice versa. Family background and school factors overlapped due to allocation of students to schools on the basis of social background. For minorities, the type of school attended assumed a larger explanatory role than social class; for Whites, the reverse.

One of the major issues at the time of Mayeske and Beaton's (1975) research was whether there were inherent differences in achievement among ethnic groups. They found that the more social background factors were taken into account, the more achievement of different ethnic groups tended to approach a common distribution and the more ethnic group differences in achievement tended to be inseparable from differences in social background. Most if not all student variation in achievement associated with ethnic group membership (i.e., achievement gaps) was explained by differences in family background and type of school attended. The role of family background factors in achievement exceeded that of school factors for all students. Minority students and students in the South tended to show a greater sensitivity to school factors than others. Differences in schools explained the differences in achievement formerly associated with geographic differences.

Although the focus of Mayeske and Beaton's (1975) analysis was on ethnicity, social class, and individual background, their study did include some school characteristics, and they hypothesized that attitudes and motivations were of more



importance than family background and school factors. Subsequently, effective schools research has addressed those areas more intensely.

To address school effectiveness, Brookover et al. (1979) studied the relationship of school outcomes to (1) school input variables, (2) social structure variables and (3) school climate variables. School input variables included a school composition index consisting of socioeconomic status and racial composition, and a personnel index that combined size of student body, average daily attendance, ratio of professional personnel/1000 students, and teacher qualifications, experience, and salary. They defined school climate as consisting of the norms, expectations, and beliefs of the people in the school social system and developed an extensive questionnaire to measure student, teacher, and principal *climates*.

Brookover et al. (1979) found a high correlation between school climate variables and the socioeconomic and racial composition of the school as well as a high correlation of those variables with mean school achievement. Multiple regression analysis revealed that the climate variable contributed 36% of the variance in student achievement in predominantly Black schools and 12% in predominantly White schools over and above the contribution of school composition when composition variables (SES and percent White) were entered into the equation first. When Black/White composition of the student body was controlled in Black and White school samples, the climate variable explained as much variance as SES. If the climate variables were entered first, the composition variables added little to the correlation. The Brookover et al. findings indicate that the differences in achievement between schools were attributable more directly to climate, although climate itself appears to be influenced by

composition. School personnel variables had little effect on achievement except in the case of predominantly Black schools in which it added about 20% to the variance contributed by school composition.

Previously accepted beliefs that SES is a predetermining influence on student achievement was not supported by White (1982) in a meta-analysis on 101 studies. He found that disparate results depended on the unit of analysis, the definition of SES, and the way the results were reported. When the student was the unit of analysis and SES was broken down into various parts, the correlation between student achievement and income was only .315, parent education .185, parent occupation .201; and the mean correlation was .245. However, home atmosphere--which he defined as parents' attitude toward education, their aspirations for their children, and family cultural and intellectual activities--had a correlation of .577. He warned that home atmosphere may be a causal agent or it may be a result of the student's achievement on the family.

With an aggregate unit of study (e.g., school), White (1982) found higher correlations between family background and student achievement. The correlation of student achievement with income was .767, with parents' education .686, and with parents' occupation .586; the mean correlation was .680.

A summary of decades of research is provided by L. S. Miller (1995). He reports achievement test scores are positively correlated with income and parents' education. Black children are more likely to experience long-term poverty. Non-poor students in schools with high concentrations of poverty have lower academic performance than those in schools with low concentrations of poverty and vice versa. He concludes gaps in student achievement are established early in elementary years as a

result of learning gaps before first grade, different rates of progress after starting school, and differences in learning gain/loss when school is not in session. Two additional obstacles to student learning associated with poverty are high student-mobility rates and health problems.

Upon reexamining the Coleman Report and National Education Longitudinal Study of 1988 data, L. S. Miller (1995) confirmed that social class accounts for a large part of variation in achievement. However, within class comparisons indicate that achievement gaps within ethnic groups (Whites, Asians, Blacks, and Puerto Ricans) are large. L. S. Miller looked at NAEP reading and mathematics scores by race/ethnicity and parents' education from 1971 to 1988 and found that achievement gaps were smaller in 1988 than in 1971. Examination of 1990 SAT data showed that the higher the parents' education the higher the student scores, but there was still a gap within education levels among Whites, Blacks, and Hispanics.

*The Black-White Test Score Gap* (Jencks & Phillips, 1998), is an up-dated analysis of some of the issues included in the 1972 book *Inequality: A Reassessment of the Effect of Family and Schooling in America* (Jencks et al., 1972). Jencks and Phillips (1998) describe the persistent Black-White test score gap and present possible causes and implications of the gap. They refute the traditional explanations for the gap—the culture of poverty, the decline of the family (especially the scarcity of two-parent Black families), and genetics. Fourteen contributing authors present alternative explanations for the gap that Jencks and Phillips characterize as essentially based on differences in culture and schooling. Phillips, Crouse, and Ralph (1998) propose that the culture of the home background is defined by a larger set of factors than just income

and education. The factors, which together account for about two-thirds of the test score gap, include household size; neighborhood characteristics; grandparents' education; and mothers' high school quality, self-efficacy, parenting practices, and cognitive skills.

The Black-White gap increases after children enter school. Phillips (1998) reports that when Black and White students attend the same schools and have the same prior scores and SES, Black children do not gain as much as White children on achievement tests from year to year. The differences can be attributed to continuing differences in background such as SES and to differences in the experiences Black and White students have in schools. Ferguson (1998) reports school characteristics that have been found to increase student achievement are smaller class size and effective teachers who themselves have high test scores. Both of these interventions have more of an effect on Black students' achievement than on White students'. Jencks and Phillips acknowledge there yet remains a portion of the test score gap that has not been explained. Their hypotheses include cultural characteristics that might affect Black students' test scores, such as fear of being seen as acting White or "stereotype threat" (i.e., fear of confirming negative stereotypes, Steele & Aronson, 1998, p. 401).

John Ogbu attributes the underachievement of American Blacks to their being an involuntary minority, a group that was brought to this country contrary to their own wishes. That condition results in the historical discrimination against the minorities within the educational system and society, which leads to distrust of White-controlled institutions and to adaptive responses. The beliefs and responses of opposition, ambivalence, and mistrust affect minority student attitudes toward schooling and their

achievement (Ogbu & Simons, 1998).

In a review of NAEP reading and mathematics trends over the last three decades (1970-2000), Lee (2002) observed that the Black-White achievement gap decreased substantially until the late 1980s, when it began to increase somewhat. The Hispanic-White gap follows the same pattern but is not as extreme. Lee proposes that correlations in the 70s and 80s that explained the achievement gaps based on economic and social conditions no longer hold up and that new multidisciplinary frameworks for identifying influences on achievement gaps are needed. He recommends attention to policy issues including the effects of standards-based reform and high-stakes testing.

In an analysis of the effects of race and class on NAEP mathematics test results, Lubienski (2002) reported that the achievement gaps based on race were larger than those based on SES, and although achievement for all races has improved from 1990 to 2000, the gap has slightly widened. The NAEP data do not include school composition variables, however, so Lubienski's observations that eighth and twelfth grade White students who qualified for free and reduced lunches scored higher than Black students who did not qualify does not take into account that middle class Black students are more likely to attend high poverty schools. In these NAEP data, school effects cannot be distinguished from individual background effects.

Recent studies of demographic factors and student achievement within the context of standards-based reform have produced conflicting results. In Louisiana, Caldas and Bankston (1997) addressed the extent to which the SES of students' peers exerts an effect on individual student achievement on the tenth grade graduation exam, independent of the student's own background. SES was defined by poverty based on

students receiving free or reduced lunch and by family social status based on students' reports of their parents' education and occupations. They also included race, percentage of African American students, and the students' self-report on time use variables: television, reading, homework, working, and school activities. Results showed the time use variables had small to non-existent correlations with measures of poverty, SES, or academic achievement. Students' race had the greatest effect on achievement. The effect of individual family poverty on achievement was small; educational and occupational status had a more important influence on achievement. When the poverty, social status, and race of peers were included in the regression, the effect of classmates' race and family social status was significant and substantial. Family poverty status of peers had a small effect in addition to the effect of race. However, the  $R^2$  for the combined factors was only .210.

In contrast, a study by Abbot and Joireman (2001) in Washington State, based on the Washington Assessment of Student Learning (WASL) and the Iowa Test of Basic Skills scores, found that ethnicity (percent White) was not a major influence, contributing only 6% of the variance in school achievement, but that the proportion of low income explained 12-29% of the variance indicating that the influence of ethnicity on achievement is indirect and a result of the association of poverty with ethnicity.

The Caldas and Bankston (1997) and Abbot and Joireman (2001) studies may not be comparable because the ethnicity of students in Washington and Louisiana are quite different. In the Washington study all non-Whites were grouped together. Hispanics were the largest minority at about 9%; Asians, who typically score higher than Whites, were 7%; and African Americans represented only about 5%. In

Louisiana, African Americans are a much larger minority, and other ethnic groups are so small they were not included in the study. This is an obvious example of the different contexts and uses of the terms *race* and *ethnicity*.

Sutton and Soderstrom (1999) conducted an extensive study of all the schools in Illinois, which has a standards-based system. The Illinois Goal Assessment Program (IGAP) measured school achievement. A regression analysis indicated strong relationships between achievement and school-level characteristics of low income, percent White, high school graduation rate, and dropout rate. There were moderate relationships for achievement with attendance, mobility, and high school pupil-teacher ratio. There were low correlations for achievement with average class size, elementary pupil teacher ratio, teacher salary, teacher experience, and expenditure per pupil.

Sutton and Soderstrom (1999) grouped the variables into *can control* (average class size, teacher experience, pupil-teacher ratio, teacher salary, expenditure per pupil) and *cannot control* (percentage White, percentage low income, percentage attendance, percentage mobility, high school dropout rate) regression models. (This research is consistent with Bloom's, 1980, alterable and non-alterable variables.) Then Sutton and Soderstrom tested the models using step-wise multiple-regression analysis for each grade level and tested subject. For Grade 3, cannot-control variables contributed 70% of the variance on the IGAP for reading and 56% for mathematics. In contrast, the can-control variables contributed 26% for reading and 18% for mathematics. For tenth grade, cannot-control variables contributed 74% of variance in reading and 62% in mathematics. The can-control variables contributed 23% of the variance in reading and in mathematics. Cannot-control variables consistently were better predictors of

achievement scores than can-control variables.

Sutton and Soderstrom (1999) point out the importance of accounting for the effects of factors outside the control of the schools because comparing schools and school districts without adjusting the outcome measures favors the advantaged schools with advantaged students and adversely affects schools with high proportions of low income students. They suggest statistical techniques to adjust outcomes based on input variables that the schools cannot control and report that Indiana school assessments are being statistically weighted as they have proposed. It should be noted that the can-control variables in their study mirror the school input variables in the Brookover et al. (1979) study. Alterable school climate variables, which accounted for considerable effect in the Brookover et al. study, were not included in the cannot-control measures by Sutton and Soderstrom.

Most of the work on the influence of context variables on student outcomes is based on static, absolute measures of achievement rather than on the extent to which achievement changes as a result of schooling. Tennessee, however, uses an assessment system based on longitudinal data on individual students (Sanders, 1998a). In Sanders' studies, race and socioeconomic factors were found to be unrelated to the cumulative gains for schools.

#### *Recent Findings in Kentucky*

In KERA studies, researchers have examined both cross-sectional absolute scores (e.g., fourth grade scores in 1990, 2000, 2001, and 2002) and change scores (e.g., change from 1999 to 2000) in analyzing achievement results. Findings have been inconsistent regarding the effects of socioeconomic factors on student achievement and



school performance both on absolute scores and on change scores. For example, a 1995 study by the KDE (cited in Petrosko, 2000) determined that the KIRIS results were fair and that school SES level did not affect the distribution of rewards and assistance. The KIRIS technical manual (cited in Petrosko) reported that the percentage of minority students, percentage of students receiving free and reduced lunches, the size of the school, and the baseline score accounted for only 17% of the variance in the percent of its improvement goal that a school accomplished, which they did not consider large enough to be a concern. Based on simple correlations, the KDE authors also concluded there was not a strong relationship between percent of students on free and reduced lunches and percent of improvement with  $r$  values of -0.06 at fourth grade, -0.26 at eighth grade and -0.11 at twelfth grade. The KDE authors also considered the correlation of percent minority students with percent of improvement negligible (grade four  $r = -0.22$ , grade eight  $r = -0.08$ , and grade 12  $r = 0.03$ ).

Although, the 1995 KIRIS technical manual (cited in Petrosko, 2000) indicates race is not an influential factor in school accountability change score measurements, another section of the manual compares the performance of Black and White students on absolute scores. "White students did better than Black students on both open-response and multiple-choice items. However, the racial gap was less on open-response items than on multiple-choice items" (p. 58).

According to Petrosko (2000), other researchers have different findings. He points out that later studies indicate background factors are influential in school success. Catterall et al. (cited in Petrosko) found that between 1995 and 1997, the poorest school districts made the least improvement on their accountability index scores.

In a three year study (1993-1995) in one district in Kentucky, Guskey (1997) found that several socioeconomic and contextual variables showed a high correlation with school accountability scores. Using stepwise multiple regression, he discovered a single socioeconomic variable, the percent of students qualifying for free and reduced lunches, accounted for most of the variance with an  $R^2$  range from 51%-66% for elementary schools, 60%-81% for middle schools, 62%-78% for high schools. For elementary and middle schools, the  $R^2$  was highest in 1993 and lowest in 1995, indicating a decreasing influence of poverty. The opposite occurred at high school, where the  $R^2$  was highest in 1995.

In contrast, Guskey (1997) found socioeconomic factors were not correlated with accountability score gains from year to year. Similarly, in a study of the implementation of the Primary Program in Kentucky, Luvisi (2000) found that demographic variables (enrollment, free and reduced lunch, and geographic region) accounted for 34% of the variance on KIRIS scores and 34% of the variance on CTBS scores but only 3.5% of the variance on Accountability Index change scores.

In a recent study of the effects of instructional strategies on seventh grade science achievement, Ennis (2002) examined individual student achievement and school change scores. At the individual level, he found that student poverty, gender, and race explained a greater portion of the variance than the instructional strategies. Students on free and reduced lunches scored lower. Female and White students scored higher. Demographic factors did not contribute significantly to the change scores, however. For change scores, instructional strategies explained 4.0% of the variance.

Roeder (1999) studied the effects of school characteristic variables and district

context variables on school Accountability Index scores for the years 1993-1997 and on Accountability Index change scores for those years. The school variables were (a) a variable indicating whether the school was an early adopter of SBDM and Family Resource/Youth Service Centers, (b) student-teacher ratio, (c) school size (total enrollment), (d) non-White enrollment, (e) students per school administrator (school bureaucracy), (f) proportion of students eligible for free and reduced lunch (poverty), and (g) the school level (i.e., elementary, middle, or high). District contextual variables were (a) the state/local revenues per pupil prior to reform (1988) and change in state and local revenue after reform (1988-1991), (b) a variable representing whether there were private schools in the district that would provide competition for the public schools, (c) a district index for academic achievement based on high school completion and college attendance, and (d) district bureaucracy. Multiple regression analysis revealed the most consistent, significant effects were due to the proportion of poor and minority children in the school. Schools with higher proportions of poor and minority children consistently had lower performance on KIRIS tests in all five years. The effects of the percent of poor children were significant and substantial for elementary, middle, and high schools. School size and percent minority predicted significantly and negatively for elementary and high schools. (Smaller schools with lower percents of minorities scored higher.) At the middle school level, however, only the percent students eligible for free and reduced lunches was significant.

The Roeder (1999) study also looked at the changes in the Accountability Indices for various intervals in the years between 1993 and 1997. For all the intervals tested, the previous year's score was the strongest predictor of change scores and was

negative, i.e., the higher the previous score the smaller the change score. In contrast to the work of Guskey (1997), Luvisi (2000), and Ennis (2002), Roeder included previous scores as a variable and found school poverty had a significant negative effect. Further, the change in accountability scores got smaller over time in high poverty schools. Percent minority was also a significant and strong negative predictor in seven of the ten intervals tested.

Roeder (1999) created an index of school disadvantage for each school using factor analysis. The index included three variables: the poverty rate (proportion subsidized meals), state/local revenue, and academic success before KERA. He compared the most advantaged 10% of schools to the least advantaged 10% and found that the gap between them on the KIRIS Accountability Index consistently increased from 1993 to 1997.

In a later study, Roeder (2000) used the 1998 KIRIS and the 1999 CATS Accountability Index scores in Kentucky's two largest school districts (Jefferson and Fayette Counties) to examine the effects of SES and race. He reported a pattern of achievement in those two counties similar to the rest of the state, except the two urban counties had more schools with high poverty. Jefferson County schools had higher proportions of poor and minority students and performed at lower levels than Fayette County schools. The performance gap between advantaged and disadvantaged schools was larger in Jefferson County than in Fayette County. demonstrated that poverty was a strong negative predictor of achievement in these two counties whereas race was not. Roeder pointed out the achievement gaps based on poverty and race between the most advantaged and disadvantaged schools in these two districts increased substantially

between 1995 and 1999.

In response to research that suggests small schools may counteract the effects of poverty on student achievement, Roeder (2002) used the CATS Accountability Index and the CTBS/5 Index from the 2000 assessments to examine school size and achievement. Multivariate models for elementary schools in the two urban districts, Fayette and Jefferson, found that (a) size and (b) size interacting with poverty had no effect on school performance; instead, poverty was the major determinant of performance. Roeder suggests policymakers in districts with many schools and diverse neighborhoods should consider drawing attendance boundaries to redistribute poor children more equitably across schools regardless of size.

Using independent variables from the Kentucky School Report Cards, Roeder (2001b) analyzed their effects on the schools' CATS accountability index and CTBS/5 percentile score. He looked at cross sectional results for 2000 and at change scores from 1999-2000. For the 2000 cross sectional results on both tests, poverty continued to be the predominant predictor with a negative coefficient, but it did not have as large an effect as the earlier results. In contrast Roeder's earlier studies, race had a significant positive effect on the 2000 school Accountability Index scores when controlling for poverty. (Note: in the earlier study, race had a positive, but not significant, coefficient.) Other significant predictors were percentage of teachers with master's degrees, student/teacher ratio, and the numbers of parent volunteer hours.

As with most other studies, Roeder (2001b) found the greatest predictor for school change scores was previous scores, with a negative correlation. Controlling for the previous score, the only other significant results with the CATS change scores were

a positive correlation of proportion of teachers with master's degrees and a high negative correlation of poverty with the CTBS change score. This distinction suggests that instruction influences CATS scores whereas background is more influential for CTBS

Pitts and Reeves (1999) studied the interaction of the geographic location in Kentucky with school poverty. As with other studies they found school accountability scores were negatively influenced by poverty. They found that metro-adjacent and small town districts had more positive effects on achievement than rural or metropolitan locations. But they also found that the negative effects of poverty were moderated by a rural location--the more rural a school district (the less dense the population), the greater the effect of moderating the effects of poverty. They also found that there were clusters of adjacent districts in various parts of the state that had similar scores. They recommend that socioeconomic and geographic factors be included in analyses of assessments in order to give a truer picture of school performance. Borland and Howsen's (1999) results support similar conclusions on rural and urban effects; they found that highly rural and highly urban districts demonstrate lower achievement than areas of more moderate density (metro-adjacent or suburban).

A comprehensive study of gender and racial differences at the individual student level on KIRIS assessment was prepared for the Kentucky Department of Education (D. C. Smith et al., 1999). This study compared the effects of gender and race on the KIRIS reading, mathematics, science and social studies theta scores and standard constructed response index (CRI) units. Whites scored higher than African-Americans, and the difference between African-Americans and Whites increased slightly over the six years.

Whites in districts and schools that had above the 3% median percent of minority enrollment were more likely to score above African Americans than in those districts and schools with less than the median 3% minority enrollment. Differences were greater at the seventh and eighth grades and less at the eleventh.

Unlike the other studies above, which used the school as the unit of study, D.C. Smith et al. (1999) examined the effects of race, class (determined by participation in Title I; the rationale for using Title I participation was not explained), and gender on individual student scores. Regression analysis revealed significant, small, independent effects of class, race, and gender in all subject areas at all grades except eleventh grade (when few students participate in Title I) and between genders in fourth grade science. Only 7-15% of the variance was explained by the three variables, but there were always effects of race and gender.

In another regression model in the same study, D. C. Smith et al. (1999) included an additional seven variables from the student questionnaire given with KIRIS tests. The seven variables included items representative of socioeconomic status, geographic mobility, and home environment. Together the ten variables represented 20-25% of variance in theta scores, and 10-20% of variance in CRI scores. The student's perception of grades was the strongest predictor at all grades. (Students' perceptions of grades could be construed as a mediating factor, a type of independent variable that taps the internal state of the students; see Ennis, 2002, for an extended discussion and analysis.) Title I participation and race had significant effects at all grades. Interestingly, the number of books in the home had a negative effect on achievement except in fourth grade--a finding that demands some explanation. School

absences and having a foreign language spoken in the home had small negative effects. Length of time in Kentucky and in the school generally had insignificant effects. The other variables included in the D.C. Smith study are included in the discussions of specific variables below.

An examination of the 2002 Kentucky Performance Report (KDE, 2002b) data disaggregation for 1999-2002 reveals that the achievement gaps persist under the CATS system. On the reading and mathematics CATS test, Asians consistently score highest, then Whites, then Hispanics, while African Americans have the lowest scores. Rather than years of schooling alleviating the gaps, they are largest at the high school level. The achievement gaps on Kentucky tests are also evident in the national NAEP test results. *Ed Watch Online* (Education Trust, 2001) reported that on the 1992-1998 NAEP, Kentucky Black students were about two years behind White students at fourth and eighth grades in reading and mathematics although the gap was smaller than the national average.

#### *Recent Jefferson County Studies*

As described in Chapter I, the performance on KERA assessments in Jefferson County Public Schools (JCPS) have generally been below the state averages. In addition, JCPS has had greater achievement gaps between groups of students based on race and SES. The disproportionate nature of the composition of the schools combined with the possible effects of poverty and ethnicity on achievement may contribute to the disproportionate assessment results in Jefferson County.

Support for this position is found in a study by Muñoz and Dossett (2001). They explored the correlation between school and social variables on the school-level CTBS



NCE (normal curve equivalent) scores in 133 JCPS schools for four years. Hierarchical multiple regression analyses were used on data from school profiles for the 1995-1996 to 1998-1999 school years. Independent variables were percent students qualified for free and reduced lunch prices, percent of Black students, number of teachers with masters degrees, number of teachers with emergency certifications, school level (elementary, middle, or high school), per pupil expenditures, and student-teacher ratio [teacher-student]. Overall the regression models explained 68% of the variance. The strongest predictor was poverty (free or reduced lunch), which explained an average of 58% of variance across all four years. The school-related variables explained an average of 24%.

Muñoz and Dossett (2000) recognized that the Kentucky Accountability model does not take into account the many differences in conditions that affect student learning, inside and outside of school. They posited that both the input variables and the output variables are complex and many-faceted. Muñoz and Dossett created a *need* factor to include in the accountability formula that would compensate for the effects of those school composition variables. The formulation of the need factor was based on an informed estimate by top-level JCPS administrators of the contribution of each of four variables. They were weighted as follows: students receiving free and reduced price lunch (.50), mobility rate (.15), percent of Exceptional Child Education [students with disabilities] (.05), and percent of single parent households (.30). This formula created the *weighted input index* for the independent variable. The dependent variable, the *weighted outcome index*, included: results of the KIRIS or CATS tests (Academic Index, .50), percent of students not academically at risk (above stanine 3 on another

unidentified test, .20), writing portfolio scores (.15), attendance percentage (.10), and the CTBS scores (.05). A simple regression analysis indicated that the weighted input index contributed significantly to the variance in the weighted outcome index: 66-75% for elementary schools, 78-86% for middle schools, and 75-79% for high schools. Comparing the standardized residuals identified schools that were performing above and below expectations. The authors point out that this more equitable comparison of schools, which compensates for student background variables, encourages attention to those instructional attributes of high-performing schools that do contribute to student learning such as variation in instructional time.

The weights of the variables in the above indices were chosen arbitrarily. Therefore, the contribution of each of the separate independent variables to the variability of the components of the weighted output index was not addressed nor was the interaction of such variables. Also, the different input variables might have differing effects on the different components of the output index. Further it should be noted that the weighted dependent variable contains the writing portfolio scores twice, as part of the Academic Index and as a separate entry, which could inflate the results.

### *Gender*

The research findings on gender and student achievement are contradictory. Mayeske and Beaton (1975) reported gender did not significantly affect student achievement or motivations; however, more recent studies have found differently. On the NAEP Long-Term trend reading assessment and on the main NAEP reading assessment, females have consistently outscored males from 1971 to 1996 (Coley, 2001; Vanneman & White, 1998).

In mathematics achievement, a review by Leder (1990) indicated few consistent differences between males and females at the primary grades, but males outperformed females in secondary school, especially at the higher achievement levels. On the NAEP Long-Term trend assessment, gender differences have not been significant, but males have slightly outscored females in mathematics (Vanneman, 1998a). On the main NAEP test (different from the Long-Term trend assessment), Coley (2001) reported that White, fourth-grade males scored higher on NAEP mathematics than females in 1996, although there was no significant difference for eighth and twelfth grade students. Lubienski (2001) pointed out the fourth grade gap was concentrated in the highest SES group where White males scored a significant 7 points higher than White females. Other mathematics score differences among SES and gender groups were not significant, and there was no gender gap for African Americans.

Zhang and Manon (2000) found a similar lack of gender differences in mathematics achievement when comparing scores on the Delaware assessments (grades 3, 5, 8 and 10). When they looked at the students at the extremes, however, they found a different pattern. In grades three and five, females outscored males in the top 10% and bottom 10% of students. At grades eight and ten, females continued to score higher than males in the bottom 10%; however, males outscored females in the top 10%. Males demonstrated an increasing variability among their scores at the higher grades. Zhang and Manon also found that the test item format had a slight effect. Males performed better than females on multiple choice questions, but females performed better on those requiring an extended written response.

This tendency for males to be disproportionately represented among the highest

achieving students in mathematics has been evident with more demanding tests. Coley (2001) reported that on the SAT I, a test for college-bound students (especially for more selective colleges), males scored slightly higher than females on the verbal test, except among Black students. Black males had a slight advantage over Black females in the early 1990's. But since 1994, Black females have scored slightly higher than males. On the mathematics portion of the SAT I, males of all ethnicities scored substantially higher than females. The trends were similar for high school students taking the Advance Placement examinations. Females scored slightly higher on the English literature and composition exam; but on the mathematics exam (Calculus AB), males scored substantially higher in all ethnic groups.

In the KIRIS technical manual (cited in Petrosko, 2000), comparison of students' KIRIS achievement based on gender and ethnicity "revealed that on average, girls did better than boys on both open-response items and multiple choice items" (pp. 57-58), and the "gender difference was greater for open-response items than multiple choice" (p. 58). Likewise, as mentioned above, D. C. Smith et al. (1999) found that gender was one of the largest predictors of variance in KIRIS reading, although not as influential in other subjects. Similarly, Ennis (2002) found that females scored higher than males at the individual level on seventh grade science tests.

The 2002 Kentucky Performance Report (KDE, 2002b) data disaggregation for 1999-2002 reveals females score substantially higher on the reading CATS tests but only slightly higher than males on the mathematics tests. Therefore, although in the national literature the findings on gender are inconsistent, on the Kentucky CATS tests, females consistently out-perform males. This study will indicate whether those trends

extend to urban Jefferson County.

### *Mobility*

Mobility is a measure of how frequently students move from one school to another during the school year. It is often associated with poverty or with disarray in family relationships, and hence may be a function of socioeconomic status (cf. *A Report from the Kids*, 1998). A study by Medsker (cited in Petrosko, 2000), which looked at all the schools in Kentucky, found that at the individual level, the more often students changed schools, the lower were their KIRIS test scores. In contrast, the D. C. Smith et al. (1999) study found that when mobility was regressed with the nine other variables against individual student achievement, it was negatively significant only for the fourth grade reading and science test, but positive for all other grades and subjects.

At the school level, Medsker (cited in Petrosko, 2000) found the higher the school mobility, the lower the accountability score; and high transience negatively affected the accountability score improvement (change score). Petrosko commented that it seems unfair to penalize schools or districts for the effects of mobility, which is not under the control of the district or school. Mao (1997) found similar results for mobility in Texas, another standards-assessment driven state.

Guskey (1997) found mobility to be moderately correlated with accountability scores at elementary schools in Kentucky, very highly correlated at middle schools, and highly correlated at high schools. However, in the stepwise multiple regression, mobility did not add significantly to the equation after the entry of percent of students on free and reduced lunch. This finding would seem to indicate that the correlation of mobility with achievement is only a consequence of its correlation with poverty.

Jennings, Kovalski, and Behrens (2000) at the University of Arizona investigated the association of mobility with student achievement. Their review pointed out there have been contradictory findings regarding mobility. Most studies have found mobility has a negative affect on student achievement. However, some have found no effect of mobility if previous achievement and other variables are controlled. Other studies have found that other variables--intelligence, maltreatment, and SES--interact with mobility. Jennings et al. deduced that differing definitions of mobility contributed to the contradictions. They tested 11 different formulas for defining mobility, using the Arizona Department of Education categories for entry and withdrawal, to find the one that contributed the most to school-level student achievement on reading and mathematics as measured by the SAT-9 for the years 1997 and 1998. In results similar to Guskey's (1997), they found that mobility added nothing to the regression model when other demographic variables (percent of students on free and reduced lunch, percent of students with limited English proficiency, and the absence rate) were included.

The Kentucky studies by Medsker (cited in Petrosko, 2000), Guskey (1997), and D. C. Smith et al. (1999) used the same data sources, so differences in the measurement of mobility do not seem to be responsible for the disparity. However, the Kentucky studies were using different dependent variables, different units of study, and different methods of statistical analysis. Smith's study used only theta test scores for the individual students, whereas Guskey used the school Accountability Index for schools as the unit of study.

### *Family Structure*

Single-parent and other non-traditional family structures are derogated in the conventional wisdom as not providing an optimum home environment for children. Some support for that position is found in the educational literature and research. In an article focused on ways changes in family trends affect students, Hofferth (1987) reviewed six studies that implicate parental divorce and separation as negatively affecting student achievement. Also, in a review of the effects of divorce on children, Hopper (1997) cites seven studies that found divorce to have a negative effect on student achievement or school success. However, the extensive study of the effects of 28 home background and behavior variables on student achievement by Phillips et al. (1998) found that when other factors such as income, education, race, and mother's background and test scores were held constant, the number of parents in the family had little or no effect.

Other studies concluded that the differences attributed to family structure are a function of the higher levels of poverty and lower resources available to many single-parent families. The association of poverty with single, divorced, or never-married mothers is well established (Hofferth, 1987). Single parent families, often headed by single mothers, are frequently identified as having less capacity to provide the resources to enable children to be successful in school. Wilson (1987) associates the increase in single female-headed households to economic conditions, which result in increasingly fewer low skill jobs and, therefore, fewer desirable fathers available for the marriage pool who can support a family.

Coleman et al. (1966) considered the absence of a father in the home as a negative influence, but Mayeske and Beaton's (1975) reanalysis of the Coleman data

found that most of the effects of family structure could be attributed to socioeconomic status. Likewise, a recent longitudinal study of early primary students in Baltimore (Entwisle & Alexander, 1995) confirmed that socio-economic factors (eligibility for subsidized meals and parents' education), rather than two parent or single parent family status, accounted for differences in achievement and that those differences occurred over the summer months when school was not in session.

Whether family structure affects student academic achievement independent of poverty remains debatable. For example, Jeynes (1999) analyzed the National Educational Longitudinal Survey data and found that both low and high SES children from divorced families in which the custodial parent has remarried have lower achievement than children from single parent families or intact families. He found the differences between intact and single parent families were inconsistent and insignificant.

The complexity of the influences affecting children who do not have the traditional intact two parent family is well demonstrated by a 12-year study by Weisner and Garnier (1992) on a wide array of non-conventional family life-styles and school achievement. They found that overall children from non-conventional families did as well or better on school grades and on intelligence tests (Wechsler Intelligence Scale for Children-Revised and Stanford Binet). Students from families that had a high commitment to their unconventional life-style did better than students whose families had a low commitment to an unconventional life-style. They point out "The categories usually used to group families, such as single parent, unwed mother, divorced, unmarried couple, or married couple, are not capturing important differences in values,



commitment, and stability, which influence children ..." (p. 628)--and, one might add, student achievement.

In spite of inconsistent evidence, educators continue to expect single-parenthood to be a detriment to student achievement. In JCPS research, Muñoz and Dossett (2000) included the percent of students without an intact two parent family as one of the components of the weighted input variable in their study of the effects of need on student achievement, but they did not analyze the independent effect of family structure nor its contribution to the variance. (Note: The JCPS family structure data classifies intact families with a mother and father as *dual* families and all other family arrangements, which include single parents, foster parents, step parents, and other relatives, as *single parent*.)

Another study conducted on JCPS family structure data used multiple regression to identify the effect of single parent family structure on student longitudinal achievement from third grade on the California Achievement Test to sixth grade on the Comprehensive Test of Basic Skills, Version 5 (Spievak, Snyder, Miles, & Burns, 2001). That study found only one significant result out of six regressions (for low and middle SES students on reading, mathematics, and language tests). Having a single parent had a small *positive* effect on middle class students' achieving above expectations in mathematics. The inconclusive results in this second study, the inaccuracy of the data, and the criticisms by Weisner and Garnier (1992) bring the use of this variable in evaluating student achievement into question.

### *Disabilities*

Disabilities are not confined to people of lower societal status, of course, but

persons with disabilities are frequently subject to discrimination and limited to lower status occupations. Also, some disabilities are more prevalent among families in poverty due to prenatal or childhood experiences. The demands of caring for children with disabilities can lower a family's income as well as confining some persons with disabilities to low income employment in adulthood. Therefore, disability is another component that contributes to societal stratification.

Students with disabilities are included in the Kentucky testing for accountability. Being identified with a learning disability would predispose one to expect that such students would not do as well as regular students. Students with disabilities are allowed to take the tests using accommodations that have been identified as needed for their academic success, however, which could compensate for their disability. Also, theoretically, including those children in the accountability system will keep them from being left out of the advantages of a standards-based system and their achievement will rise along with other students. Of course, the fact that ECE students are tested along with other students eliminates the incentive to classify low-achieving students as ECE in order to remove them from the testing pool to raise school scores.

Results from 1993-1996 KIRIS testing reported by Trimble (1999, cited in Petrosko, 2000) indicate that students with disabilities do have lower scores than other students in reading and mathematics. As hoped, the gap narrowed substantially from 1993 to 1996 for grade four students, but not as much for grades eight and eleven. This trend supports the claim that raising expectations for students with disabilities will raise their achievement.

The study by Guskey (1997) of school-level effects demonstrated the

multicollinearity of the percent of children with disabilities in a school with other disadvantaged background factors. In the Kentucky district he studied, schools with higher percentages of students with disabilities also had higher percentages of children on free and reduced lunches, higher percentages of minority children (Black), higher mobility rates, and more retentions. The regression analysis revealed that neither percent of students with disabilities nor the other factors added to the regression after the percentage of students on free and reduced lunches was entered, indicating that poverty level is the characteristic that is responsible for the concentration of the other manifestations of disadvantage in a school.

The Spring 2002 Kentucky Performance Report (2002b) reveals that children with disabilities score substantially lower on the CATS tests and on the CTBS/5 than other students. Those that were tested with accommodations scored lower than those tested without accommodations. This does not necessarily indicate that the accommodations are not helpful to students in compensating for their disabilities. It may be that students who qualify for accommodations are generally lower achieving than students who do not, and therefore, it may be a selective factor for lower-achieving students.

### *Giftedness*

Students are identified as gifted and talented in Kentucky based on various procedures including testing. Although the identification may be based on giftedness in art, music, or interpersonal abilities as well as academics, many of the gifted students have demonstrated that they are advanced beyond their grade level in academic ability. These students compose the highest track of school stratification, and the association of

economic class and high test scores or academic achievement is well documented (Bowles & Gintis, 1976; Guskey, 1997; Jencks et al., 1972; Oakes, 1985; Persell, 1977; Rothstein, 1998). Identification as a gifted student, therefore, can be considered a proxy for socioeconomic class as described by Dika and Singh (2002).

Logically, gifted and talented students in a school will by definition have higher scores than their classmates. At question is what effect the presence of students identified as gifted and talented might have on the other students and on the scores of the school as a whole.

The percentage of gifted and talented students in a school is a composition variable that introduces the issue of access to knowledge. The operating assumption of gifted education is that gifted students need a more challenging or advanced curriculum than provided by the regular curriculum, and, hence, their curriculum must be differentiated (Maker, 1982). If a school has a program for gifted and talented students that is differentiated from the other students' curriculum, conceivably the regular students access to knowledge will be affected. The results would also be affected by the type of program provided, for example, whether the gifted and talented students and their curriculum were segregated from the other students via tracking or ability grouping or if all students have access to the highest level curriculum. The identification of gifted and talented students also raises the issue of the effects of their presence on the attitudes of teachers and other students.

Research regarding the effects of ability grouping and tracking has been contradictory. Some researchers maintain that ability grouping is detrimental to the achievement of lower level classes and high level students are not affected (Braddock &

Slavin, 1993; Oakes, 1985; Persell, 1977). Others argue all students profit from ability grouping. Kulik (1993) reviewed the research on grouping using meta-analysis. He reports that grouping of children has little effect on achievement when all levels have the same curriculum and instruction. Gifted students showed the greatest gains in achievement when they were in classes that provided the greatest amount of differentiation with acceleration and enrichment for the high level classes. In such a program, the access to knowledge would not be equal, a situation that would seem to contradict the standards-based system that requires high standards for all students.

Gifted and Talented students scored substantially higher than other students on the 2000 CATS and CTBS tests (KDE, 2002b). None of the studies of Kentucky accountability, however, have included that variable in their analysis of the variables affecting school scores.

### Demographic Factors and Accountability

#### *Individual Student Achievement, School Performance, and Change*

There are various dimensions to the effects of demographic factors on student achievement within an educational system. One is the effect of demographic or background factors on individual students versus the effect on school-level accountability measures. Another dimension is the effect of demographic factors on absolute scores, the yearly scores, versus the effect on change scores between years (both individual change scores and school-level change scores). Most of the research in Kentucky cited above has looked at only one or two of these dimensions. Most of the studies have examined school-level accountability measures at the absolute level. A few have looked at change scores.

That demographic factors may have different effects at the individual and aggregate levels is generally accepted (cf. White, 1982). Thus, presumptions that the effects of a demographic factor on individual scores and on school performance measures might be the same or similar could be mistaken. In Kentucky, for example, D. C. Smith et al. (1999) found that race had a small, negative effect on student achievement at the individual level independent of poverty (Title I), but Roeder (2001b) found that the percentage of [Black] minorities in a school had a small, positive effect on school achievement. Also, there are studies that indicate the influence of certain factors may be waxing or waning depending on the circumstances, as in the case of the decreasing reading achievement gap between students with disabilities and regular fourth graders (Trimble, cited in Petrosko, 2000) and the increasing gap between low poverty and high poverty schools (Roeder, 1999). Whether these findings are related to increased individual poverty or other factors awaits more complete information.

Several studies have indicated that the effects of demographic factors are more intense in the aggregate. Because the effects of high poverty on school scores is well established, several researchers (cf. Guskey, 1997; Lee & Coladarci, 2001; Linn, 2000; Roeder, 2000) suggest the need to take background factors into consideration when assessing for purposes of accountability.

Change scores have only been studied under KERA at the school level and do not reflect the same trends as the absolute scores. Whereas aggregate school-level scores are universally negatively correlated with poverty, most Kentucky studies (Ennis, 2002; Guskey, 1997; Luvisi, 2000) have found demographic factors to be insignificantly correlated with change scores. On the other hand, Roeder (1999, 2001b)

found that poverty significantly affected change scores when previous scores were included in the analysis. He consistently found previous scores to be the greatest predictor of change scores and negatively so.

In his 1999 study, Roeder found poverty to have a significant and negative effect on school accountability change scores. He found that the score changes got smaller over time in schools with high poverty. He also found percent minorities to be a significant and strong negative predictor in seven of the ten testing intervals. In the 2001b study, Roeder found poverty to be a negative predictor on the CTBS but not on the CATS, and he found school percent minorities to be a positive predictor (although not significant).

Since previous scores are negatively correlated with change scores, and since high poverty schools generally have initially low scores, a low previous score would predict a larger change score, while poverty would predict a lower score. These are offsetting trends; thus omitting previous scores from a regression could disguise any negative effects of poverty on change scores. Therefore, the inclusion of previous scores in the regression analysis is essential to measuring the effects of poverty. Also, since change scores are negatively correlated with previous scores, as schools improve their scores, the change scores can be expected to decrease. It should be noted that these conclusions must be offered with great caution because Roeder's (1999, 2001b) empirical findings have not been widely confirmed.

If, as Ennis, Guskey, and Luvisi have found, change scores are not affected by demographics and change due to instructional strategies is no greater for schools with high proportions of poor, Black, and male students, then those schools that are behind

will remain behind. If Roeder's results prove more representative, one can expect high-poverty schools to fall increasingly farther behind, but perhaps high-minority schools will improve relative to other schools. Overall, however, these findings demonstrate that schools have not been able to compensate for the effects of demographic factors on aggregate school scores and accountability measures.

### *High-stakes Testing Affects Poor and Minorities*

Orfield and Wald (2000) assert that high-stakes testing hurts poor and minority students the most. Their studies show that dropouts and retentions for minorities and poor students are increased, primarily because of promotion and graduation requirements but also because of the erosion of the quality of instruction. It follows that since at risk students need the most expert instruction and up-to-date resources to meet high standards, any diminution in the quality of instruction will affect them most. Teachers from high minority classrooms report more pressure to teach to the test, more use of standardized tests, and more time spent on test preparation than teachers in low minority classrooms and, therefore, a narrowing of the curriculum to test preparation (Madaus & Clarke, 2001).

Rotberg (2001) posits that high-stakes testing as part of accountability measures has a detrimental effect on the high standards it was meant to support. High-stakes testing can hurt disadvantaged children by "...encouraging, and even requiring, policies that may not be in the best interest of the children" (p. 170). She points out that policies that assign low scoring students to special education classes or encourage retention of low scorers may result in higher school scores but may not provide the best placement for the students.



Other concerns Rotberg (2001) describes may not apply only to low-achieving or disadvantaged students, but since those students are the most vulnerable with the fewest outside resources to compensate for deficiencies in their schooling, any diminution in the quality of education can be expected to affect them the most. She finds teachers must teach to the test in high-stakes systems, and the test then determines the curriculum "...when the test becomes the education program" (Rotberg, p. 170). That assertion is debatable as it applies in Kentucky where the test is designed to ensure teaching of the core curriculum. Whether it does have the effect of focusing instruction on effective in-depth teaching of the curriculum or on increased time spent teaching test-taking skills (Kannapel, Coe, Aargaard, & Moore, 1996) is still under investigation. Rotberg also asserts high-stakes testing "... discourages the most qualified teachers and principals from remaining in the profession" (p. 170). Most damning in Rotberg's view is that attention to high-stakes test preparation diverts attention and energy from identifying and addressing the root causes of low achievement.

The high stakes attached to accountability scores in Kentucky raise the question of whether the high stakes contribute to the failure of high poverty schools to reach their growth goals. To address this question, Roeder (2001a) devised several formulas for predicting future school scores based on past averages in change in scores under KIRIS from 1992 to 1998 and under CATS from 1999-2001. From the results, he projected that less than half the Kentucky schools and less than one third of the urban schools will reach the statewide goal of 100 by the year 2014 unless additional changes are made in the accountability system. Correlations between a school's score and its previous score indicate increasing stability in Kentucky accountability scores. The correlation between

the 1993 and 1994 scores was  $r = .56$ . In contrast the correlation between the 1999 and 2000 scores was  $r = .92$  and between 2000 and 2001 scores  $r = .91$ . The high correlation of scores with the previous year's scores means low performing schools do not show much likelihood of changing.

### *Alternative Accountability Systems*

Seeing the need for improvements in accountability systems because of the effects of background factors (which may or may not be controllable by schools) several researchers have suggested alternative approaches to accountability. Some researchers such as Bolon (2000) question the whole process of testing for accountability. He makes a point that with lack of validity confirmation, it is not known if we are testing anything more than social privilege. That position deserves serious contemplation and critique; but reversing the political momentum of the accountability movement seems unlikely in the near future. A more likely avenue for increasing equity and student achievement is for researchers and policy makers to address possible ways to create more accurate, equitable, and motivating accountability systems.

Lee and Coladarci (2001) explored the use of hierarchical linear modeling to take student and school background characteristics into account when assessing school effectiveness. They used the 1996 eighth grade NAEP data to compare the use of such a model in Maine and Kentucky, both of which have instituted a standards-based reform system. They tested three models to measure the extent of the effects of schooling on student achievement (school effects). The first model, which did not control for student background factors or school composition factors, simply partitioned the variance in mathematics achievement into within- and between-school effects. Model 2 controlled

for the student's race and SES, and indicated the extent to which students' race and SES contribute to the students' scores. Model 3 controlled for student background, race and SES, and also for school composition predictors, percent White and average SES, to explain between-school variation and indicate school performance.

Using the regression models for Maine and Kentucky, Lee and Coladarci (2001) found that individual race and SES were significant predictors of student achievement in both states. The school composition variables of race and SES, however, did not have a significant effect in Maine. In Kentucky only school SES composition had a significant effect on school mean achievement. They also found that the racial and social gaps varied very little between schools in the respective states, which they interpreted to mean that the gaps are not a result of school effects. (They may be underestimating how similarly schools treat students of different backgrounds or overestimating the effects of other school inputs.) The authors suggested that other background factors such as prior achievement and mobility be considered in order to create a more comprehensive model. They propose that an estimation of true school effects requires a model with school input variables, including school practice and context.

The Dallas school system uses an elaborate two-stage model to determine school and teacher (classroom) effects (Webster, Mendro, Orsak, & Weerasinghe, 1998). In the first stage, multiple regression is used to control the effects of what they call *fairness variables* (p. 82) including gender, SES, ethnicity, and limited English proficiency. In the second stage, the equalized residuals from the first stage regression (from which the effects of the fairness variables have been removed) serve as the y in a

hierarchical linear regression to distinguish the contribution of prior achievement and school context variables (school mobility; overcrowdedness; average SES; percentages on free or reduced lunch, minority, Black, Hispanic, and Limited English Proficient; and instructional days lost to unfilled vacancies) from classroom effects.

Linn (2001) reported that Tennessee, North Carolina, and South Carolina and some other school districts around the country also use hierarchical linear modeling to control for demographic variables. He cited a study by Clotfelter and Ladd (1996) in which they tested various accountability formula models. The correlations between the models were generally low, indicating the choice of model makes an important difference in results. The model that controlled for SES and prior achievement had a very low correlation with the other models.

Linn (2001) points out that a regression-based approach still requires judgments about what variables to include. He recommends putting the emphasis on improvement over time rather than on current status. A longitudinal approach tracks individual student achievement from year to year. He points out the mobility of children makes this approach difficult (but computer tracking will make that easier). A quasi-longitudinal approach would be to assess all students in one grade one year (e.g., grade 4) and the next year to assess all the students in the next grade (e.g., grade 5). In this case, the change scores should be used rather than cross sectional absolute scores. Linn explains that any longitudinal approaches require comparable tests across each grade.

The Tennessee Value-Added Assessment System (TVAAS) is a different approach developed by William Sanders based on Henderson's mixed-model equations using gain scores (change scores) as the dependent variable (Sanders, Saxton, & Horn,

1997). This is a longitudinal system wherein each individual student's achievement test scores from year to year are maintained in a database. Student progress is determined by the change in scores. Sanders explains that using gain scores controls for background and socio-demographic variables that influence the absolute scores. The student gains and the aggregated students' gains are gauged by how they diverge from the student mean gain scores. Teacher and school effectiveness are gauged by how the aggregated scores diverge from the mean. Sanders does not demonstrate the details of his statistical analysis and whether it accommodates for problems associated with the use of change scores, for example, problems of skewness because of ceiling and floor effects (Tabachnick & Fidell, 2001) and problems with unreliability due to the negative correlation of change scores with prescores (J. Cohen & Cohen, 1975). Sanders does assert that the TVAAS system is flexible and can adjust for missing data.

Western Oregon State College (Schalock, Schalock, & Girod, 1997) has developed a system that is both longitudinal, in that each student's progress is evaluated, and contextual, in that the characteristics of the child and the conditions of learning are taken into account. Teachers develop their own student assessments (pre and post instruction) and calculate the percentage gain on the outcomes the teacher has determined are desired, thereby aligning outcomes with instruction and assessment. It has been used at Western Oregon State College for teacher preparation and the authors propose it be used to measure teacher effectiveness. However, as Airasian (1997) and Stufflebeam (1997) point out, the quality of the assessments is not controlled for validity or reliability and there is much potential for teacher bias, especially if the system were used for high-stakes decisions, such as teacher evaluation or certification.

They recommend against using this method for school, teacher, or district evaluation.

Wheelock (2000) protests that the current outcomes-based approach to school reform (that relies primarily on state tests, rewards, and sanctions) is inadequate to realize long-lasting reforms in education. She acknowledges that knowledge of outcomes is important to know how students are progressing and if they are meeting generally accepted standards. In order to improve teaching and learning, it is necessary for teachers to take responsibility for making judgments about, and changes in, classroom practices with an understanding of the effects on learning. However, the focus on outcomes, rather than process, does not provide teachers with effective incentives or guidance to change their instructional practices. Although she is not specifically addressing the effects of demographic factors, the plan she proposes would shift the emphasis to improving teaching for all children including those who are disadvantaged. She proposes a hybrid system of educational accountability that combines measurement of student outcomes with evaluation of classroom practices. This approach would address the connection between the student background and low achievement.

### *Continuing the Work*

The preceding review demonstrates that students' family and cultural backgrounds are frequently, if inconsistently, correlated with how well individual students achieve in school. In addition, individual student achievement is affected by the characteristics of the schools they attend. Some of the school characteristics may be alterable (Bloom, 1980) and under the control of the school. Some may be beyond the power of the school to change. To determine the value a school adds to student

achievement--to assess how well the school is serving the students--it is necessary to sort out the unalterable characteristics from those that are alterable.

In general, demographic characteristics of the school (school composition) are considered unalterable, just as inseparable from the students as are their personal poverty or race. As Roeder (2002) has pointed out, that may not always be true because a school system, by its powers of student assignment, may change the demographic composition of a school. In addition, although the demographic conditions (such as poverty) may not be alterable within a school system, the mediating factors between the demographic factors and student achievement (such as school climate) may be alterable.

In Kentucky since KERA, the effects of demographic factors on student achievement have been studied only to a limited degree. Kentucky research indicates that poverty, gender, race, and disabilities can negatively affect student achievement, but the extent and direction of the effects are not established. Individual student background may have a different impact than the aggregate background of the student population of a school--the school composition. For example, measurements of the effects of individual poverty and percent school poverty have yielded different results. Therefore, research is needed to distinguish between the effects of individual background variables and school composition variables on student achievement and to identify more clearly the causal pathways of the impact of background factors on student achievement.

Only when those factors are included and alterable factors distinguished from unalterable may accurate evaluations be made of the value that a school adds to the knowledge and skills that students bring to school. Valid measurements of the value

added by schools are necessary to create equitable and productive accountability systems. Such research is needed to inform policy-makers' attempts to improve schools and student achievement by use of accountability and high-stakes testing.

### Summary

The Kentucky Education Reform Act was launched within a national movement toward standards-based education reform that continues under the NCLB Act of 2001. The principles of standards-based reform--assessment and local school accountability--form the core of the KERA curricular provisions.

Standards-based reform is based on the position that holding all students to high standards will increase equity. However, critics have pointed out that having the same standards for all students without providing the support and resources for disadvantaged students to reach those standards is inequitable. Other critics fear that a reliance on assessment and school accountability detracts from addressing the root causes of low achievement.

Overall, Kentucky students have improved their scores on the tests used in the KIRIS and CATS accountability systems and have also improved on the national NAEP tests, although not as dramatically. National studies have demonstrated conflicting evidence about the role of socioeconomic and race/ethnicity variables in student achievement.

Kentucky results, since KERA, indicate that poverty has a substantial negative effect on individual and on school-level test scores. Gender has a significant effect at the individual level. The evidence regarding race, mobility, and disability is mixed at both the individual and school levels. Family structure and giftedness have received



little attention. The correlation of demographic factors and change scores does not reflect the correlations in individual and aggregate absolute scores. In Kentucky, the distinctions between demographic effects at the individual level and those at the school level have not been extensively explored.

Considering the implications of high-stakes testing and the effects of demographic factors on student achievement, several researchers have explored alternative approaches to accountability including longitudinal models emphasizing change scores, statistical techniques to control for demographic variables, and a more comprehensive approach to accountability that include inputs as well as outputs. Research is needed to improve equity in accountability systems by measuring the effects of individual and school-level background variables on student achievement.

## CHAPTER III

### METHODOLOGY

#### Overview of the Study

This study examined and compared the relationship between student background factors and student achievement at the individual student level and at the school level. The measures of student achievement were taken from the Commonwealth Accountability and Testing System (CATS). The study provides the opportunity to compare the association of student background variables to student achievement on both norm-referenced test results and on criterion-referenced test results. Elementary, middle, and high school levels were analyzed independently.

First, the study addressed the extent to which individual background factors predict individual student achievement on reading and mathematics achievement tests. In this case the subjects were the individual students in Jefferson County Public Schools (JCPS). Second, the study addressed the relationship of school demographic composition to aggregate student achievement. In this case the subjects were the individual schools in Jefferson County. The aggregate of the individual student background characteristics formed the school composition variables. Individual test scores are aggregated to create school-level scores as measures of school performance on the norm-referenced tests and on the criterion-referenced tests. In addition, the study addressed the relationship of school composition variables with school performance on

the school-level Academic Index and the CTBS Index, which are used for accountability.

### Sources of Data

This study was a secondary analysis of quantitative data collected by the Kentucky Department of Education (KDE) and by Jefferson County Public Schools (JCPS). The measures of student achievement were the achievement tests used for accountability in Kentucky schools--the Kentucky Core Content Tests (KCCT) and the CTBS. The sources of the data were the electronic files obtained from the Kentucky Department of Education and Jefferson County Public Schools for year 2002.

### *Subjects*

#### *Students*

All of the JCPS students who took the KCCT and CTBS tests in 2002 were potential subjects at the individual student level. Only those students for whom there were complete data in regard to the background variables were included in the analyses for this study. The specifics of any missing cases are addressed in Chapter IV, Results. The numbers of students taking the tests at each grade as calculated from the electronic files obtained from the KDE on November 5, 2002 are shown in Table 1.

#### *Schools*

There are 87 elementary, 24 middle, and 21 high schools in the Jefferson County Public School District included in the CATS accountability system. The elementary schools in Jefferson County include grades K-5. Many elementary schools also have preschool programs, but neither the preschool students nor the early primary students (kindergarten, first, and second grade) were tested for accountability purposes.

Table 1

*Number (N) of JCPS Students Tested per Grade Level*

Year	Grade level								
	3	4	5	6	7	8	9	10	11
2002	7154	7153	7529	7290	6761	6533	7848	6632	5619

*Note.* In twelfth grade, students are assessed only on writing. Therefore, they were not included in this study.

Most JCPS middle schools are grades 6-8, and high schools are grades 9-12. The Brown School, which is K-12, was the only exception. For accountability purposes, however, the Brown elementary, middle, and high school levels are reported separately as though they were different schools. Alternative schools for students in special circumstances (such as adjudication or disciplinary action) or with special needs (such as hospitalized illnesses) were not included in the analysis. Students' scores from those schools were included with the regular schools they would be attending if not in special circumstances (KDE, 2001b).

*Peculiar Aspects of JCPS Student Assignment*

The distribution of students with various background characteristics in the populations of the schools was affected by the student assignment plan in JCPS. Unlike school districts that assign students to schools based solely on students' residence, JCPS has a managed-choice system of student assignment for the purposes of desegregating its schools (JCPS, 2002a). The managed-choice system provides students and their

families with choices of where a child may attend school within parameters set by the school system. Generally, the population at each school was held to a range of 15% to 50% Black in grades 1-12 based on the proportion of Black students in the system, which was about 33%.

At the elementary level, students are assigned a home school by residence location, although that school may not necessarily be the nearest school to their residence. The elementary schools are divided into clusters of 5 to 10 schools per cluster (JCPS, 2001a). Each family has a choice of sending their children to one of several schools in their assigned cluster. The primary limitations on placement within a cluster are space available and the student's ethnicity, e.g., if a neighborhood home school enrollment is 50% Black, then additional Black students wishing to enroll must choose another school in the cluster in which there is space available.

Middle schools and high schools are primarily desegregated by the geographic assignment of students to schools and by magnet schools and programs. School attendance zones are drawn to produce desegregated schools. Since the Black population is concentrated in certain areas of the school district, the attendance zones are not all contiguous to the school, but geographic satellites are assigned to distant schools. Predominantly it is the Black students that are bused across the district, but some gerrymandering results in predominantly White neighborhoods being assigned to predominantly Black neighborhood schools. In addition, there is an open enrollment policy at the high school level. Ninth graders may choose any high school if there is space available, if it supports the desegregation goals, and if the students provide their own transportation.

Only four schools were exempt from the racial balance guidelines of the district. As a result of litigation, the federal court removed the racial guidelines for one of the magnet high schools, and the district subsequently removed the racial balance restrictions from three other magnets. However, only one of those exempted schools was not within the 15% to 50% guidelines in 2001-2002 (JCPS, 2001a, 2001c).

The elementary cluster plan and the geographic assignment are complemented by a variety of magnet schools and programs to which students may apply for admittance. The magnets offer special curricula, programs, or instructional approaches designed to appeal to special interests. At the elementary and middle school levels, many of the magnets are located in predominantly Black neighborhoods and thereby draw White students into those neighborhoods voluntarily. In addition, within schools there are varying degrees of segregation of students (tracking) by academic program or by ability depending on the policies and practices of the individual schools. Within-school tracking was not addressed in this study but is another mediating factor between student background and achievement that could be influential.

Although the redistribution of students by a managed-choice assignment plan results in a school system that is among the most racially desegregated in the country; the schools were very different in the proportions of free and reduced lunch students, gifted and talented students, and students with disabilities--Exceptional Child Education (ECE)-in their respective populations (JCPS, 2001a, 2001b, 2001c). The combination of the self-selection by families, ECE and Advance Program placements, and admission requirements of the magnet schools contributed to creating those disparities.

### *Demographic Factors*

The predictor variables in this study were student background factors. In Bloom's (1980) terms, these are non-alterable givens. The policy issues addressed by this study relate to the extent that these factors affect (or account for) variation in the various achievement measures. In one respect or another all of these factors are aspects of the larger stratification system in the United States (cf. Bowles & Gintis, 1976; Dika & Singh, 2002; Persell, 1977). Because of the hierarchical nature of society, groups located in lower strata based on race, socioeconomic class, or gender are subject to the repressive effects of inequality in society and in school. Typically, they perform less well on various measures of societal success including school achievement tests.

In this study, based on conceptual and feasibility parameters, the factors of gender, race, SES, mobility, family structure, academic giftedness, and disability status--(Exceptional Child Education (ECE))--were analyzed for their relationship to student achievement. The same factors were used in both individual student level analysis and in school-level analysis.

### *Gender*

Gender was included in the analysis because it consistently has a differential association with achievement at the individual level between males and females, and the percentages of males and females vary somewhat from one school to another among schools in this study. For example, in 2001, at the elementary level, the district average for females was 48% with a range from 43% to 54% (JCPS, 2001a).

### *Socioeconomic Status*

SES is the predominant social classification associated with academic achievement. The term that includes both the economic conditions of a person or

family and their social status based on other factors such as parents' education and occupations. This study used only an indicator of income. The most consistently available measure of income is the eligibility of students for free and reduced meals (F&R) based on federal poverty guidelines. It is a gross measure because (a) it does not distinguish among levels of poverty nor among the broad range from those students who are barely above the qualifying level to those who are from extremely affluent families, (b) all families who might be eligible do not apply for the benefit, and (c) it does not necessarily encompass other factors that constitute elements of socioeconomic status and might have a more direct effect on academic achievement, e.g., educational level, values.

### *Ethnicity*

Black students have historically been subject to discrimination and inequality in Kentucky schools. This variable was included to measure enduring effects of racism and inequality.

The KDE records ethnicity as White, African-American, Hispanic, Asian, and other. For this study, only Black (African American) and other (non-Black) students were considered because the Black minority was a substantial proportion of the student population and represented the principal oppressed minority group in Kentucky. As a result of the student assignment process, JCPS schools had a Black student population ranging from 16% to 71% in 2001-2002.

Ethnic classifications other than Black were not analyzed in this study because they comprise a very small proportion of the countywide enrollment. In the 2000-2001 school year Hispanics, the most numerous of those remaining groups, represented only



1.8% of the elementary school population in JCPS. Asians represented 1.0% of the elementary school population. Asians or Hispanics represented sizable portions of the populations at certain schools (JCPS, 2001a, 2001b, 2001c). However, the multiplicity of the origins and ethnicities of those subpopulations in different schools would require a more extensive examination than the scope of this study allows in order to make meaningful comparisons. For example, a high proportion of new immigrant Asian refugees at a regular school would not be comparable to a mixed group of first and second-generation Asian American children at a selective magnet school. Also, during the period these data were collected, students with limited English proficiency were not tested with CATS instruments for the first two years they were enrolled.

### *Mobility*

Mobility is the number of times a student leaves one school and enters another school in the same school district. A high mobility rate, representing frequent changes of schools, addresses, and possible homelessness, is characteristic of families in crisis or disarray and may be a proxy for children whose families are under the most extreme economic stress (cf. Mao et al., 1997; *A Report from the Kids*, 1998).

### *Family Structure*

School registration forms require the name and relationship of the persons with whom the student resides. JCPS staff members transcribe that information to data files. This study compared those students who reside with both their mother and their father in an intact, dual-parent family to those in other family arrangements (including step parents, foster parents, and other relatives). Although most studies have found this variable to have insignificant effects when SES is controlled and although the process

of recording of the JCPS data is quite vulnerable to error, the variable was included because school personnel persistently referenced it.

### *Gifted and Talented*

In JCPS, students are identified as Gifted and Talented or eligible for Advance Program placement based on a process that includes recommendations, test scores, and school records (JCPS, 2000). The classifications are transmitted to KDE, and the students are recorded as Gifted and Talented (G&T) students in the KDE data. In JCPS, most students are not identified as G&T until the fourth grade, although a few (3%) are so identified in the third grade (KDE, 2002b). At the individual level, identification for gifted programs is essentially based on a student's previous achievement, so by definition such students can be expected to perform at high levels on achievement tests.

At the school level, special programs for gifted and talented students result in high-achieving students being transferred from some schools and concentrated in other schools. This process can be expected to affect directly the relative scores of these particular schools.

This classification does not include all academically talented students, only those that were identified by the JCPS procedures. Academically talented or advanced students who are in the first and second grades have not been identified and may or may not be present in a school proportional to the older students. Other students in the third through eleventh grades may have declined participation--a phenomenon that is not necessarily systematic. Therefore, any pattern that reinforces or contradicts the proportions of the identified students (creating bias) cannot be discerned without further investigation.

## *Disability*

Students are identified by the local district as eligible for services in Exceptional Child Education (ECE) based on evaluation procedures in conformance with the requirements of federal legislation. This represents a diverse group of disabilities, some of which might affect academic success greatly and others that would not. At the individual level, the category has a very limited usefulness without an indication of the nature of the disability, but at the school level, this construct represents a different dynamic. Because the greatest percentage of students identified as ECE have disabilities that affect school performance, the percentage of ECE children in a school would be expected to affect the scores directly by the inclusion of those scores in the aggregated school scores and indirectly by any effect the concentration of ECE students might have on other students or on the culture of a school (cf. Wilson, 1987). For example, the number and percent of students who have emotional and behavioral disorders (EBD) can vary considerably from school to school, which could be especially relevant for the overall sense of discipline in a school. Therefore, disability was included in this study because of its relationship to societal stratification and school climate.

## *Achievement Scores*

The outcome variables in this study all derived from the achievement test results that are part of the CATS, which was first implemented in the 1998-1999 school year. The purpose of CATS is to assess school performance, and it includes measures of student achievement and non-cognitive measures. Testing occurs in the spring. The following descriptions of the tests are based on information from the KDE

implementation and interpretive guides (KDE, 2001a, 2001b) and the Kentucky Core Content Test 2000 Technical Report (KDE, 2002a).

CATS is designed so that students at different grade levels take tests in different subjects. In the last year of primary (third grade), in sixth grade, and in ninth grade students take the Comprehensive Test of Basic Skills (CTBS). Fourth and seventh grade students are assessed by The Kentucky Core Content Tests (KCCT) in reading, science, and on-demand writing and by their writing portfolios. Fifth and eighth graders are assessed by the KCCT in mathematics, social studies, arts and humanities, and practical living. Tenth graders take the Reading and Practical Living KCCT. Eleventh graders take the mathematics, science, social studies, and arts/humanities KCCT. Twelfth graders are assessed by the on-demand writing test and evaluation of their writing portfolios.

Because the KCCT and the CTBS both have reading and mathematics components, those sections offered the opportunity to compare the effects of background factors on two different types of tests, criterion-referenced and norm-referenced, in similar subject areas. Therefore, the reading and mathematics KCCT and CTBS results were analyzed because of their comparability and because of the importance of reading and mathematics as core subjects.

#### *Kentucky Core Content Tests (KCCT)*

The KCCT are criterion-referenced tests (KDE, 2002a) created specifically to assess the Kentucky educational standards delineated in the Kentucky Academic Expectations and Core Content for Assessment (KDE, 2002a). They are considered criterion-referenced because the scoring of the test is determined by performance

standards that represent what students are expected to know and be able to do.

In 2001, the Kentucky Board of Education adopted new performance standards for CATS. The performance standards were developed, evaluated, and refined and aligned with the KCCT by a yearlong process that included the involvement of teachers, school administrators, university educators as well as the KDE staff and professional testing consultants. The National Technical Advisory Panel on Assessment and Accountability (NTAPAA) directed the process for identifying cut scores for each performance level represented by each of these categories with the help and input of approximately 1600 Kentucky teachers (KDE, 2002a)

Since the primary purpose of the KCCT is for school accountability rather than student accountability, all students do not take the same forms of the tests. Each form has different questions based on a sampling of sub domains in each content area. Reading, mathematics, science and social studies tests each have 12 forms at each grade level. Each form includes six open-response questions and 24 multiple-choice questions. Therefore, a complete matrix of test questions includes 36 open-response and 144 multiple-choice questions for each grade in each school in reading, mathematics, science, and social studies (KDE, 2002a).

*KCCT Scores.* The scoring of the tests is based on the categories of student achievement that constitute the designations of competency on Kentucky standards. Generally on assessments, student performance is categorized as Novice, Apprentice, Proficient, or Distinguished. On reading, mathematics, science, and social studies assessments, the first two categories are further divided into sublevels: Novice non-performance, Novice medium, Novice high, Apprentice low, Apprentice medium, and

Apprentice high. These are termed *performance standards* because they are considered measures of what students know and what they are able to do. The performance standard categories are the focus of the reports sent to parents and schools (KDE, 2002a).

Each open-response question is given a score of from 0 to 4 points. Each correct response to the multiple-choice questions counts one point. The open-response score is doubled and combined with the multiple-choice score to create the raw score. The raw scores are converted to scale scores by statistical procedures derived from item response theory, a method of test construction. The scale scores range from 325 to 800. The scale scores are converted to the performance standard categories--Non-performing Novice, Medium Novice, Low Apprentice, and so forth--based on the established Kentucky standards (KDE, 2002a).

The scale scores are interval data and more reliable than the performance categories. For this study, therefore, scale scores were used for the analysis of individual student data because of the increase in precision they provide for analytical purposes. Similarly, at the school level, aggregate scale scores were available for each subject; therefore, scale scores were used for analysis of aggregate reading and mathematics KCCT results at the school level in order to compare trends at the individual and aggregate levels.

The scale scores offer the best opportunity for analysis of student achievement that is sensitive to minor differences related to demographic factors. Those are not the scores that are used directly for accountability purposes in school-level achievement, however. For accountability purposes, the KCCT scale scores are converted into an

Academic Index (see Table 2).

Table 2

*Academic Index Formula for Reading, Mathematics, Science, and Social Studies*

Level	Weight	X	Percent of students/100	
Non-Performing	0	X	___%/100	= _____
Medium Novice	13	X	___%/100	= _____
High Novice	26	X	___%/100	= _____
Low Apprentice	40	X	___%/100	= _____
Medium Apprentice	60	X	___%/100	= _____
High Apprentice	80	X	___%/100	= _____
Proficient	100	X	___%/100	= _____
Distinguished	140	X	___%/100	= _____
Academic Index			Total	= _____

*Note.* The information for this table was adapted from the *Kentucky Core Content Tests 2000 Technical Report [Electronic Version]* (KDE, 2002a, p. 9).

*Academic Index.* For accountability purposes, at the school level, an Academic Index is computed in each subject. The percentage of students scoring at each performance level is multiplied by the assigned weight and the results summed to produce an Academic Index (between 1 and 140) for each subject. For example, the formula used to compute the Academic Index for reading, mathematics, science, and social studies is demonstrated in Table 2.

The *school-wide* Academic Index is a composite of the subject academic indices

in reading, writing, mathematics, social studies, science, arts/humanities, and practical living. Following is the school Academic Index formula from the *2001-2002 District Assessment Coordinator Implementation Guide for the Commonwealth Accountability Testing System* (KDE, 2001b, p. 199; Note: KDE uses the asterisk as a symbol for multiplication).

#### School Academic Index Formula

##### Grade 4/5

$$\begin{aligned} & (\text{Reading index} * .20 + \text{Mathematics index} * .20 + \text{Science index} * .15 + \\ & \text{Social studies index} * .15 + \text{Writing index} * .15 + \text{Arts/Humanities index} \\ & * .05 + \text{Practical Living index} * .05) / .95 \end{aligned}$$

##### Grade 7/8/9/10/11/12

$$\begin{aligned} & \text{Reading Index} * .15 + \text{Mathematics index} * .15 + \text{Science index} * .15 + \\ & \text{Social studies index} * .15 + \text{Writing index} * .15 + \text{Arts/Humanities index} \\ & * .075 + \text{Practical Living index} * .075) / .90 \end{aligned}$$

The formulas for the elementary grades 4/5 vs. the secondary grades 7-12 academic indices give a greater weight to reading and mathematics at the elementary level than at the higher grades and give a higher weight to the non-cognitive components at middle and high school than at the elementary level. The school Academic Index is subsequently combined with the CTBS results (see next section) and non-cognitive factors (attendance, drop-outs, and successful transition to adult life) to create the school's Accountability Index by which school improvement is ultimately judged for rewards and sanctions.

#### *Comprehensive Test of Basic Skills (CTBS)*



CATS requires the Reading/Language and Mathematics components of the CTBS/5 Survey Edition (KDE, 2001b). The questions are multiple-choice. The Level 13 test given in Kentucky at third grade has 50 items in reading and 30 items in mathematics. The Level 16 test given at sixth grade has 55 items in reading and 30 items in mathematics. The Level 19 test given at ninth grade has 55 items in reading and 25 items in mathematics (KDE).

The results of the CTBS tests are reported in Normal Curve Equivalents (NCE), scale scores, and National Percentiles. A Total Battery score, which is a composite of the Reading/Language and Mathematics components, is also reported and is converted into a norm-referenced index by KDE. The percent of students performing in each quartile on the Total Battery is multiplied by performance level weights to create a Norm-Referenced Test (NRT) Index on the same scale as the KCCT Academic Index (KDE, 2001b) as in Table 3.

In the current study, the NCE scores in reading and mathematics were used in the analyses at the individual student level and at the school level because they represent continuous, interval data similar to the KCCT scale scores. The NRT Total Battery Index is an academic component (5%) of the school Accountability Index. Because both the Academic Index and the NRT Index are the academic components of the Accountability Index formula, they were analyzed in addition to the aggregate reading and mathematics scores.

Table 3

*Formula for National Norm Referenced Index*

National percentile	Weight	X	Percent of students/100	=	_____
1-24	0	X	____%/100	=	_____
25-49	60	X	____%/100	=	_____
50-74	100	X	____%/100	=	_____
75-99	140	X	____%/100	=	_____
NRT Index			Total	=	_____

*Note.* The information for this table is from the *Commonwealth Accountability Testing System Spring 2001, Interpretive Guide, Kentucky Evaluator's Edition* [Electronic version], (KDE, 2001a, p. 14).

Variables in the Study

For both the independent and the dependent variables in this study, there are two subsections describing variables: one at the individual student level and the other at the school level. The same variables were used in the separate elementary, middle, and high school analyses.

*Independent Variables*

The independent variables represent social and demographic background characteristics of the students in the student level analyses. In the school-level analyses, they represent social and demographic characteristics of the pupil population of each school.

### *Individual Student Level*

Except for mobility and family structure, the student level (demographic) variables were obtained from KDE. Local districts, including JCPS, collect the student background data and local school staff records the information in the KCCT Student Response Booklets. These data then become part of the KDE database. In contrast, mobility and family structure status are collected by JCPS staff for district purposes and are available only from JCPS data files.

For the individual student level, all of the predictor variables were nominal except for mobility. The variable, variable label code, operational definition, and level of measurement for each predictor are given below.

*Female (Female).* For this nominal variable, males were recorded as 0, females as 1.

*Socioeconomic Status (SES).* SES was measured at the ordinal level. Students who were eligible for free meals were coded 1. Those eligible for reduced meals were coded 2. The other students were coded 3. As noted above, this was a gross measure of economic class.

*Black (Black).* Ethnicity was represented by Black and was measured as a categorical construct; Black students were coded 1, White, Asian, Hispanic and other students were coded 0.

*Mobility (Mobil).* Mobility was the lone ratio level variable at the individual level. It represented the number of times a student left and reentered any school in the district the previous year. Students who stayed in one school for the entire school year had a mobility of 0. The mobility of students who changed schools during the school

year was the number of the student's reentries--1, 2, 3, 4, ....

*Dual-parent Family (Two Parents).* Students whose entry forms indicate they were residing with both their mother and father (not step parents or other guardians following the distinctions made by JCPS in Muñoz & Dossett, 2000) were coded 1. Students residing with one parent or other guardians were coded 0.

*Gifted and Talented (Gifted).* Students identified as G&T were coded 1. Other students were coded 0.

*Disability (ECE).* Students who were identified as having an IEP (Individual Educational Plan) for disability were coded 1. Other students were coded 0.

#### *School Level*

The school-level demographic (school composition) predictor variables addressed mirrored those addressed in the individual student analysis. The variables represented the percentage of students in the school identified by each demographic factor except for mobility, which is a rate devised by JCPS. The school-level data referred to the composition of all the grades at each school even though only the third through twelfth grades were included in testing. At the elementary school level, therefore, students in grades k-5 were included in the school demographic data. The school-level demographic data were secured from the JCPS data files except for the data on G&T students, which was from the KDE database.

The variables that were measured as a simple percentage--the number of students in each classification divided by the total number of students in the school multiplied by 100--constitute a ratio scale, as does the mobility rate. The variable, variable label, and operational definition are given for each construct.

*% Females (%Female)*. This variable was simply the percent of female students in each school.

*% Higher Socioeconomic Status (%HiSES)*. This variable was the percent of students in a school that were not participating in the free or reduced meals program.

*% Blacks (% Black)*. This was the percent of students in a school who were identified as Black.

*Mobility Rate (Mobil Rate)*. The mobility rate was a ratio of the number of reentries (excluding R1 and R6) divided by the total enrollment of the school written as a percentage (JCPS, 2001a).

*% Dual-parent Families (%Two Parents)*. This was the percentage of students whose entry forms indicate they were residing with both their mother and father (not step parents or other guardians).

*% Gifted & Talented (%Gifted)*. This was the percentage of students in the school that were identified as G&T. At the elementary school level only third, fourth, and fifth graders are so classified. However, for consistency with the other composite variables, the number of G&T students (in third, fourth, and fifth grades) was divided by the total number of students in the school (k-5) to compute the percentage.

*% Disability (%ECE)*. This variable was the percentage of students in a school that were identified as ECE by having an IEP.

### *Dependent Variables*

The dependent or outcome variables in this study were the CATS reading and mathematics measures of achievement for JCPS in 2002.

### *Individual Student Level*

Both the KCCT scale scores and the CTBS NCE scores are measured at the ratio level. The variable name, variable label code, and operational definition are given.

*CTBS Reading (CTBS READ).* Students were tested with the CTBS/5 Survey Edition at the third, sixth, and ninth grades. The students' individual NCE scores on the reading tests were the dependent variables.

*CTBS Mathematics (CTBS Math).* As with reading, the CTBS mathematics test was given in the third, sixth, and ninth grades. The students' individual NCE scores on the mathematics tests were the dependent variables at each grade.

*KCCT Reading (KCCT Read).* Students were tested by KCCT in reading in the fourth, seventh, and tenth grades with a test that is specific for each of those grades. The individual reading scale scores for the students in those grades were the dependent variables.

*KCCT Mathematics (KCCT Math).* Fifth, eighth, and eleventh grade students take the mathematics KCCT designed for their grade level. The individual mathematics scale scores for the students in those grades were the dependent variables.

#### *School Level*

The dependent variables at the school level were parallel to those at the individual student level, except that in addition to the reading and mathematics tests, the NRT Index and Academic Index were included because of their importance in school accountability. All of the variables were measured at the ratio level. The variable name, variable label code, and operational definition are given.

*CTBS Reading (SCTBS Read).* School-level scores on the Reading CTBS were the average NCE scores of the students who took the CTBS Reading test at each school.

For elementary schools the average third grade NCE reading score was the school CTBS Reading score, for middle schools the average sixth grade score, for high schools the average ninth grade score.

*CTBS Mathematics (SCTBS Math).* School-level scores on the Mathematics CTBS were the average NCE scores of the students who took the Mathematics CTBS at each school. The average NCE reading score for the third graders was the school Mathematics CTBS score for elementary schools, for sixth graders at the middle school, and for ninth graders at high school.

*KCCT Reading (SKCCT Read).* The school-level Reading KCCT scores were the average scale scores of all students in a school who took the reading test at each level (elementary, middle, and high school levels). Therefore, for elementary schools the school KCCT Reading score was the average scale score for the fourth graders who took the reading test, for middle schools the school reading score was the average reading score of the seventh graders, and for high schools the school reading score was the average reading score of the tenth grade students.

*KCCT Mathematics (SKCCT Math).* The average scale score on the Mathematics KCCT for all the students who took the test at each level was the school-level variable. For elementary schools, therefore, the school KCCT Mathematics score was the average scale score for the fifth graders, for middle schools it was the average mathematics score of the eighth graders, and for high schools it was the average score of the eleventh grade students.

*Academic Index (Acad Index).* The Academic Index for each school is a scale ranging from 0-140. It is computed by a weighted formula (see description, p.109,

above) that includes the Academic Indices of the KCCT assessments (which are also weighted percentage scales, see Table 2, p. 108) in all subjects (tests and portfolios) for all grades in a school.

*National Norm-Referenced Test Index (NRT Index)*. This index is computed by the KDE based on the percent of students in a school (at third, sixth, and ninth grades) who score in each quartile of the National Percentile Range on the CTBS/5 Total Battery (see Table 3, p. 111). Like the Academic Index, this index represents a weighted percentage scale ranging from 0-140.

### Empirical Research Questions

The data were analyzed separately for the different school levels--elementary, middle, and high school. For each of the three school levels, the following research questions were addressed.

Research Question 1. At the individual student level, to what extent do demographic variables:

- a. Female (Female)
- b. Socioeconomic Status (SES)
- c. Black (Black)
- d. Mobility (Mobil)
- e. Dual-parent Family (Two Parents)
- f. Gifted and Talented (Gifted)
- g. Exceptional Child Education (ECE)

predict individual student achievement on

- a. CTBS Reading (CTBS Read)



- b. CTBS Mathematics (CTBS Math)?
- c. KCCT Reading (KCCT Read)
- d. KCCT Mathematics (KCCT Math)

Research Question 2. To what extent do the school-level, composite demographic variables:

- a. % Females (%Female)
- b. % Higher Socioeconomic Status (%HiSES)
- c. % Blacks (%Blacks)
- d. Mobility rate (Mobil Rate)
- e. % Dual-parent Families (%Two Parents)
- f. % Gifted and Talented (%Gifted)
- g. % ECE (%ECE)

predict aggregate school performance on

- a. CTBS Reading (SCTBS Read)
- b. CTBS Mathematics (SCTBS Math)
- c. KCCT Reading (SKCCT Read)
- d. KCCT Mathematics (SKCCT Math)
- e. Norm-Referenced Test Index (NRT Index)
- f. Academic Index (Acad Index)?

#### Statistical Analysis

The analyses in this study focused on two distinct, but related, levels. The first focus was the proportion of the variance in individual level student achievement and the relationship of the variance to selected student background factors. The second focus

was the extent of the relationship of school-level composition factors with school-level achievement (school performance). Multiple regression analysis was used for both of these levels to examine the effects of demographic variables on achievement, as measured by the reading and writing portions of the KCCT and of the CTBS and by the Academic Index and the NRT Index. This study used the SPSS 11.5 computer software for the statistical analysis, which also provided descriptive statistics, correlations, and tests of significance for the regressions (Shannon & Davenport, 2001).

The demographic and social variables in this study were chosen because of their known association with student achievement and their availability in school records. However, a high degree of multicollinearity among the independent variables was expected, and in fact, a major purpose of the study was to determine which variables affect student achievement. In effect, it was an exploratory study to develop a model representing the influence of those socio-demographic variables on student achievement and to compare the relative effects of the model components at the individual and school levels. Forward stepwise multiple regression is considered a model-building procedure that is useful both for eliminating superfluous variables and for exploratory efforts to identify variables for future research (Tabachnick & Fidell, 2001).

In forward stepwise entry of the independent variables into the regression formula, the variable with the highest simple correlation with the dependent variable is entered first. Then the variable with the next highest semi-partial correlation is entered. This continues until the subsequent variable entered does not add significantly to the equation.

As a strategy to examine the effect of each variable on the dependent variable,

the default criteria for entry of variables used by SPSS was modified. The default entry criterion of  $F$  to enter and  $F$  to remove variables was overridden to allow entry of all predictors. This allowed the calculation of  $R$  squared and change in  $R$  squared for each predictor variable. For reporting purposes, regression equations presented in the results identify predictor variables that account for a minimum of 2% of incremental variance in the dependent variable. The latter criterion has been suggested as a criterion for a small effect size (J. Cohen, Cohen, West, & Aiken, 2003).

Stepwise multiple regression is criticized for its over-reliance on chance and inherent over fitting of the data. Both of these problems refer to the use of stepwise regression in inferential statistics because sampling may result in decisions being based on minor differences across samples, so that an equation fitting only that sample may not generalize to the population. Those problems are discussed in the interpretation and discussion of the results in Chapters IV and V.

The results of this study cannot be generalized to other school districts, yet generalizing the results to other years of testing in the same school district was an ultimate goal of the study. Comparison of the results against other years in future research is necessary for validation of the resultant regression equations from this study.

The current study was exploratory, in that its purpose was to determine the extant empirical relationships among seven predictor variables and the CATS accountability system, and stepwise multiple regression is recommended for that purpose (see Tabachnick & Fidell, 2001). On balance, given the parameters of the data set and the purpose of the study, the stepwise procedure can be justified as the most appropriate procedure under these conditions.

### *Individual Level*

There were seven student-background variables included in the regression analysis at the individual level. Because there were approximately 7,000 students tested in each grade, the number of subjects compared to the number of variables provided more than adequate statistical reliability (Stevens, 2002).

### *School Level*

For the analysis of the effects of school composition on school-level scores, there were 87 elementary schools. With seven predictors, the ratio of subjects to predictors was 12.42, which falls below the parameters for a reliable regression equation according to Stevens (2002). At the middle and high school levels, there were only 24 and 21 schools respectively. With seven school composition aggregate variables used in the regression, the ratio of subjects to predictors in the middle and high school-level regressions was an even more radical violation of Stevens' criterion of 15 subjects per predictor. The resolution of this problem involved several considerations.

First, this study did not infer from a sample to a larger population as in traditional inferential statistics, but rather included all of the JCPS schools in the analyses. However, the probability levels computed in multiple regression analyses based on the principles of inferential statistics were still guides to the strength of the relationships among the variables in the population. Despite the proviso that for a population, a relationship is meaningful or not simply by its de facto strength, there is no prima facie relationship that is as easy to interpret as the computed probability level and effect sizes that are products of the inferential analyses.

Furthermore, even for the description of a population, it is still appropriate to reduce the number of predictor variables in order to increase the validity of the regression equation, particularly in data sets where the ratio of subjects to predictor variables is suspect (Stevens, 2002). That situation applies in the current study when the unit of analysis was the school, and especially so for the middle school and high school analyses. The number of variables was narrowed down by the use of forward stepwise multiple regression at the school level to identify predictors that contribute at least 2% of increment in variance to the regression equation. In addition, the regression equations were generated with full sets of predictor variables in the event that some variables that were not important predictors at the individual level turned out to be important at the school level.

### *Multiple Regression Models*

The individual level analysis required four multiple regressions at each school level (elementary, middle, and high). The school-level analysis involved six multiple regressions--four for the aggregate achievement scores and two for the performance index scores at each of these school levels. Therefore, there were a total of 30 planned regressions. Schematic representations of the planned multiple regressions at each level for each research question are shown in Figures 1 and 2. To review, the core research questions follow:

1. To what extent do individual student background factors--gender, SES, ethnicity, mobility, family structure, gifted and talented status, and disability--predict individual student achievement?
2. To what extent do school composition variables--proportions representing

gender, SES, ethnicity, mobility, family structure, gifted and talented status,  
and disability--predict school performance?

*Human Subjects Approval*

The letter of approval from the Human Studies Committee is attached in  
Appendix A.

Individual Student Level

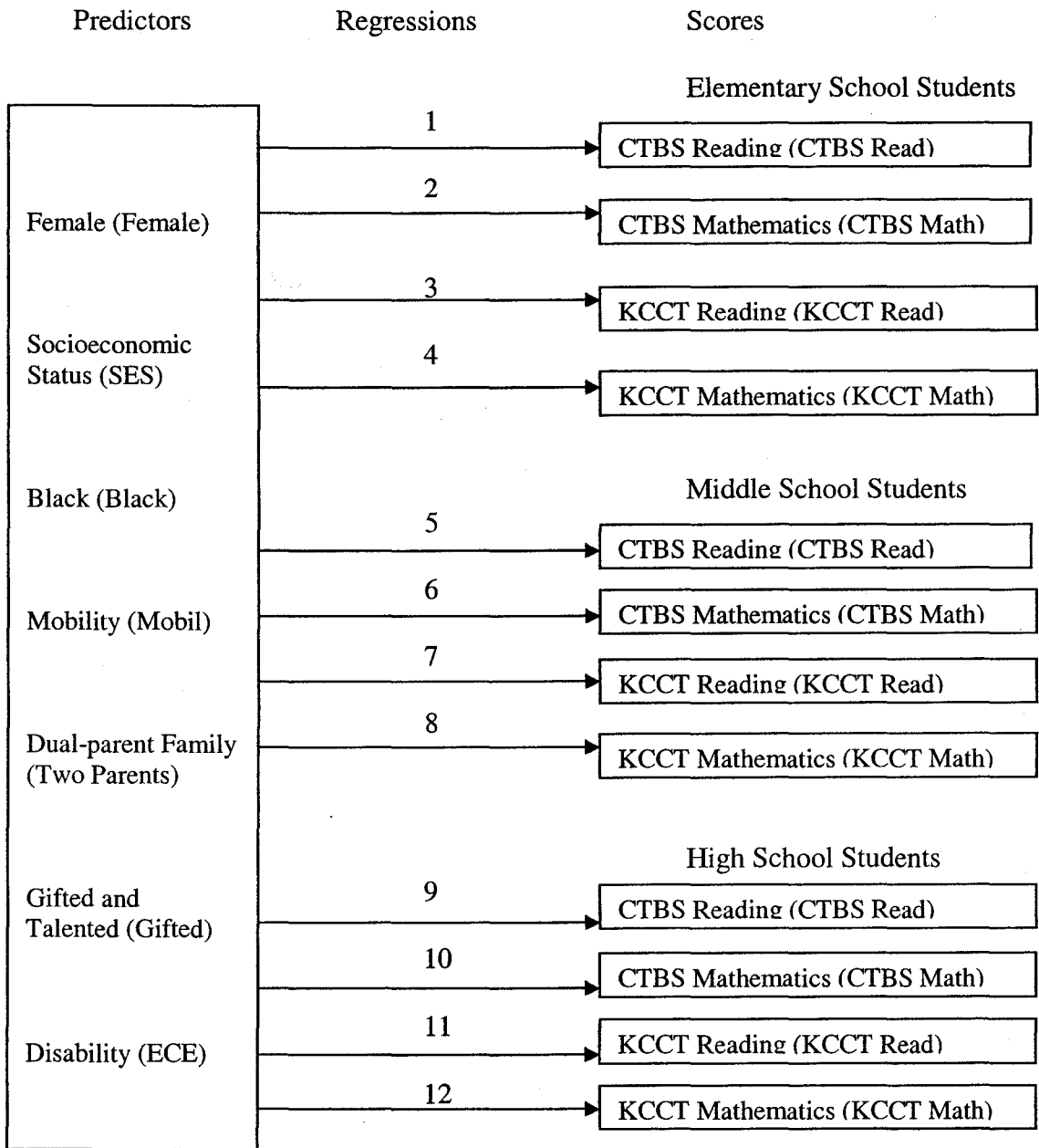


Figure 2. Step-wise regressions for Research Question 1.

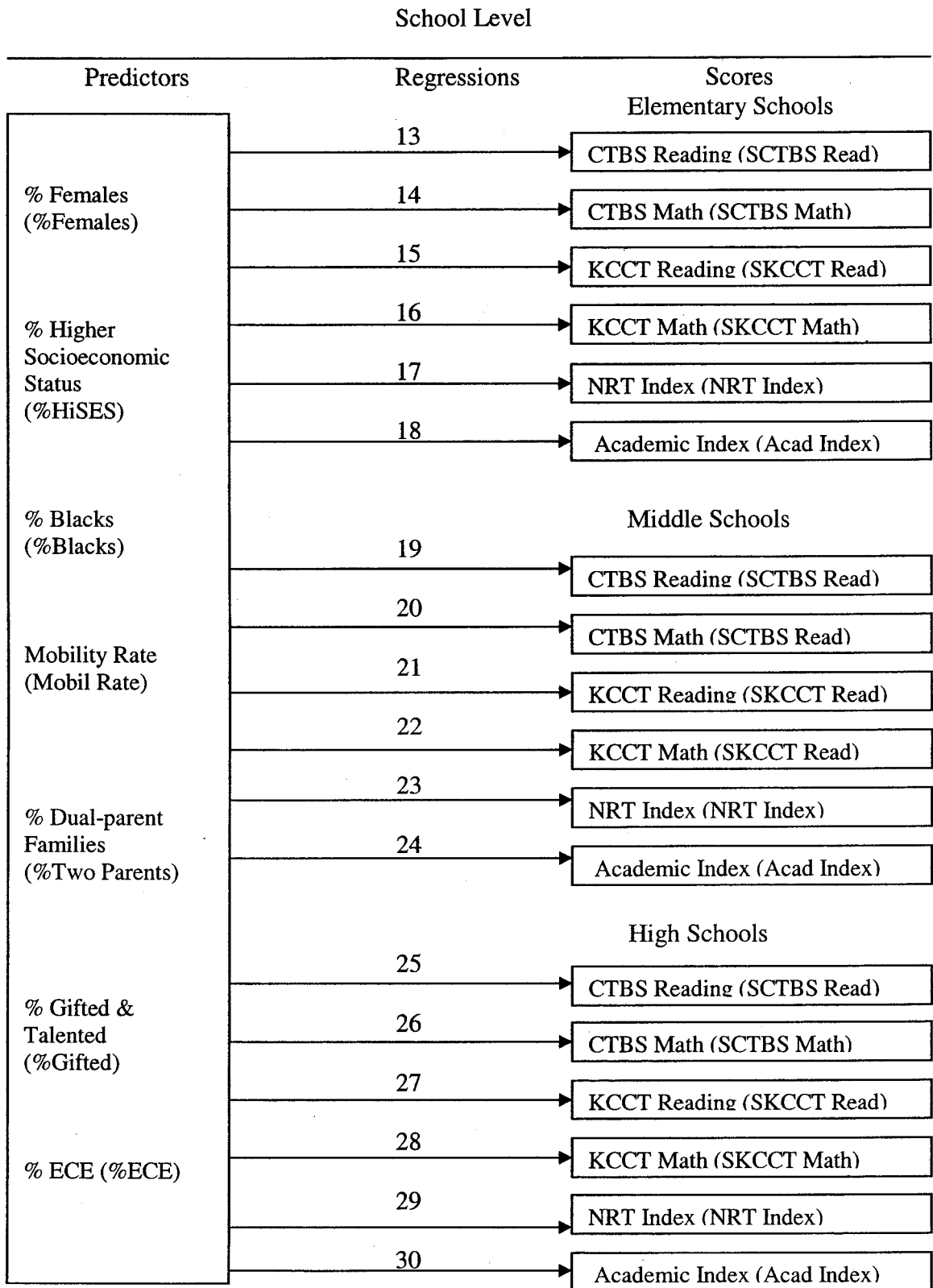


Figure 3. Step-wise regressions for Research Question 2.



## Summary

This study was a secondary analysis of data collected by the Jefferson County Public Schools (JCPS) and the Kentucky Department of Education (KDE). It examined the relationships between student background factors and student achievement at the individual student level and at the school level on achievement tests used for accountability in the Commonwealth Accountability Testing System in Kentucky. The tests were the KCCT, which is a criterion-referenced test aligned with the Kentucky standards, and the CTBS/5 Survey Edition, a national norm-referenced test.

The individual student predictors included in the analysis were gender, socioeconomic status, ethnicity, mobility, family structure, giftedness, and disability. The school-level demographic predictor variables addressed were parallel, representing the percentages of students in each category that composed the school population and school mobility rate.

Each level of students and schools--elementary, middle and high schools--was analyzed independently for the two levels of data aggregation (individual and school) addressed by the research questions. The data were analyzed by multiple regression using SPSS 11.5, which also provided descriptive statistics and tests of significance.

The study first addressed the association of individual student demographic variables with individual student achievement on the KCCT and CTBS reading and mathematics tests.

Second, the study examined the extent to which school population demographic (composition) variables predict aggregate student achievement, school-wide scores, in reading and mathematics on the KCCT scores and the CTBS scores. The study also

examined the relationship of those demographic variables with the CATS measures of school performance--the Academic Index and Norm-Referenced Test Index.

## CHAPTER IV

### RESULTS

#### Introduction

The purpose of this study was to contribute to an understanding of the interrelatedness of demographic factors and student achievement and to inform evaluation of accountability formulas and processes. The specific objectives were to ascertain and compare (a) the effects of individual student background factors on individual student achievement and (b) the effects of demographic school composition factors on composite school performance measures. These investigations were a secondary analysis of data collected by the Jefferson County Public Schools (JCPS) and the Kentucky Department of Education (KDE). The student achievement measures and the school performance measures included the spring 2002 achievement test results from the Kentucky Commonwealth Accountability Testing System (CATS)--the reading and mathematics Kentucky Core Content Tests (KCCT), the reading and mathematics CTBS/5 Survey Edition (CTBS), and the Academic and Norm Referenced Test (NRT) Indices that were computed from the KCCT and CTBS test results (as described in Chapter III).

Multiple regression analysis was used to determine the relationships outlined in the core research questions, which follow.

At each level: Elementary, Middle, and High School,

1. To what extent do individual student background factors--gender, SES, ethnicity, mobility, family structure, gifted and talented status, and disability --predict individual student achievement?
2. To what extent do school composition variables--proportions representing gender, SES, ethnicity, mobility, family structure, gifted and talented status, and disability--predict school performance?

The null hypothesis for these research questions was that none of the variables in either question demonstrate a relationship with student achievement or school performance.

The organization of this chapter is directed at answering the core research questions. First, descriptive statistics of the independent and dependent variables are reported as context. Then, multiple regression analyses at the student level and at the school level are presented to address the effects of background variables on test results.

### Descriptive Context

This study included all the schools and essentially the population of the third through eleventh grade students in the Jefferson County Public Schools. Therefore, the population parameters are reported and include frequencies, range, means, standard deviations, and variance as appropriate, depending on whether the data were at the nominal, ordinal, interval, or ratio level of measurement. These tables are located in Appendices B, C, D, and E. In addition, correlation matrices are reported for between-variable comparisons because they constitute the basis of the multiple regression analyses that form the core of this study.

### *Student Level Parameters*

In this section, the parameters of the student background characteristics are

presented and trends noted for elementary, middle and high school-level students.

Within the three school levels, the parameters are reported by individual grades because students take different tests at different grades and different numbers of students were involved with each test.

Five of the background variables were nominal dichotomies: gender, ethnicity, disability, giftedness, and family structure. Socioeconomic status (SES) was measured as ordinal data because students were coded as qualifying for free lunch (1), reduced-price lunch (2), and not qualifying (3). Frequencies are reported for those six variables. Mobility was a ratio level variable, and the means and standard deviations are reported as well as the frequencies for that variable.

#### *Missing Data*

Because the primary purpose of this study was to examine the relationship of background variables to CATS achievement test scores, only those students for whom there were complete background data were included in the analyses. Those cases with missing data for the variables under study were excluded, which reduced the number of subjects from the original population.

Most of the student data--test scores, ethnicity, gender, ECE, gifted, and SES--were obtained from the KDE electronic student data files. The mobility and family structure data were obtained from JCPS. The files were merged using the student ID numbers. There were instances of missing data in both the KDE and the JCPS data. That included missing student ID numbers in the KDE data, which prevented any merger of those cases. The merged data had fewer cases in each grade than the original KDE data. The largest proportion of missing data was from the family structure and

mobility data from JCPS. According to Tabachnick and Fidell (2001), this did not represent a major problem for most of the grades because the total missing data represented less than, or slightly over, 5.0% (grades 3, 4, 5, 6 and 10). Grades 11 and 8 were missing slightly more at 6.0% and 6.8%, respectively.

For the ninth grade, the discrepancy was more pronounced with more than 23% fewer complete cases than in the original KDE data. Part of that difference, 1.7% was missing data among the KDE variables similar to the pattern of missing data at the other grades, but 1589 cases were missing from the JCPS family structure and mobility data. Examination of the data revealed that the missing cases were the block of 1589 cases with the highest student ID numbers and that the same cases were missing from both the family structure data and the mobility data. For those reasons, it appears that the missing data were caused not by systemic factors related to the students' characteristics but that a block of students was omitted due to clerical error.

Incorrect coding of some of the students caused a different missing-data problem. During data analysis it was discovered that one of the JCPS middle schools had failed to code its seventh grade gifted and talented (G&T) students. According to the school's principal (personal communication, May 16, 2003), the seventh grade had a high proportion of G&T identified students--25% rather than the 3% reported by KDE--which was supported by the data for previous and subsequent years. That means about one fourth of the seventh grade students were incorrectly coded in that school. Since it was unknown which students were miscoded, all seventh grade student scores and independent variables from that middle school were omitted from the student-level analysis, 268 cases. Because there remained more than 6300 seventh grade students in

the analysis, this omission was expected to have essentially no effect on the results of the student-level analysis.

### *Independent Variables*

*SES.* Socioeconomic status was measured by the qualification of the students for federally subsidized meals (free and reduced lunch). In the elementary grades (Table B1) over 55% of the students qualified for free and reduced meals. The proportions decreased through middle school (Table B4) and high school (Table B7) until in grade 11 only 29% of the students were qualified. Most of the qualified students were eligible for free subsidy while a smaller number qualified for reduced price meals. That number on reduced price meals decreased from a high of 9.4% in the fourth grade to 6.7% in the eleventh grade.

*Mobility.* Mobility is a measure of how many times a student changes schools during the school year. At all grade levels (Tables B2, B5, B8), most students had a mobility of 0 (91.5% to 94.7%) because they remained in the same school during the entire school year. At the elementary level (Table B2) about 5% changed schools only one time, and only a small percentage, about 1%, changed schools more frequently than once during the year. In middle and high school (Tables B5 and B8), the pattern was similar except that the mobility was somewhat higher. Although the numbers of students who change schools during the year was less than 10%, the mobility was increasingly higher at higher grades. Table B10 shows the means for mobility were 25% higher in middle school and about 75% higher in ninth and tenth grades school than in the elementary grades, although at eleventh grade mobility was similar to that at elementary.

*Gender.* Examination of Table B3 shows that elementary students were almost equally divided between male and female with slightly more males than females. In grades 6 through 10 (Tables B6 and B9, respectively), both the proportions of males and of females were very close to 50%, but there were 4% more females than males in the eleventh grade.

*Ethnicity.* The percentage of Black students was stable in the elementary grades (Table B3) at about 37%. The percentage of Black students gradually becomes fewer at the higher grades (Tables B6 and B9). At the eleventh grade, there were only 26.4% Black students.

*Disabilities.* The percentage of students identified as ECE was about 10% in third grade and slightly larger, over 11%, in grades four and five (Table B3). It was about 10% throughout middle school (Table B6). In contrast, the percentage of ECE students was substantially lower in high school (Table B9)--only 5.7% in the eleventh grade.

*Giftedness.* Only three percent of the students were identified as gifted and talented (G&T) in the third grade, but that increased to 8.0% and 9.6% in the fourth and fifth grade when students are formally selected for the programs (Table B3). The percentage of students identified as gifted and talented increased across the middle school grades (Table B6) from 8.7% to 13.1%. In high school (Table B9), gifted percentage was increasingly higher at each grade from 10.1% in ninth grade to 14.8% in the eleventh grade; however, the total numbers of both students and gifted students was lower in the eleventh grade.

The number of gifted students identified was increasingly larger within a school



from year to year. In elementary school, more gifted students were in fifth grade than in fourth. Then in middle school, the number was lower at sixth grade than at fifth, but progressively greater at seventh and eighth grades. At ninth grade the number was again lower, increased substantially at tenth grade, then was only a little lower at eleventh (although the percentage was higher).

*Family Structure.* Contrary to traditional expectations about families, registration forms indicate that only 42.2% to 45.5% of elementary (Table B3), middle (Table B6), and high students (Table B9) in JCPS live with both their mother and father. The other students' registration forms indicate they live with single parents, stepparents, other relatives, or in some other arrangement. There was no apparent trend in the fluctuation of the percentages, which were similar at all three levels.

#### *Dependent Variables*

Students are tested in third, sixth and ninth grades with the CTBS/5. They take a KCCT Reading test in fourth, seventh, and tenth grades and a KCCT Mathematics test in fifth, eighth, and eleventh grades. The KCCT tests at the different grade levels are independent and assess skills and content specific to the different levels. The CTBS/5 test scores are reported in normal curve equivalents (NCE) scores. The KCCT tests results are reported in scale scores that range from 325 to 800.

For this study, only those students with complete data were included in the calculations of the mean and standard deviation because those values were used in the correlation and regression analyses. As mentioned above, one school's seventh grade students were omitted from the analysis. That school's mean KCCT Reading score for those seventh grade students was higher than the district mean; therefore, the mean for

the district-wide seventh grade KCCT Reading score without them was slightly lower than it would be had those students been included.

Tables C1, C2, and C3 report the number, range, mean, and standard deviation for the tests given to students in the various grades. The test scores show different trends for the CTBS and the KCCT.

For the KCCT tests, the middle and high school scores appear to be lower than elementary scores. However, the KCCT are criterion-referenced tests, and the test scores are not comparable among subjects or grades because performance standards and cut-off scores are set independently (e.g., one test may be relatively more difficult than others). Therefore, results for each KCCT test can only be compared to the results for the same test at the same grade. For example, the KCCT scores in mathematics were consistently higher than those in reading, but that does not mean that students were more skilled in mathematics because the scores are based on different, discipline-specific criteria rather than comparisons of students.

The standard deviation and the variance, which are measures of the dispersion of the scores, were much larger at the high school level than at the middle and elementary school levels for both the KCCT Reading and Mathematics tests. This represents a difference in the spread of the students' scores at the different grades, which may be attributable to either the design of the tests or the differential achievement of the students.

In contrast to the KCCT results, the CTBS test scores can be compared across grade levels and subjects because the normal curve equivalent scores are based on comparison with students at each grade in the norming population. Therefore, on both

the CTBS tests, the scores indicate students' achievement relative to a normal curve. On the CTBS, the mathematics scores were slightly higher than the reading scores in third grade. In sixth grade, reading was slightly higher than mathematics, and in high school, reading was substantially above average, while mathematics was slightly below average. The CTBS middle school scores were substantially lower than those of elementary and high school students. For the CTBS, the standard deviations and variances were similar at elementary, middle, and high school.

### *Correlation Matrices*

For each of the 12 CATS achievement tests, the correlations among the variables are given in Tables 4 through 12. These data describe the total populations of the students in the tested grades in JCPS in 2002. Because of the large numbers of subjects, essentially all correlations were statistically significant at very low probability values. Consequently, the usefulness of statistical significance is limited to eliminating those variables with very negligible correlations from consideration of effects or of covariance.

The strength of the correlation itself is more indicative of the strength of effects related to the variables. Typically, guidelines for interpreting the strength of the correlation are similar to those by Best and Kahn (1989), which suggest that correlation below .20 is negligible, .20 to .40 is low, .40 to .60 is moderate, .60 to .80 is substantial, and .80 to 1.00 is high to very high.

For the purposes of this study, the author considered the possible impact of the variables on a test score the important criterion for deciding if the relationship were important in practical terms. Because the basic analysis concerns the effects of

demographic variables on test scores, a few points can make a difference in accountability decisions. Therefore, based on her calculations of the impact of the relationships on test scores, the author has adopted the following standards for interpretation of the scores: Correlations below .13 that are statistically significant are considered to have *negligible* correlation. Those between .13 and .25 ( $r^2 = .02$  to  $.06$ ) are considered to be practically important and to have a *low* correlation because a correlation of .13 representing 2% of the variance in a test score could be important in an accountability system. Those with correlations between .25 and .50 ( $r^2 = .06$  to  $.25$ ) are considered *moderate*, .50 to .75 ( $r^2 = .25$  to  $.56$ ) are *substantial*, and above .75 ( $r^2 = .56 - 1.0$ ) are *high*.

Overall the statistical tests of probability show that all of the socio-demographic variables had a significant correlation with students' test scores at all grades except for Female, which was not significant for the CTBS Mathematics test at third grade nor the KCCT Mathematics tests at grades 5 and 11. Generally, the student-level correlations were not very large--the greatest correlation with a test score ( $r = .426$ ) was between SES and the CTBS Mathematics test in the sixth grade.

The complete set of relationships between the socio-demographic variables and between the socio-demographic variables and student test scores are presented in Tables 4 through 12. Notable findings are highlighted below.

*Elementary Students.* The CTBS Reading and Mathematics tests were administered to the same set of students in third grade so the correlations are exhibited together in Table 4. Among the correlations of demographic variables with the third grade CTBS Reading test, the greatest was with SES--the higher the student's income

(1-3), the higher the test score. Black and ECE have low-moderate, negative correlations, and Two Parents (student has a two parent family) has a low-moderate, positive correlation. The correlations of the third grade CTBS Mathematics test scores show the same pattern as the third grade CTBS Reading test: moderate, positive correlations between the mathematics scores and SES and Two Parents; moderate, negative correlations between the mathematics scores and Black and ECE; and negligible or insignificant correlations for all the other relationships.

The demographic variables in third grade shown in Table 4 reveal that the strongest relationships were the moderate, positive correlation between SES and Two Parents and the negative correlations between Black and both SES and Two Parents. Gifted had a negligible correlation with SES.

The two highest correlations between the Fourth Grade KCCT Reading score and the demographic variables in Table 5 were moderate, .319 with SES and .296 with Gifted. The positive correlation for Two Parents and the negative correlations for Black and ECE with the KCCT Reading score were in the low range. Among the socio-demographic variables, there was a moderate positive correlation between students' SES and Two Parents and moderate negative correlations between Black and both SES and Two Parents. Other relationships were very low or negligible.

Table 4

*Correlations for Independent Variables and Third Grade Student CTBS Reading and Mathematics Scores*

Variable	CTBS Read	CTBS Math	Female	Black	ECE	Gifted	SES	Mobil	Two Parents
CTBS Read	--		.070**	-.267**	-.217**	.181**	.347**	-.140**	.231**
CTBS Math		--	.005	-.300**	-.242**	-.185**	.358**	-.131**	.250**
Female			--	.002	-.100**	.028**	.011	-.004	.002
Black				--	.059**	-.082**	-.400**	.101**	-.371**
ECE					--	-.039**	-.110**	.093**	-.055**
Gifted						--	.128**	-.034**	.072**
SES							--	-.155**	.418**
Mobil								--	-.124**
Two Parents									--

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

Table 5

*Correlations Between Independent Variables and Fourth Grade Student KCCT Reading Scores*

Variable	KCCT Read	Female	Black	ECE	Gifted	SES	Mobil	Two Parents
KCCT Read	--	.121**	-.187**	-.164**	.296**	.319**	-.079**	.165**
Female		-	-.001	-.136**	.014	.005	-.009	-.006
Black			--	.084**	-.154**	-.391**	.088	-.370**
ECE				--	-.094**	-.131**	.077**	-.081**
Gifted					--	.237**	-.058**	.150**
SES						--	-.147**	.423**
Mobil							--	-.100**
Two Parents								--

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

Table 6 shows that in fifth grade, the correlations of background variables with the KCCT Mathematics test scores were slightly stronger in general than those for the fourth grade KCCT Reading test. Gifted and SES had a moderate, positive correlations, and Black and ECE had low-moderate, negative correlations with the KCCT Mathematics scores. The correlations among the fifth grade background variables were similar to those in fourth grade (Table 5).

*Middle School Students.* Sixth grade students were tested on CTBS Reading and CTBS Mathematics tests. The correlations are shown together in Table 7. On both tests the correlations of demographic variables with the test scores were stronger than were evident on the third grade CTBS tests. All variables except Female show a low or moderate correlation with the test scores. The strongest correlation in sixth grade was between the test scores and SES--.415 for reading and .426 for mathematics--followed closely by the positive correlations between the tests and Gifted. Other moderate correlations with the tests include negative correlations with both Black and ECE and a positive correlation with Two Parents.

Table 8 contains the correlations for background variables and the seventh grade KCCT Reading test scores. All, including Female, show low or moderate correlations. Gifted, SES, and ECE had the strongest correlations. The low correlations between Black and KCCT Reading (-.196) and between Two Parents and KCCT Reading (.198) were much less than the moderate correlations in Table 7 between Black and CTBS Reading (-.320) and between Black and Two Parents (.265). The correlation between SES and KCCT Reading, although moderate, was also somewhat less than for CTBS Reading in Table 7.



Table 6

*Correlations for Independent Variables and Fifth Grade Student KCCT Mathematics Scores*

Variable	KCCT							
	Math	Female	Black	ECE	Gifted	SES	Mobil	Two Parents
KCCT Math	--	.017	-.235**	-.224**	.345**	.334**	-.120**	.183**
Female		--	.009	-.124**	.001	-.016	-.035**	-.027**
Black			--	.099**	-.167**	-.404**	.113**	-.345**
ECE				--	-.110**	-.140**	.086**	-.106**
Gifted					--	.250**	-.064**	.193**
SES						--	-.157**	.403**
Mobil							--	-.123**
Two Parents								--

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

Table 7

*Correlations for Independent Variables and Sixth Grade Student CTBS Reading and Mathematics Tests Scores*

Variable	CTBS Read	CTBS Math	Female	Black	ECE	Gifted	SES	Mobil	Two Parents
CTBS Read	--		.113**	-.320**	-.317**	.368**	.415**	-.164**	.265**
CTBS Math		--	.042**	-.345**	-.323**	.400**	.426**	-.179**	.284**
Female			--	-.014	-.126*	-.002	.016	-.046**	-.019
Black				--	.094**	-.158**	-.428**	.137**	-.335**
ECE					--	-.097**	-.140**	.071**	-.090**
Gifted						--	.217**	-.066**	.174**
SES							--	-.180**	.402**
Mobil								--	-.135**
Two Parents									--

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

Table 8

*Correlations for Independent Variables and Seventh Grade Student KCCT Reading Scores*

Variable	KCCT							
	Read	Female	Black	ECE	Gifted	SES	Mobil	Two Parents
KCCT Read	--	.183**	-.196**	-.302**	.356**	.360**	-.153**	.198**
Female		--	-.007	-.151**	.030**	.020	-.038**	.007
Black			--	.097**	-.170**	-.412**	.116**	-.359**
ECE				--	-.112**	-.174**	.111**	-.099**
Gifted					--	.268**	-.074**	.162**
SES						--	-.156**	.387**
Mobil							--	-.137**
Two Parents								--

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

The correlations for the results of the eighth grade KCCT Mathematics test scores and demographic variables in Table 9 are similar to those with KCCT Reading test scores, except that Female had a lower, negligible correlation. The strongest negative correlation was between KCCT Mathematics and ECE. The strongest positive correlations for KCCT Mathematics were with Gifted and SES. The relationship between CTBS Mathematics (Table 7) and KCCT Mathematics was similar to the pattern between the CTBS and KCCT reading tests--generally the KCCT was lower. The one exception was for ECE where the negative correlation was higher for the KCCT Mathematics test.

In all of the middle school grades (Tables 7, 8, and 9) the patterns of the correlations among the demographic variables were generally the same as the pattern of the fifth grade correlations. The strongest were the negative moderate correlations between Black and SES and between Black and Two Parents and the positive correlation between SES and Two Parents. There were positive, low correlations between ECE and Mobil, between SES and Two Parents and between SES and Gifted. There were negative low correlations between ECE and Gifted, between Gifted and Black, between SES and Mobil, and between SES and Two Parents.

*High School Students.* In high school, Tables 10, 11, and 12 show the same pattern of correlations among the demographic variables as that in middle school. Across the three years, however, there was a trend of the correlations among the demographic variables being generally weaker in the higher grades.

Table 9

*Correlations for Independent Variables and Eighth Grade Student KCCT Mathematics Scores*

Variable	KCCT							
	Math	Female	Black	ECE	Gifted	SES	Mobil	Two Parents
KCCT Math	--	.029*	-.228**	-.372**	.376**	.361**	-.148**	.195**
Female		--	.001	-.104**	.009	-.015	-.023*	-.001
Black			--	.097**	-.164**	-.418**	.105**	-.341**
ECE				--	-.125**	-.147**	.118**	-.112**
Gifted					--	.245**	-.072**	.167**
SES						--	.160**	.378**
Mobil							--	-.117**
Two Parents								--

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

Table 10

*Correlations for Independent Variables and Ninth Grade Student CTBS Reading and Mathematics Scores*

Variable	CTBS Read	CTBS Math	Female	Black	ECE	Gifted	SES	Mobil	Two Parents
CTBS Read	--		.135**	-.329**	-.318**	.346**	.357**	-.187**	.254**
CTBS Math		--	-.060**	-.391**	-.298**	.381**	.400**	-.182**	.292**
Female			--	.014	-.112**	.027*	-.019	-.042*	-.013
Black				--	.112**	-.140**	-.420**	.110**	-.350**
ECE					--	-.088**	-.165**	.113**	-.124**
Gifted						--	.197**	-.078**	.146**
SES							--	-.175**	.359**
Mobil								--	-.151**
Two Parents									--

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

Table 11

*Correlations for Independent Variables and Tenth Grade Student KCCT Reading Scores*

Variable	KCCT							
	Read	Female	Black	ECE	Gifted	SES	Mobil	Two Parents
KCCT Read	--	.166**	-.265**	-.305**	.419**	.358**	-.200**	.239**
Female		--	.005	-.103**	.023*	-.012	-.024*	-.035**
Black			--	.115**	-.175**	-.389**	.102**	-.308**
ECE				--	-.102**	-.154**	.136**	-.099**
Gifted					--	.219**	-.086**	.189**
SES						--	-.140**	.330**
Mobil							--	-.122**
Two Parents								--

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

Table 12

*Correlations for Independent Variables and Eleventh Grade Student KCCT Mathematics Scores*

Variable	KCCT							
	Math	Female	Black	ECE	Gifted	SES	Mobil	Two Parents
KCCT Math	--	.016	-.294**	-.350**	.395**	.352**	-.144**	.228**
Female		--	.035**	-.082**	.043**	-.042**	.008	-.033**
Black			--	.120**	-.165**	-.408**	.108**	-.291**
ECE				--	-.096**	-.146**	.029*	-.095**
Gifted					--	.200**	-.083**	.153**
SES						--	-.138**	.311**
Mobil							--	-.117**
Two Parents								--

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



In ninth grade the correlations of demographic variables with the CTBS Reading and Mathematics test scores shown in Table 10 are similar to those in sixth grade. The correlations for demographic variables with the KCCT Reading test scores in tenth grade (Table 11) and the KCCT Mathematics scores in eleventh grade (Table 12) were generally stronger than for middle school students.

#### *Trends of the Independent Variables Across Grade Levels*

Trends are evident among the independent variables and between each independent variable and the test scores. Table 13 presents a summary of the correlations for just the dependent variables with each of the seven socio-demographic variables as derived from Tables 4-12.

*Female.* Female had an insignificant or negligible, positive correlation with the test scores (Table 13) for all years, except that it had a somewhat stronger but low correlation with KCCT Read in grades seven and ten and a very negligible negative correlation with CTBS Mathematics in ninth grade. Overall, the correlations of Female with the various tests, although negligible or low, were higher for reading scores than for mathematics scores for all years. In reading, Female correlated higher with the KCCT tests than the CTBS, but this trend did not hold for mathematics.

Female had an insignificant correlation with Black, except in the eleventh grade when a positive correlation was significant but negligible indicating a slightly higher proportion of females among the Black students. Female had a consistent, negative, negligible-low correlation with ECE ranging from  $-.082$  to  $-.151$ . Female had an insignificant or negligible correlation with Gifted, SES, Mobil, and Two Parents for all grades.

Table 13

*Correlations of Students' Background Variables with CATS Achievement Test Results for all Grades*

Grade	Test	Female	Black	ECE	Gifted	SES	Mobil	Two Parents
3	CTBS Read	.070**	-.267**	-.217**	.181**	.347**	-.140**	.231**
3	CTBS Math	.005	-.300**	-.242**	.185**	.358**	-.131**	.250**
4	KCCT Read	.121**	-.187**	-.164**	.296**	.319**	-.079**	.165**
5	KCCT Math	.017	-.235**	-.224**	.345**	.334**	-.120**	.183**
6	CTBS Read	.113**	-.320**	-.317**	.368**	.415**	-.164**	.265**
6	CTBS Math	.042**	-.345**	-.323**	.400**	.426**	-.179**	.284**
7	KCCT Read	.183**	-.196**	-.302**	.356**	.360**	-.153**	.198**
8	KCCT Math	.029*	-.228**	-.372**	.376**	.361**	-.148**	.195**
9	CTBS Read	.135**	-.329**	-.318**	.346**	.357**	-.187**	.254**
9	CTBS Math	-.060**	-.391**	-.298**	.381**	.400**	-.182**	.292**
10	KCCT Read	.166**	-.265**	-.305**	.419**	.358**	-.200**	.239**
11	KCCT Math	.016	-.294**	-.350**	.395**	.352**	-.144**	.228**

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

*Black.* Black had a moderate, negative correlation with the test scores (Table 13), ranging from -.187 for fourth grade KCCT Read to -.391 for ninth grade CTBS Math. The correlations were consistently stronger for Black with the CTBS scores than with the KCCT scores and stronger for mathematics scores than for reading scores.

Black had a negligible, positive correlation with ECE and Mobil. The correlation with ECE increases from third to eleventh grades. Black had a low-moderate, negative correlation with gifted that was consistent from fourth grade to eleventh. Black had a high moderate, negative correlation with SES and a moderate, negative correlation with Two Parents for all grades.

*ECE.* ECE had a moderate, negative correlation with the test scores, which was somewhat greater in middle and high school than in elementary (Table 13). With the exception of the CTBS ninth grade tests, the relationship was somewhat stronger for mathematics than for reading. Among the demographic variables, ECE had a negligible, negative correlation with Gifted and Two Parent, a negligible positive correlation with Mobil and a low, negative correlation with SES.

*Gifted.* Gifted had a moderate, positive correlation with test scores (Table 13) that was stronger in the higher grades (e.g.,  $r = .419$  for KCCT Read at tenth grade). Gifted had a low-moderate positive correlation with SES and a weaker, low positive correlation with Two Parent. The correlation with Mobil was negative and negligible.

*SES.* Of the demographic variables, SES generally had the greatest correlations with the test scores. Table 13 shows a consistently moderate, positive correlation, ranging from .319 to .426. The correlations with SES were stronger for the CTBS than

for the KCCT tests with the exception of high school reading tests. Among the CTBS scores, the correlations were stronger for mathematics scores than for reading.

Among the demographic variables, SES had a low, negative correlation with Mobil and a moderate correlation with Two Parent at all grades, in addition to the moderate, negative correlations with Black and positive correlations with Gifted already mentioned.

*Mobil.* Mobil generally had a low, negative correlation with the test scores (Table 13). It had negligible, negative correlations with Two Parents as well as the negligible or low correlations with the other demographic variables already mentioned.

*Two Parents.* Two Parents had a low to moderate correlation with the test scores. It had a stronger relationship with the CTBS test scores than with the KCCT except in high school. Among the demographic variables, the strongest relationships with Two Parents were a moderate, negative correlation with Black and a somewhat higher, moderate, positive correlation with SES.

Overall, among the demographic factors, the strongest relationships were among Two Parents, SES, ECE, and Black with correlations of .3 to .4, which were in the moderate range. These relationships raise questions about the overlap or redundancy of those variables that were explored in the regression analysis. The three trends revealed by the correlations were these. First, except for KCCT Math in the eighth and eleventh grades, SES consistently had the highest correlations with the tests scores, followed by Gifted and ECE on the KCCT scores and by Gifted, ECE, and Black on the CTBS scores. Second, CTBS tests generally had higher correlations with the demographic factors than the KCCT. Third, the mathematics tests generally had higher correlations

with the demographic factors than the reading tests.

### *School Level Parameters*

The population parameters for the school composition variables are found in Appendix D. School composition was described in terms of percentages of students in the school who were characterized by the background factors under study, i.e., the percentages of females, Blacks, ECE students, gifted students, students not eligible for free and reduced meals (higher SES), and two parents. Mobility is a rate (devised by JCPS) expressed as a percentage computed by dividing the number of reentries (students who enter a school during the year) in a school by the number of students in the school. These percentages represented the entire population of the schools, grades K-12, although only grades 3-12 were tested.

During analysis, the percentage of gifted students in the seventh grade at one of the middle schools was found to be incorrect. As mentioned previously, the principal of that school reported that the actual percentage was not 3% as recorded at KDE, but closer to 25% (personal communication, May 16, 2003). The principal's estimate agreed with the percentage of gifted students in the eighth and seventh grades in previous years. Therefore, the eighth grade percentage was used as the most unbiased estimate of that school's seventh grade gifted percentage that was available. Because there were only 24 middle schools and because the percent of gifted students identified at that school was among the highest in the school district, omitting the school or the seventh grade from the analysis or substituting with a district average percentage would have been more likely to bias the results.

In addition to the means and standard deviations of the school composition

percentages, the lowest and highest percentages and the range are reported in order to describe better the differences in student body composition among the schools.

Table D1 reports the parameters for the elementary schools. The greatest differences among the elementary schools were found in their percentages of gifted, higher SES (%HiSES), and two-parent families (%Two Parents). %HiSES had the greatest range (84) and a large standard deviation (21). A high degree of skewness was evident in the %Gifted, which had a mean of 8.1% and a median of 2.1% with a standard deviation of 12. The mode for gifted was 0.0% because there were 34 elementary schools that had no identified gifted and talented students. Although the range for %Black was 34, the mean and median were both 37% and the standard deviation was 9.7 indicating that the Black percentage of most schools was similar.

Middle school percentages are shown in Table D2. For %Black, the range and standard deviation were slightly less, indicating more equal dispersion of Black students among middle schools than in elementary schools. The range and standard deviation of %HiSES were slightly smaller, but the mean was considerably higher. The skewed pattern of %Gifted was similar to the elementary with a mean of 11%, a median of 3%, and 10 middle schools having 0.0% Gifted. The Mobility Rate mean was 20% higher in middle school than in elementary.

Table D3 indicates high schools had more variability than middle and elementary schools in proportions of females and Blacks. The range of %Females was 25, and the standard deviation was 5.6. The range of %Black was 52 and the Standard Deviation was 13.2. The median was considerably lower than the mean indicating a skewed distribution. In contrast, the ranges of %HiSES, %Mobil, and %Two Parents

were narrower than in middle school. The range of %Gifted was 46, which was the same as in elementary school but wider than in middle school. The dispersion of the %Gifted in high school was similar to that in elementary with 9 schools having 0.0% gifted, a mean of 10% and a median of 5% with the highest percentage being 45%.

### *School-Level CATS Test Results*

The range, mean, and standard deviation are given for the school-level scores in Appendix E. The tests included the CTBS Reading and Mathematics tests, the KCCT Reading and Mathematics tests, the NRT Index, and the Academic Index. (The NRT and Academic Indices are calculated from the CTBS and KCCT test results respectively and are used for accountability. See Chapter III for a complete explanation.)

The means were slightly different from those reported for the students in this study. Since the sizes of the schools differ, the grand mean of the school mean scores differs from the district-wide aggregated student means.

As with the student-level scores, the KCCT scores were higher for elementary schools (Table E1) than for middle and high schools (Tables E2 and E3). For the CTBS, the scores were near 50 (expressed in Normal Curve Equivalents) for elementary and Grade 9 reading but somewhat lower for middle school and Grade 9 mathematics. The same trend was seen in the NRT Indices. The mean elementary NRT Index was 76.1. At middle school, it was 65.9, and at high school, it was 66.6. For the mean Academic Index, however, the scores were more similar at each level--60.6 for elementary schools, 58.2 for middle schools, and 61.8 for high schools.

The standard deviations and variances of the CTBS scores were somewhat higher at the higher grades, but for the KCCT scores the standard deviations and

variances were considerably larger at the high school level. They also increase considerably for the NRT Index and the Academic Index. It is important to note that the variances at the school level were much less extreme than at the student level. For example, the variance for the high school KCCT Mathematics test was 902.9, whereas the variance at the student level for KCCT Mathematics was 4264.7. The standard deviations, of course, follow the same pattern.

### *Correlation Matrices*

At the school level, the values of the independent variables applied to the entire school although the various CATS tests were taken at different grades. In this section, two tables are presented for each school level: elementary, middle, and high school. The first table reports the correlations among the independent variables at that school level. The second table presents the correlations between the independent (school composition) variables and the six dependent variables: the scores of the four CATS tests and two performance indices.

In the school-level analyses, the number of subjects (unit of analysis) was much less than with the student level analysis in which the  $N$  was close to 7000 for each calculation. At the school level,  $N = 87$  for elementary schools,  $N = 24$  for middle schools, and  $N = 21$  for high schools. As a consequence, only relatively strong coefficients were found to be significant in contrast to the correlations at the student level, which were significant at very low values. But the rationale for interpreting the strength of the correlations remains the same at the school level as at the student level, which is that the relationship of the correlations to school accountability scores can have important implications even for small differences in scores. Therefore, the same



guidelines were used to interpret the strength of the correlations. Correlations below .13 have *negligible* correlations; between .13 and .25, *low* correlations; between .25 and .50, *moderate*; .50 to .75, *substantial*; and above .75, *high*.

For every variable at all levels, except gifted at the middle school level, the correlations among the demographic variables were much stronger at the school level than at the individual level. The correlations of the CTBS and KCCT test scores with the demographic variables were also much stronger at the school level, again except for the correlation of the scores with %Gifted for middle schools.

The correlations for the elementary school composition variables are reported in Table 14. The strongest relationship was a very high correlation between %HiSES and %Two Parents, .930. There was a substantial negative correlation of %Blacks with %HiSES, -.643; and with %Two Parents, -.686. The strongest correlations for %Mobil were substantial, negative correlations with %HiSES, -.608, and %Two Parents, -.614. %Gifted had a moderate positive correlation with %HiSES and with %Two Parents. The correlation of %ECE with %Mobil was positive and moderate. The correlations of %ECE with %HiSES and with %Two Parents were moderate and negative. The %Females was not significantly correlated with any of the other composition variables, but had low, positive correlations with %HiSES and %Two Parents and a low, negative correlation with %Mobil.

Table 14

*Correlation Matrix for Elementary School Independent Variables*

Variable	%Females	%Blacks	%ECE	%Gifted	%HiSES	Mobil Rate	%Two Parents
%Females	--	-.017	-.017	.091	.170	-.156	.169
%Blacks		--	.162	-.221*	-.643**	.254**	-.696**
%ECE			--	-.232*	-.378**	.348**	-.373**
%Gifted				--	.479**	-.294**	.497**
%HiSES					--	-.608**	.930**
Mobil Rate						--	-.614**
%Two Parents							--

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

The correlations between the composition variables and the CATS test and performance indices scores (dependent variables) for the elementary schools are shown in Table 15. Every composition variable except for %Female had a moderate, substantial, or high correlation with the CATS scores. The strongest was a high positive correlation between all of the scores and %HiSES. The greatest correlation was between %HiSES and the Academic Index at .848. The correlations of the scores with %Two Parents was similar, and only slightly less strong overall than with %HiSES. The scores and %Mobil have a substantial, negative correlation. %Gifted had a moderate, positive correlation and %Black a moderate, negative correlation with the scores.

At middle school, the correlations among the middle school composition variables shown in Table 16 are generally stronger than those in elementary school. The strongest was again the high positive correlation between %HiSES and %Two Parents. Also, %Mobil and %ECE had a high positive correlation. Correlations between %Mobil and %Two Parents, between %Two Parents and %ECE, %HiSES and %ECE, and between %Mobil and %HiSES were all high and negative. There were substantial positive correlations of %Blacks with %ECE and with %Mobil, and substantial negative correlations of Blacks with %HiSES and with %Two parents. %Females shows a moderate positive correlation with %HiSES and %Two Parents and a moderate negative correlation with %Mobil and %ECE. The exception to the stronger correlations was %Gifted, which unlike the patterns at elementary and high school, had a negligible or low correlation with the other school composition variables.

Table 15

*Correlations Between Independent Variables and Elementary School CATS Achievement Tests and Performance Indices Scores*

Variable	%Females	%Blacks	%ECE	%Gifted	%HiSES	Mobil Rate	%Two Parents
SCTBS Read	.108	-.402**	-.341**	.463**	.754**	-.637**	.708**
SCTBS Math	.089	-.423**	-.362**	.468**	.723**	-.606**	.702**
SKCCT Read	.132	-.440**	-.271**	.479**	.711**	-.528**	.737**
SKCCT Math	.142	-.395**	-.381**	.423**	.729**	-.622**	.721**
NRT Index	.100	-.442**	-.355**	.450**	.749**	-.603**	.720**
Acad Index	.204	-.513**	-.384**	.455**	.848**	-.636**	.827**

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

Table 16

*Correlation Matrix for Middle School Independent Variables*

Variable	%Females	%Blacks	%ECE	%Gifted	%HiSES	Mobil Rate	%Two Parents
%Females	--	.022	-.350*	.031	.358*	-.291	.357*
%Blacks		--	.630**	.025	-.614**	.626**	-.627**
%ECE			--	-.026	-.851**	.814**	-.841**
%Gifted				--	.223	-.191	.214
%HiSES					--	-.824**	.943**
Mobil Rate						--	-.801**
%Two Parents							--

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

Table 17 shows the correlations of the middle school composition variables with the CATS tests and performance indices scores. %ECE and %Mobil had high negative correlations with all of the scores. %HiSES and %Two Parents had high positive correlations with all of the scores. Those correlations were all much stronger than the comparable correlations in elementary school. %Blacks had low-substantial, negative correlations with the scores; and %Females had low-moderate positive correlations with the scores, both of which were also stronger than in elementary school. In contrast, the correlations between %Gifted and the scores were weaker than in elementary school and weaker than most of the student-level correlations of Gifted with test score.

Table 18 shows the correlations among the school composition variables at high school. As at elementary and middle school, the strongest correlations were the very high positive ones between %HiSES and %Two Parents (.964) and between %Mobil and %ECE (.936). The correlations among high school variables were generally similar to those among middle schools. However, differences include much stronger correlations between gifted and the other school composition variables; a stronger, substantial, negative correlation for %Females with %Mobil and %ECE; and a much weaker positive correlation for %Black with %ECE and %Mobil.

Table 17

*Correlations Between Independent Variables and Middle School CATS Achievement Tests and Performance Indices Scores*

Variable	%Females	%Blacks	%ECE	%Gifted	%HiSES	Mobil Rate	%Two Parents
SCTBS Read	.238	-.499**	-.811**	.308	.906**	-.853**	.876**
SCTBS Math	.267	-.490**	-.820**	.328	.904**	-.851**	.879**
SKCCT Read	.253	-.523**	-.825**	.307	.847**	-.895**	.860**
SKCCT Math	.325	-.492**	-.750**	.387*	.881**	-.837**	.857**
NRT Index	.269	-.509**	-.833**	.268	.914**	-.863**	.890**
Acad Index	.310	-.468**	-.786**	.339	.902**	-.862**	.887**

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

Table 18

*Correlation Matrix for High School Independent Variables*

Variable	%Females	%Blacks	%ECE	%Gifted	%HiSES	Mobil Rate	%Two Parents
%Females	--	.081	-.605**	.086	.283	-.564**	.327
%Blacks		--	.364	-.384*	-.632**	.333	-.679**
%ECE			--	-.624**	-.789**	.936**	-.802**
%Gifted				--	.670**	-.509**	.638**
%HiSES					--	-.770**	.964**
Mobil Rate %Two Parents						--	-.777**
							--

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



The correlations of the school composition variables with the CATS achievement tests and performance indices scores for high schools are shown in Table 19. %HiSES and %Two Parents had very high positive correlations ranging between .887 and .936 for all of the test scores. %ECE and %Mobil had high negative correlations with the test scores ranging from -.801 to -.846. %Gifted had a substantial, positive correlation with the scores, and %Females had a moderate, positive correlation with the scores. For %Blacks, the negative correlations with the scores at high school were high-moderate, similar to those at middle and elementary school.

Overall, there were some fluctuations in the correlations among the school composition variables with those in high school and middle school generally being slightly stronger than those in elementary (Tables 14, 16, and 18). The highest and most consistent correlation was the very high positive correlation between %HiSES and %Two Parents, .930 to .964 and with Mobil Rate and %Two Parents (-.617 to -.801). The negative correlation for %ECE and Mobil Rate was only moderate in elementary school, but was higher in middle school and very high at .936 at high school. The negative correlation of %ECE with %HiSES and with %Two Parents was moderate at elementary school and high at middle and high school. The negative correlations of %Females with %ECE and with %Mobil increase substantially from elementary to middle to high school. %Blacks had a consistently strong, negative correlation with %HiSES and %Two Parents, but the correlations of %Blacks with %ECE, %Gifted, and %Mobil fluctuate.

Table 19

*Correlations Between Independent Variables and High School CATS Achievement Tests and Performance Indices Scores*

Variable	%Females	%Blacks	%ECE	%Gifted	%HiSES	Mobil Rate	%Two Parents
SCTBS Read	.353	-.419*	-.826**	.660**	.887**	-.823**	.897**
SCTBS Math	.315	-.554**	-.820**	.700**	.915**	-.801**	.936**
SKCCT Read	.389*	-.434*	-.840**	.670**	.891**	-.817**	.891**
SKCCT Math	.460*	-.514**	-.856**	.609**	.929**	-.847**	.932**
NRT Index	.375*	-.480*	-.836**	.646**	.896**	-.821**	.915**
Acad Index	.402*	-.461*	-.829**	.640**	.907**	-.827**	.919**

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

The most dramatic aspect of the correlations was the strong relationships between the school composition variables and the school scores for every grade. Most of the correlations of the school composition variables with the CATS tests scores were markedly stronger in middle and high school than at elementary (Tables 15, 17, and 19). The only variable that had similar correlations with the CATS scores across all three levels was %Blacks, which was moderate and negative. The variable %Females had low, positive correlations with the scores in elementary but moderate correlations in middle school and slightly stronger in high school. %ECE had moderate negative correlations with the scores in elementary school but high negative correlations in high school. %Gifted was moderately correlated with the test scores in elementary school, a little lower in middle school, but substantially correlated with the scores in high school. The high positive correlations of %HiSES and of %Two Parents were almost identical, between .708 and .848 in elementary school and between .847 and .936 in middle and high school. %Mobil follows the same pattern although not quite as strong. %Mobil had substantial negative correlations with the scores in elementary school and high negative correlations in middle and high school.

The correlations of the school composition variables with the school-level CATS tests and performance scores were far stronger than the corresponding correlations at the student level as seen by comparing the figures in Table 13 for the student level scores with Tables 15, 17 and 19 at the school level. For example, the range of the correlations for the student level SES was .319 to .426, whereas the range for %HiSES at the school level was .711 to .929. Except for Blacks and %Blacks, the other variables demonstrated the same pattern. The correlations with the scores and

%Blacks at school level were stronger than the correlations for scores and Black at student level but only slightly so. Both were generally in the moderate negative range.

### Multiple Regression Analyses

Forward linear regression analyses were run to determine to what extent individual student background factors affect individual student achievement and to what extent demographic school composition factors affect school-level CATS achievement scores and performance index scores. Because the purpose was to compare the effects of various variables in a population across nine grade levels, statistical analysis of probability was used to indicate the strength of the relationships but not to eliminate variables from the regression analysis. In order to produce regression models that would include all of the variables, the default entry criterion for  $F$  to enter and  $F$  to remove variables was overridden. The probability of  $F$  to enter was set for .98 and the probability of  $F$  to remove was set at .99 for all regression analyses. In spite of this effort, the probabilities for some of the variables in the school-level analyses were so high that the SPSS program excluded them, and only six models were produced.

With the default entry criterion overridden, each analysis resulted in six or seven regression models, but the purpose of the study was to identify those variables that have an important influence on student achievement scores from among those variables. Therefore, a rationale was necessary that would enable building a best-fit model--a regression equation that resulted in an inclusive, yet practical, equation that retained important variables and excluded those with trivial effects. J. Cohen et al. (2003, cited in Chapter III) recommend a criterion of a .02 increment in the  $R^2$  for small effect size. Therefore, following their recommendation; retaining variables that contribute .02 more

than the previously entered variable (in a simpler model) constituted the decision rule to determine the optimum model for each regression for this study.

The tables are arranged to compare the regression equation models. For each model the unstandardized regression coefficient  $B$  and the standardized regression coefficient  $\beta$  are reported. The unstandardized regression ( $B$ ) provides a direct indication of how many points on a test could be attributable to a particular variable. The standardized coefficient ( $\beta$ ) provides a comparison of how influential a particular variable was in terms of standard deviation units within each regression. The  $R^2$  for each regression model, which indicates the composite effect of all of the variables in a model, is included below each respective model, along with the  $R^2$  change as each variable is entered into the equation.

In the student-level regressions, most of the regression coefficients were significant at  $p < .001$  because of the large number (almost 7000) of subjects tested at each grade. At the school level, there were only 87 elementary schools, 24 middle schools, and 21 high schools. Since the number of cases ( $N$ ) was small, only the first two or three regression equations and coefficients were significant at  $p < .10$ ,  $p < .01$ , or  $p < .001$ . Below, each regression is described table by table, and comparative trends are summarized at the end.

### *Student Level Regression Results*

Models for the third grade CTBS Reading Scores regressed on the student background variables are given in Table 20. Model 4 is the optimum equation because it had an  $R^2$  of .188, which is .017 or .02 (rounded) greater than the  $R^2$  of Model 3. The  $R^2$  of .188 indicates that the optimum regression equation, Model 4, contributed almost

20% of the variability in the test results. The variables added in Models 5, 6 and 7 contributed negligibly to the  $R^2$ . Using the unstandardized  $B$  coefficients from Model 4 produces the following regression equation:

$$\text{Predicted CTBS Reading Score } (\hat{y}) = 42.99 (\text{constant}) + 5.70(\text{SES}) - 12.64(\text{ECE}) - 6.41(\text{Black}) + 16.16(\text{Gifted})$$

The interpretation of this equation would be as follows. The  $B$  coefficients for each demographic variable multiplied by the value of the variable are then added to the constant to produce the predicted score for a student. For each unit increase in SES (which is on a rank order scale of 1-3) while holding other predictors constant, a student's third grade CTBS reading score would increase by 5.70 (on average). ECE is a dichotomous variable with a child identified as ECE = 1 and a child not identified as ECE = 0. Therefore, ECE students can be expected on average to score 12.64 points lower than non-ECE students on the test holding the other predictors constant. Black students (also coded as 1 on a dichotomous variable) would be expected to average 6.41 points lower than other students holding other predictors constant, and Gifted (gifted and talented) students would be expected to average 16.16 points higher than other students holding other predictors constant.

Table 20

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Third Grade CTBS Reading Student**Scores on Student Background Variables (N = 6948)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	35.80		37.97		42.97		42.99		43.61		43.05		41.90	
SES	7.80	(.35)	7.36	(.33)	6.02	(.27)	5.70	(.25)	5.51	(.25)	5.06	(.23)	5.10	(.23)
ECE			-13.04	(-.18)	-12.87	(-.18)	-12.64	(-.17)	-12.28	(-.17)	-12.27	(-.17)	-11.89	(-.17)
Black					-6.62	(-.15)	-6.41	(-.15)	-6.27	(-.14)	-5.60	(-.13)	-5.60	(-.13)
Gifted							16.16	(.13)	16.07	(.13)	15.99	(.13)	15.81	(.13)
Mobil									-4.34	(-.07)	-4.13	(-.06)	-4.14	(-.06)
Two Parents											2.73	(.06)	2.72	(.06)
Female													2.23	(.05)
$R^2$	.121		.153		.172		.188		.193		.196		.199	
$R^2$ Change	.121		.033		.019		.017		.004		.003		.003	

 $p < .001$  for all values.

The standardized coefficient,  $\beta$ , indicates the relative contribution of each variable to the prediction of the dependent variables, the test scores. The variable with the greatest influence was SES, which had a  $\beta$  of .25. The ECE  $\beta$  was -.17 and Black was -.15. Gifted had the least influence of Model 4 with a  $\beta$  of .13. The proportions of these effects seem different from the interpretation of the regression equation as explained above because the  $\beta$  coefficients represent standard deviations with a common z score unit. A change of one standard deviation in the independent variable results in a change in the dependent variable that is  $\beta$  times the standard deviation of the test score (dependent variable). Since the standard deviations of the independent variables differ, the standardized effects are influenced by the size of the standard deviation for each respective predictor. As with the  $B$  coefficients, the relative size of the  $\beta$  coefficients is affected by which predictors are included in an equation.

Table 21 shows that the unstandardized regression coefficients ( $B$ ) for third grade CTBS Mathematics scores on student background variables follows the same general pattern as the reading test. Again, Model 4 represents an  $R^2$  change of .017. However, the Model 4  $R^2$  of .214 was slightly higher than that for the CTBS Reading test. The influence of SES with a  $\beta$  of .25 was similar to the influence of SES on the reading test, but the negative coefficients for ECE and Black were greater and apparently responsible for most of the increase in  $R^2$ .



Table 21

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Third Grade CTBS Mathematics Student Scores on Student Background Variables (N = 6947)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	36.05		38.43		44.48		44.51		43.86		44.34		44.69	
SES	7.88	(.34)	7.38	(.34)	5.77	(.26)	5.46	(.25)	4.91	(.22)	4.79	(.22)	4.78	(.22)
ECE			-14.46	(-.21)	-14.27	(-.20)	-14.04	(-.20)	-14.01	(-.20)	-13.74	(-.20)	-13.86	(-.20)
Black					-7.97	(-.18)	-7.76	(-.18)	-6.95	(-.15)	-6.88	(-.16)	-6.88	(-.16)
Gifted							15.96	(.13)	15.85	(.13)	15.79	(.13)	15.85	(.13)
Two Parents									3.31	(.08)	3.17	(.08)	3.18	(.08)
Mobil											-3.15	(-.05)	-3.14	(-.05)
Female														-.68 <sup>a</sup> (-.02)
$R^2$	.128		.169		.198		.214		.219		.221		.221	
$R^2$ Change	.128		.042		.028		.017		.005		.002		.000 <sup>a</sup>	

$p < .001$  unless otherwise noted. <sup>a</sup> $p = .131$ .

For the fourth grade KCCT Reading regression (Table 22) Model 2, the optimum regression equation included only variables SES and Gifted with an  $R^2$  of .153. Thus, background variables accounted for about 15% of the variance on the KCCT Reading tests and were less influential here than at any other grade.

Although not meeting the .02  $R^2$  increase criterion, the addition of Female to the equation in Model 3 was interesting. Female had consistently the weakest correlations with the test scores, when SES and Gifted were controlled, but it seems to account for some variability in this fourth grade test, a .014 increase in  $R^2$ . Female also had the lowest correlation with the other background variables, so its entry into the equation had little effect on the coefficients of the other variables.

The regression equations for fifth grade KCCT Mathematics scores and background variables are in Table 23. In this case, Gifted was the most influential variable, followed by SES and ECE with an  $R^2 = .209$ . As with the third grade CTBS test, the background variables accounted for a somewhat larger proportion of the variance on the mathematics test than on the reading test. Here Female was statistically insignificant. Black was again in the top four with a coefficient of -10.3 but not included in the optimum equation. For the KCCT Mathematics test, Model 3 with an  $R^2$  increase of .025, which includes only Gifted, SES and ECE, provides the optimum regression equation.

Table 22

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Fourth Grade KCCT Reading Student**Scores on Student Background Variables (N = 7143)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	507.9		510.1		504.7		508.1		512.0		512.4		512.2	
SES	16.0	(.32)	13.2	(.26)	13.2	(.26)	12.7	(.25)	11.7	(.23)	11.6	(.23)	11.4	(.23)
Gifted			41.1	(.23)	40.8	(.23)	39.7	(.23)	39.2	(.22)	39.1	(.22)	39.0	(.22)
Female					11.1	(.12)	9.9	(.10)	9.9	(.10)	9.9	(.10)	9.9	(.10)
ECE							-14.6	(-.10)	-14.4	(-.09)	-14.2	(-.09)	-14.2	(-.09)
Black									-5.3	(-.05)	-5.2	(-.05)	-5.1	(-.05)
Mobil											-2.6 <sup>a</sup>	(-.02)	-2.5 <sup>a</sup>	(-.02)
Two Parents													0.7 <sup>b</sup>	(.01)
R <sup>2</sup>	.102		.153		.167		.176		.178		.178		.178	
R <sup>2</sup> Change	.102		.052		.014		.009		.002		.000		.000 <sup>b</sup>	

$p < .001$  unless otherwise noted. <sup>a</sup> $p < .10$ . <sup>b</sup> $p = .547$ .

Table 23

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Fifth Grade KCCT Mathematics Student Scores on Student Background Variables (N = 7357)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	548.3		519.6		525.2		532.9		534.0		534.1		534.1	
Gifted	63.7	(.35)	51.5	(.28)	49.1	(.27)	48.0	(.26)	47.8	(.26)	46.8	(.26)	47.8	(.26)
SES			15.0	(.26)	13.9	(.24)	12.0	(.21)	11.6	(.20)	11.7	(.20)	11.7	(.20)
ECE					-27.7	(-.16)	-27.1	(-.16)	-26.6	(-.15)	-26.6	(-.15)	-26.6	(-.15)
Black							-10.3	(-.09)	-10.0	(-.09)	-10.0	(-.09)	-10.0	(-.08)
Mobil									-8.9	(-.05)	-8.9	(-.05)	-8.9	(-.05)
Two Parents											-.22 <sup>a</sup>	(-.00)	-.22 <sup>a</sup>	(.00)
Female													0.1 <sup>a</sup>	(.00)
$R^2$	.119		.184		.209		.216		.219		.219		.219	
$R^2$ Change	.119		.065		.025		.007		.002		.000 <sup>a</sup>		.000 <sup>a</sup>	

$p < .001$  unless otherwise noted. <sup>a</sup> $p > .800$

In the sixth grade the CTBS Reading (Table 24) and Mathematics (Table 25) regressions were similar. As in third grade, the background variables had a greater influence on the mathematics scores than on the reading. The same four variables--SES, ECE, Black, and Gifted--are included in the Model 4, which is the optimum regression equation for both reading and mathematics. In Model 4, Gifted had the strongest coefficient, followed by SES and then ECE. The  $R^2$  for CTBS Reading was .328 and for CTBS Mathematics .363, indicating that background variables had a greater influence on the scores at the sixth grade than they did on the third grade CTBS tests. The increase is attributable primarily to the increase in the influence of Gifted on the scores, but it must be remembered that the percentage of gifted students was lower at the third grade, before students have been identified, than in any subsequent year. ECE also had a greater negative influence on the scores in sixth grade than in the third grade, which also contributed to the greater  $R^2$ .

At seventh grade, the optimum equation for the seventh grade KCCT Reading scores with an  $R^2$  of .272 includes four variables in Model 4. The regression coefficients in Table 26 show that SES ( $\beta = .25$ ) and Gifted ( $\beta = .26$ ) continued to contribute in similar proportions as in the optimum models at lower grades. The influence of ECE had increased to  $\beta = -.21$  (similar to those in sixth grade), and Female was the fourth variable with a  $\beta$  of .14. Black was statistically insignificant and the weakest coefficient of the excluded variables.

Table 24

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Sixth Grade CTBS Reading Student Scores on Student Background Variables (N =6917)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	29.79		30.67		33.72		39.37		36.75		37.43		36.97	
SES	8.68	(.42)	7.35	(.35)	6.72	(.32)	5.50	(.26)	5.50	(.26)	5.32	(.25)	4.99	(.22)
Gifted			20.56	(.29)	19.36	(.27)	18.69	(.27)	18.78	(.27)	18.68	(.27)	18.40	(.20)
ECE					-16.28	(-.25)	-15.99	(-.24)	-15.33	(-.23)	-15.18	(-.23)	-15.09	(-.16)
Black							-5.91	(-.14)	-5.90	(-.14)	-5.72	(-.14)	-5.30	(-.13)
Female									3.13	(.08)	3.03	(.08)	3.10	(.08)
Mobil											-3.29	(-.06)	-3.14	(-.05)
Two Parents													2.17	(.05)
$R^2$	.172		.253		.312		.328		.334		.338		.340	
$R^2$ Change	.172		.081		.059		.016		.006		.004		.002	

$p < .001$  for all values.

Table 25

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Sixth Grade CTBS Mathematics Student Scores on Student Background Variables (N =6913)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	26.83		27.94		31.37		37.38		38.14		37.61		37.48	
SES	9.90	(.43)	8.26	(.36)	7.53	(.32)	5.95	(.26)	5.74	(.25)	5.32	(.23)	5.32	(.23)
Gifted			25.33	(.32)	23.99	(.31)	23.12	(.29)	23.01	(.29)	22.65	(.29)	22.66	(.29)
ECE					-18.30	(-.25)	-17.93	(-.24)	-17.72	(-.24)	-17.63	(-.24)	-17.58	(-.24)
Black							-7.64	(-.17)	-7.42	(-.16)	-6.89	(-.15)	-6.89	(-.15)
Mobil									-3.91	(-.07)	-3.74	(-.06)	-3.72	(-.06)
Two Parents											2.73	(.06)	2.74	(.06)
Female													.24 <sup>a</sup>	(.01)
R <sup>2</sup>	.181		.280		.340		.363		.367		.370		.370	
R <sup>2</sup> Change	.181		.099		.060		.022		.004		.003		.000 <sup>a</sup>	

$p < .001$  unless otherwise noted. <sup>a</sup> $p = .577$

Table 26

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Seventh Grade KCCT Reading Student**Scores on Student Background Variables (N =6389)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )
Constant	471.4		474.3		481.1		474.8		476.3		476.2		477.7	
SES	16.1	(.36)	12.7	(.29)	11.2	(.25)	11.2	(.25)	10.8	(.24)	10.3	(.23)	10.0	(.22)
Gifted			38.5	(.28)	36.3	(.26)	36.0	(.26)	35.7	(.26)	35.3	(.26)	35.2	(.26)
ECE					-32.5	(-.23)	-29.5	(-.21)	-28.8	(-.20)	-28.7	(-.20)	-28.6	(-.20)
Female							11.8	(.14)	11.7	(.14)	11.7	(.14)	11.78	(.14)
Mobil									-6.6	(-.07)	-6.4	(-.07)	-6.3	(-.07)
Two Parents											2.8 <sup>a</sup>	(.03)	2.3 <sup>b</sup>	(.03)
Black													-2.1 <sup>b</sup>	(-.02)
$R^2$	.130		.202		.253		.272		.276		.277		.277	
$R^2$ Change	.130		.073		.050		.019		.005		.001		.000 <sup>b</sup>	

$p < .001$  unless otherwise noted. <sup>a</sup> $p < .01$ . <sup>b</sup> $p < .10$ .



The regression equations for the eighth grade KCCT Mathematics test on student background variables are in Table 27. In this case, only three variables produced the most optimum equation, Gifted ( $\beta = .28$ ), ECE ( $\beta = -.30$ ), and SES ( $\beta = .25$ ) with an  $R^2$  of .306. Thus, Gifted, ECE, and SES contributed almost 31% to the variance on KCCT Mathematics scores in comparison to 27% on the reading scores. At middle school, therefore, the background variables contributed more to mathematics scores than to reading and more to the CTBS scores than to KCCT scores.

The ninth grade CTBS Reading and Mathematics test regressions on background variables, presented in Tables 28 and 29, show that the same four variables are included in the optimum equations as have been in all of the CTBS test regressions: SES, Gifted, ECE, and Black. For the reading test (Table 28), the order of the entry of the variables was the same as for the sixth grade CTBS test regressions. For the CTBS Mathematics regression (Table 29), however, Black entered before ECE and had a stronger value. Once again, the influence of these four student background factors was stronger for mathematics ( $R^2 = .349$ ) than for reading ( $R^2 = .296$ ).

In Table 30 for the tenth grade KCCT Reading regression, four variables were included. Gifted entered the equation first and had a higher  $\beta$  coefficient, .34, than it or any other variable had in the other optimum equations. The other variables in Model 4 were SES, ECE and Female with  $R^2 = .322$ . SES, Gifted and ECE were common to most of the optimum equations. In contrast, Female was only included in the optimum equations for the middle and high school KCCT Reading tests. The  $R^2$  (.322) was higher than the  $R^2$  (.296) for the ninth grade CTBS Reading test.

Table 27

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Eighth Grade KCCT Mathematics Student Scores on Student Background Variables (N =6089 )*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	512.2		518.9		489.6		491.1		495.5		495.3		495.4	
Gifted	58.2	(.38)	51.8	(.34)	42.9	(.28)	42.7	(.28)	42.0	(.27)	41.9	(.27)	41.9	(.27)
ECE			-56.8	(-.33)	-51.8	(-.30)	-50.9	(-.30)	-50.7	(-.30)	-50.6	(-.29)	-50.6	(-.29)
SES					13.7	(.25)	13.3	(.24)	12.2	(.22)	12.1	(.22)	12.1	(.22)
Mobil							-6.5	(-.07)	-6.3	(-.06)	-6.2	(-.05)	-6.2	(-.05)
Black									-5.7	(-.05)	-5.4	(-.05)	-5.4	(-.05)
Two Parents											1.2 <sup>a</sup>	(.01)	1.2 <sup>a</sup>	(.01)
Female													-0.2 <sup>b</sup>	(.00)
R <sup>2</sup>	.141		.249		.306		.309		.311		.312		.312	
R <sup>2</sup> Change	.141		.108		.057		.003		.002		.000 <sup>a</sup>		.000 <sup>b</sup>	

$p < .001$  unless otherwise noted. <sup>a</sup> $p = .328$ . <sup>b</sup> $p = .882$

Table 28

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Ninth Grade CTBS Reading Student**Scores on Student Background Variables (N =6014)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )
Constant	35.57		36.44		39.62		45.55		43.26		44.25		43.66	
SES	7.37	(.36)	6.21	(.30)	5.41	(.26)	3.86	(.19)	3.95	(.19)	3.70	(.18)	3.42	(.17)
Gifted			17.94	(.29)	17.03	(.27)	16.36	(.26)	16.19	(.26)	15.98	(.26)	15.77	(.25)
ECE					-19.25	(-.25)	-18.67	(-.24)	-17.69	(-.23)	-17.17	(-.22)	-16.96	(-.22)
Black							-7.65	(-.19)	-7.72	(-.19)	-7.59	(-.19)	-7.02	(-.17)
Female									4.09	(.11)	3.97	(.11)	4.01	(.11)
Mobil											-2.96	(-.09)	-2.81	(-.08)
Two Parents													2.26	(.06)
$R^2$	.128		.206		.268		.296		.308		.315		.318	
$R^2$ Change	.128		.079		.061		.029		.012		.007		.003	

 $p < .001$  for all values.

Table 29

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Ninth Grade CTBS Mathematics Student Scores on Student Background Variables (N = 6000)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	27.27		28.39		37.90		40.64		42.75		43.87		43.06	
SES	9.72	(.40)	8.21	(.34)	5.74	(.24)	5.06	(.21)	4.98	(.21)	4.70	(.19)	4.31	(.18)
Gifted			23.19	(.31)	22.11	(.30)	21.25	(.29)	21.41	(.29)	21.19	(.29)	20.87	(.28)
Black					-12.07	(-.25)	-11.57	(-.24)	-11.50	(-.24)	-11.38	(-.24)	-10.59	(-.22)
ECE							-19.10	(-.21)	-20.00	(-.22)	-19.45	(-.22)	-19.14	(-.21)
Female									-3.77	(-.09)	-3.92	(-.09)	-3.68	(-.09)
Mobil											-3.28	(-.08)	-3.07	(-.08)
Two Parents													3.16	(.07)
$R^2$	.160		.255		.306		.349		.356		.362		.366	
$R^2$ Change	.160		.095		.051		.043		.007		.006		.004	

$p < .001$  for all values.

Table 30

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Tenth Grade KCCT Reading Student Scores on Student Background Variables (N =6316)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	489.7		443.4		453.4		443.6		447.3		457.2		455.4	
Gifted	80.7	(.42)	68.9	(.36)	65.7	(.34)	65.2	(.34)	64.2	(.33)	62.5	(.33)	61.3	(.32)
SES			20.2	(.28)	17.9	(.25)	18.2	(.25)	17.4	(.24)	14.9	(.21)	13.9	(.19)
ECE					-60.1	(-.23)	-56.3	(-.22)	-53.2	(-.21)	-52.0	(-.20)	-51.6	(-.20)
Female							18.0	(.14)	17.8	(.14)	17.9	(.14)	18.2	(.14)
Mobil									-12.5	(-.11)	-12.1	(-.10)	-11.6	(-.10)
Black											-13.5	(-.10)	-11.8	(-.08)
Two Parents													8.3	(.06)
$R^2$	.176		.250		.302		.322		.332		.340		.344	
$R^2$ Change	.176		.074		.053		.019		.011		.008		.003	

$p < .001$  for all values.

Gifted entered first in the regression equation for eleventh grade KCCT Mathematics in Table 31 with a standardized  $\beta$  coefficient of .32. Model 3 is the optimum equation and includes gifted, ECE, and SES entered in that order with an  $R^2 = .311$ . Although not included in the optimum equation, Black was the fourth variable entered with a  $\beta = -.13$  in Model 4. At the high school level, the background factors accounted for a greater proportion of the variance on the KCCT than at middle and elementary school.

Table 31

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Eleventh Grade KCCT Student Mathematics Scores on Student Background Variables (N = 5282)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	518.5		524.3		477.1		492.0		494.4		492.9		493.3	
Gifted	72.6	(.40)	67.0	(.37)	58.5	(.32)	56.6	(.31)	55.9	(.30)	55.1	(.30)	55.1	(.30)
ECE			-88.4	(-.32)	-79.5	(-.28)	-77.5	(-.28)	-77.5	(-.28)	-76.9	(-.27)	-77.0	(-.27)
SES					19.3	(.25)	15.4	(.20)	14.8	(.19)	13.9	(.18)	13.9	(.18)
Black							-19.2	(-.13)	-18.6	(-.13)	-17.0	(-.12)	-17.0	(-.11)
Mobil									-10.7	(-.07)	-10.1	(-.07)	-10.1	(-.07)
Two Parents											7.8	(.06)	7.8	(.06)
Female													-0.7 <sup>a</sup>	(.01)
$R^2$	.156		.254		.311		.325		.330		.333		.333	
$R^2$ Change	.156		.098		.057		.014		.005		.003		.000 <sup>a</sup>	

$p < .001$  unless otherwise noted. <sup>a</sup> $p = .629$ .

### *Trends for the Student Background Regressions*

Some trends are apparent by comparing the regressions across the grade levels, between KCCT and CTBS results, and between reading and mathematics results.

#### *Across Grade Levels*

The effects size for demographic variables on student-level CATS test scores are demonstrated in Table 32 by the  $R^2$  for the optimum models of each regression, which were the last equations that added .02 to the  $R^2$  of the previous model. This resulted in the regressions having different numbers of demographic variables represented in the different optimum equations for different grades. Most of the equations included four variables, but some contained only two or three. These exceptions are noted.

Table 32

#### *Effect Size for Optimum Regression Models of Student Background Variables on Student-Level CATS Tests at Elementary, Middle, and High School*

CATS Tests	$R^2$		
	Elementary	Middle	High
CTBS Reading	.19	.33	.30
CTBS Mathematics	.21	.36	.35
KCCT Reading	.15 <sup>a</sup>	.27	.32
KCCT Mathematics	.21 <sup>b</sup>	.31 <sup>b</sup>	.31 <sup>b</sup>

*Note:* The optimum models contained four background variables unless otherwise noted.

$p < .001$  for all  $R^2$ .

<sup>a</sup>Optimum model included two variables. <sup>b</sup>Optimum model included three variables.



It is apparent from Table 32 that the effects of student background variables were substantially greater at middle and high school than at elementary school. Demographic factors explain close to 20% of the variance on the elementary tests, but they explain from 27% to 36% of the variance at middle and high school. The contribution of each background variable to student scores was indicated by the coefficients produced by the regression of the students' CATS test scores on student background variables.

Table 33

*Unstandardized (B) Coefficients for Optimal Equation Variables for the Regression of Student Background Variables on CATS Test Scores*

CATS Tests	Variables				
	SES	Gifted	ECE	Black	Female
CTBS Reading (Grade 3)	5.7	16.2	-12.6	-6.4	
CTBS Mathematics (Grade 3)	5.5	16.0	-14.0	-7.8	
KCCT Reading (Grade 4)	13.2	41.1			
KCCT Mathematics (Grade 5)	12.0	48.0	-27.1		
CTBS Reading (Grade 6)	5.5	18.7	-16.0	-5.9	
CTBS Mathematics (Grade 6)	6.0	23.1	-17.9	-7.6	
KCCT Reading (Grade 7)	11.2	36.0	-29.5		11.8
KCCT Mathematics (Grade 8)	13.7	42.9	-51.8		
CTBS Reading (Grade 9)	3.9	16.4	-18.7	-7.7	
CTBS Mathematics (Grade 9)	5.1	21.3	-19.1	-11.6	
KCCT Reading (Grade 10)	18.2	65.2	-56.3		18.0
KCCT Mathematics (Grade 11)	19.3	58.5	-79.5		

*Note.* Scores on the CTBS tests range from 1 to 99, on the KCCT tests from 325 to 800.

Table 33 is a listing of the unstandardized (*B*) coefficients for each variable included in the optimum regression equation for each test at each grade. At the student level, the *B* coefficients translate directly into points contributed by the variable to the student scores as noted below.

The Gifted, ECE, Black, and Female variables were dichotomous; thus, the unstandardized *B* coefficient represents the difference in predicted CATS scores (the dependent variable) for the two levels of the dichotomous predictor, when the other variables are controlled. For example, for the dependent variable CTBS Reading Grade 3, the *B* coefficient for Gifted was 16.2. Thus, the predicted CTBS Reading score for gifted students was 16.2 points higher than the predicted CTBS Reading score for students not in the Gifted Program, controlling for other predictor variables.

SES was an ordinal variable with values of 1 (free lunch), 2 (reduced price lunch), and 3 (not subsidized). For the SES variable, the unstandardized *B* coefficient gives the amount of change expected in the dependent variable per unit increase in SES, when the other variables are controlled. However, because it was an ordinal scale, the regression coefficient represents only the average increase per unit. The increase from free lunch to reduced lunch does not represent the same interval as the increase from reduced lunch to non-subsidized lunch. For example, for the dependent variable CTBS Reading 3, the coefficient for SES was 5.7, but one cannot conclude that the predicted CTBS Reading score for students receiving reduced price lunch (SES = 2) was 5.7 points higher than the predicted CTBS Reading score for students on free lunch (SES = 1); likewise this holds for the change from SES of 2 to SES of 3 on this scale. However, because 5.7 was the average unit increase (on a scale of 1-3), one can

estimate that the changes in the CTBS Reading score for students who did not receive the subsidy (SES = 3) versus students who qualify for free lunches, controlling for other predictor variables (SES = 1) was double the 5.7 unstandardized *B* coefficient for a change of 11.4 points.

It is evident in Table 33 that income level (SES), identification as gifted (Gifted), and identification as a special needs student (ECE) had a substantial influence in predicting student scores. They were consistently included in the optimal equations. Black and Female were included in the optimal equations only in some cases.

Standardized coefficients ( $\beta$ ) are useful to compare the strength of the variables within a regression. The standardized  $\beta$  coefficients for the regressions of student background variables on the CTBS and KCCT tests are shown in Table 34.

Across all of the grade levels and tests, SES was consistently included in the optimum equations, having a  $\beta$  coefficient of from .19 to .26 (Table 34). From fourth grade (when gifted students are identified), Gifted was one of the first two variables entered in the equation because of its strong correlation with test scores. Also, from fifth grade through eleventh grade, Gifted consistently had the strongest standardized  $\beta$  coefficient in the optimum equations, ranging from .26 to .34 (Table 34). From fifth grade through eleventh grade, ECE was also included in the optimum equations with a  $\beta$  coefficient from -.16 to -.30 (Table 34).

Mobil and Two Parents were not included in any of the optimum equations and had low impact on the regression equations (Tables 20-31). Mobil had a negative standardized  $\beta$  coefficient that ranges from -.02 to -.11 when the other variables are controlled. Two Parents had a  $\beta$  coefficient of .03 to .07 in those equations for which it

was significant controlling for the other variables.

Table 34

*Standardized ( $\beta$ ) Coefficients of Optimal Equation Variables from the Regression of Student Background Variables on CATS Test Scores*

CATS Test	Independent variables				
	SES	Gifted	ECE	Black	Female
CTBS Reading (Grade 3)	.25	.13	-.17	-.15	
CTBS Mathematics (Grade 3)	.25	.13	-.20	-.18	
KCCT Reading (Grade 4)	.26	.23			
KCCT Mathematics (Grade 5)	.24	.27	-.16		
CTBS Reading (Grade 6)	.26	.27	-.24	-.14	
CTBS Mathematics (Grade 6)	.26	.29	-.24	-.17	
KCCT Reading (Grade 7)	.25	.26	-.21		.14
KCCT Mathematics (Grade 8)	.25	.28	-.30		
CTBS Reading (Grade 9)	.19	.26	-.24	-.19	
CTBS Mathematics (Grade 9)	.21	.29	-.21	-.24	
KCCT Reading (Grade 10)	.25	.34	-.22		.14
KCCT Mathematics (Grade 11)	.25	.32	-.28		

*Differences in the CTBS and KCCT Regressions*

Although SES and ECE have similar effects across the grades on both the CTBS and the KCCT, other variables demonstrate distinctive differences. On the high school KCCT Reading test regression, Gifted has the greatest effect of any of the independent variables with a  $\beta$  coefficient of .34 for reading and .32 for mathematics. At middle

school, however, the  $\beta$  coefficients for Gifted are similar on the CTBS and the KCCT. At elementary they are higher for the KCCT than for the CTBS, but many more students are identified as Gifted in the fourth and fifth grades than in third. In third grade when the CTBS is given, few students have been identified as gifted and talented.

The KCCT and CTBS tests differed dramatically in their relationships with the variables Black and Female (Table 34). Black was not included in the optimum equations for the KCCT scores, contributing less than .02 to the  $R^2$  when it was added to those models. But Black was included in all of the optimum equations for the CTBS tests with strong, standardized  $\beta$  coefficients. The greatest was a  $\beta$  coefficient of -.24 on the eleventh grade CTBS Mathematics test. That means Black students could be expected to score 5.34 points lower (i.e., -.24 times 22.25, the standard deviation of the test) than White students on that test. In contrast, Female was strong enough to be included in the optimum equations only on the KCCT Reading tests at seventh and tenth grades (Table 34).

Except in middle school, the values of  $R^2$  for the effects of background variables on CTBS and KCCT tests appear to be only slightly different (Table 32). The extent of differences between the effects of background on the CTBS and KCCT tests are not obvious in a cursory examination of the  $R^2$  values because of differences in the tests that complicate the analysis: (a) the tests are constructed on different scales, and (b) they have different variances. The variances are reflected in the standard deviations, which are measures of variability (Tables C1-C3). Because  $R^2$  indicates the proportion of the variance that is accounted for by the variables in each optimum regression equation (Tabachnick & Fidell, 1983), the size of the variance and the scale of the test

must be taken into consideration to understand the effects. The standard deviations for the test are displayed in Table 35 for comparison.

The standard deviations are generally consistent on the CTBS Test across the grades, although there are small differences. On reading scores, the standard deviation values are progressively less across the levels (elementary to high school), whereas the values are slightly higher from elementary to high school for mathematics scores. In contrast, the KCCT scores show a large jump in standard deviations at high school representing a much greater variability in the scores on the criterion-referenced test in high school than at the lower grades

Table 35

*Standard Deviations of the Student-Level CTBS and KCCT Tests*

*Comparing Reading and Mathematics Results*

Test	SD		
	Elementary	Middle	High
CTBS Reading	21.41	19.95	18.89
CTBS Mathematics	21.00	22.17	22.25
KCCT Reading	47.79	42.54	64.69
KCCT Mathematics	54.47	52.13	65.31

Except for the high school KCCT tests, both the CTBS and the KCCT tests regressions demonstrated greater effects ( $R^2$ ) of background variables on the mathematics tests than on the reading tests (Table 32). A large part of the differences between reading and mathematics seem to be due to the inclusion or exclusion of Black

and Female from the optimum equations. Female had a small, but different, influence on reading and mathematics. It was the fourth variable in the optimum equations for the seventh and tenth grade KCCT Reading test regressions (Tables 26 and 30) and the third variable, although excluded from the optimum equation, in the fourth grade KCCT Reading test regression (Table 22). The standardized  $\beta$  coefficients for Female fluctuated close to .10 except for the sixth grade CTBS Mathematics test and the eighth and eleventh grade KCCT Mathematics tests when they were insignificant. The regression coefficients for Black were greater for mathematics than for reading but were only large enough to be included in the optimum equations for the CTBS tests (Table 35).

### *Multicollinearity*

The multicollinearity of the background factors at the student level was low, which was demonstrated in two ways. First, as additional variables were entered into the successive regression models, they had little effect on the value of the already entered variables. Secondly, the SPSS 11.5 statistical program produces collinearity statistics including tolerance values (Shannon & Davenport, 2001). Tolerance indicates the proportion of variance for that variable that is independent of the other variables. Tolerance is equal to  $1 - R^2$ , where  $R^2$  is from the regression of the independent variable on all the other independent variables (Allison, 1999). According to Allison, he considers tolerance levels below .4 problematic and those variables should be given careful attention regarding multicollinearity.

Examination of the tolerance levels for all grades reveals that there was little shared variance among the variables at the student level. The trends were similar for all

grades. ECE, Gifted, Mobil, and Female had tolerance values between .914 and .989. SES, Black, and Two Parents show some multicollinearity with tolerance levels between .714 and .809 in elementary and middle school. Among high school students the tolerance levels for SES, Black, and Two Parents were slightly lower, between .741 and .860. Even with the lower values, however, the tolerance figures indicate that 70% to 80% of the variance of those predictors was not related to the other predictors.

Overall at the student level, Gifted, SES, and ECE consistently made the greatest contribution to test scores with Black and Female having a fluctuating effect depending on the grade and test. The  $R^2$  values for all models in each regression were higher in middle and high school than in elementary, indicating a greater influence of the background variables at the higher grades. The effects of background variables were generally greater on mathematics scores than on reading.

#### *School Level Regression Analyses*

The independent variables at the school level were the percentages of the student-level background factors that constituted the school composition and, in the case of mobility rate, a characteristic of the school as a whole. The dependent variables were the school test scores and performance indices. The school test scores were the average scores for the students in the school for the various CTBS and KCCT tests. The school performance indices were the Norm Referenced Test (NRT) Index and the Academic Index. They were both constructed by formulas based on the number of students scoring at each performance level--Novice, Apprentice, Proficient, or Distinguished--in the various subjects at the school (see Chapter III for the formulas and explanations).



Table 36 presents the  $R^2$  values for the school-level regressions. The number of variables included in the different optimum school-level regressions varied, therefore, the number of variables in each optimum equation is indicated.

Table 36

*Effect Size for Optimum Regression Models of School Composition Variables on Aggregate CATS Tests and Performance Indices at Elementary, Middle, and High School*

CATS Tests and Indices	$R^2$ (# variables)		
	Elementary	Middle	High
CTBS Reading	.62 (2)	.89 (4)	.85 (3)
CTBS Mathematics	.59 (3)	.87 (3)	.90 (2)
KCCT Reading	.56 (2)	.86 (2)	.84 (2)
KCCT Mathematics	.58 (2)	.85 (3)	.91 (2)
NRT Index	.60 (2)	.89 (3)	.88 (2)
Academic Index	.74 (2)	.89 (3)	.89 (2)

*Note.* The NRT Index and the Academic Index are both on a scale of 1-140.

The school-level regression effects were much larger than for the student-level regressions with  $R^2$ s ranging from .56 to .91 at the school level, in contrast to values from .19 to .36 at the student level. This may seem a dramatic difference at first glance. It must be remembered, however, that the unit of analysis was at different levels. The school regressions represent aggregated data, and the comparisons were between groups (schools) rather than individuals. As Pedhazur (1997) has warned,

When individuals are used as the unit of analysis,  $R^2$  indicates the proportion of the total variance accounted for by the independent

variables. When, on the other hand, aggregates (e.g., classes, schools) are used as the unit of analysis,  $R^2$  indicates the proportion of variance between aggregates that is accounted for by the independent variables. Consequently, when the variance between groups is relatively small, one should be careful not to be overly impressed even with a high  $R^2$ . (p. 686)

In the current study, therefore, the extent of the variance of the dependent variables must be taken into consideration when assessing the effects of the independent variables. A comparison of the range and standard deviations of the CATS test at the individual level versus those at the school level in Appendices C and E reveals that the variability at the school level was indeed much smaller than the variability at the individual level. However, the range and standard deviation of the school scores were still substantial and there was still a high degree of variability among schools. Since the size of the variance of both the CTBS and KCCT were substantial at the school level--even if much smaller than the variance at the student level--a large  $R^2$  indicates that a substantial portion of the variance can be attributed to the independent variables included in the model.

The tolerance values of the school-level independent variables are shown in Table 37. They were more disparate than the tolerance values at the student level. At elementary school level, the tolerance values for %HiSES (.132) and %Two Families (.103) indicate that those two variables shared a great deal with the other variables. Mobil Rate and %Black had mid-range values that indicated moderate multicollinearity with the other variables--sharing 45% and 55% of their variance respectively. %Gifted

had less overlap with the other variables. %ECE and %Female had very high tolerance values indicating they share little of their variance and thus had greater independent influences on the regression equations.

Table 37

*Tolerance Values for School Composition Variables*

Variable	School Level		
	Elementary	Middle	High
%HiSES	.132	.092	.063
Mobil Rate	.542	.256	.102
%Gifted	.715	.790	.383
%Black	.432	.453	.393
%Female	.944	.733	.412
%Two Parents	.103	.101	.052
%ECE	.819	.195	.069

The results of the multiple regression analyses at the school level are in Tables 38 through 55. The tables include all seven variables in the order in which they entered the equations. However, in five of the regressions, in spite of the effort to include all of the variables in each regression by overriding the default settings in SPSS (explained in Chapter III) and because of the low number of subjects, the error probabilities were so high that only six regression equations were created by the program. As mentioned, at the school level, only the first two or three regression equations had statistically significant differences ( $R^2$  changes) and the last entered variables had very high

probabilities of error and are likely not meaningful. Therefore, only six regression models are reported based on the  $R^2$  change--not on the statistical significance--for all of the regressions with essentially no loss of pertinent information.

Table 38 presents the regression models for school-level third grade CTBS Reading results. The optimum model had two variables with  $R^2$  of .620. %HiSES (the percentage of students who are not eligible for free and reduced meals) had the largest  $\beta$  coefficient, .58. Mobil Rate (school mobility rate) had a negative unstandardized  $\beta$  coefficient, -.28. %Gifted and %Black were the third and fourth variables entered in subsequent models. Both had small positive coefficients, in contrast to the findings at the student level when Black consistently had a negative coefficient.

The regression models for third grade CTBS Mathematics are presented in Table 39. The optimum model, with an  $R^2$  of .585, includes three variables. As an example of the practical effects of the school composition variables on test scores, the optimum regression equation based on the unstandardized  $B$  coefficients is as follows:

$$\hat{y} = 46.70 + .18(\%HiSES) - .42 (\text{Mobil Rate}) + .11 (\%Gifted)$$

For one typical JCPS school, the calculations of the expected score would be as follows (based on the value of the three predictor variables in the optimum equation):

$$\hat{y} = 46.70 + .18 (52.8) - .42 (10.5) + .11 (8.9)$$

$$\hat{y} = 52.8$$

In this case, the %HiSES of the school was predicted to raise the score by 9.44 points, Mobil Rate to lower it by 4.36 points, and %Gifted to raise it by .93 points. The school actually had a score of 47.7 on the Third grade CTBS Mathematics test, which was lower than the expected value based on only those three variables.

Table 38

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Third Grade CTBS Reading School Scores on School Composition Variables (N = 87)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	39.55***		46.35***		46.50***		43.95***		50.57***		51.61***	
%HiSES	.26***	(.75)	.20***	(.58)	.18***	(.52)	.20***	(.57)	.20***	(.58)	.21***	(.62)
Mobil Rate			-.42***	(-.28)	-.42***	(-.28)	-.40***	(-.27)	-.40***	(-.27)	-.41***	(-.28)
%Gifted					.08*	(.13)	.08	(.13)	.08	(.13)	-.08	(.13)
%Black							.05	(.06)	.05	(.07)	.04	(.05)
%Female									-.14	(-.04)	-.14	(-.04)
%Two Parents											-.03	(-.06)
%ECE												
R <sup>2</sup>	.569		.620		.633		.635		.637		.637	
R <sup>2</sup> Change	.569***		.050***		.013*		.002		.002		.000	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$ .

Table 39

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Third Grade CTBS Mathematics School Scores on School Composition Variables (N = 87)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	39.70***		46.50***		46.70***		48.06***		56.09***		55.26***	
%HiSES	.27***	(.72)	.21***	(.56)	.18***	(.49)	.17***	(.47)	.18***	(.48)	.16**	(.42)
Mobil Rate			-.42***	(-.27)	-.42***	(-.26)	-.40***	(-.25)	-.41***	(-.26)	-.40***	(-.25)
%Gifted					.11*	(.16)	-.10*	(.15)	.10*	(.15)	.10	(.15)
%ECE							-.18	(-.06)	-.17	(-.06)	-.17	(-.06)
%Female									-.17	(-.05)	-.17	(-.05)
%Two Parents											.04	(.07)
%Black												
R <sup>2</sup>	.522		.566		.585		.588		.591		.591	
R <sup>2</sup> Change	.522***		.044***		.019*		.003		.002		.001	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$

Table 40 has the regression results for the fourth grade KCCT Reading test. The optimum equation had an  $R^2$  of .560 and included only two variables. In this case, %Two Parents was the first to enter and had a standardized  $\beta$  coefficient of .66. The second variable was %Gifted with a  $\beta$  coefficient of .15. %HiSES had a  $\beta$  coefficient of only .14, and was not included in the optimum equation.

The results of the fifth grade KCCT Mathematics regression are in Table 41. %HiSES was again the first variable entered with a standardized  $\beta$  coefficient in the optimum equation of .56. As in the elementary school CTBS regressions, Mobil Rate was the second variable with a similar  $\beta$  of -.28. Only those two variables were included in the optimum regression equation.

Table 42 shows the results of the regression of the elementary school NRT Index. The optimum model includes two variables: %HiSES with a  $\beta$  of .61 and Mobil Rate with a  $\beta$  of -.24. The  $R^2$  was .595. %Gifted had a  $\beta$  of .12 in Model 3. For this test, %Two Parents was not entered in the models until sixth and had a low coefficient.

The regression of the elementary school Academic Indices shown in Table 43 had the largest  $R^2$  among the elementary school regressions, .742. Like the NRT Index the optimum model includes only %HiSES with a standardized  $\beta$  of .73 and Mobil Rate with a  $\beta$  of -.19. %Black, which will be seen subsequently, had a major effect on both indices at middle and high school levels, was the least influential variable at the elementary level.

Table 40

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Fourth Grade KCCT Reading School Scores on School Composition Variables (N = 87)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )	<i>B</i>	( $\beta$ )
Constant	508.1***		509.9***		517.2***		518.8***		510.8***		508.5***	
%Two Parents	.82***	(.74)	.74***	(.66)	.65***	(.59)	.52**	(.46)	.59**	(.53)	.60**	(.54)
%Gifted			.18*	(.15)	.19*	(.15)	.18*	(.15)	.17	(.14)	.17	(.14)
Mobil Rate					-.36	(-.12)	-.33	(-.12)	-.27	(-.09)	-.29	(.10)
%HiSES							.09	(.14)	.10	(.14)	.10	(.15)
%Black									.11	(.08)	.12	(.08)
%ECE											.21	(.04)
%Female												
<i>R</i> <sup>2</sup>	.543		.560		.569		.572		.574		.576	
<i>R</i> <sup>2</sup> Change	.543***		.017*		.010		.003		.003		.001	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$



Table 41

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Fifth Grade KCCT Mathematics School Scores on School Composition Variables (N = 87)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	533.4***		549.0***		549.2***		553.1***		547.8***		535.4***	
%HiSES	.57***	(.73)	.44**	(.56)	.40***	(.51)	.39***	(.49)	.26*	(.34)	.27*	(.34)
Mobil Rate			-.96***	(-.28)	-.96***	(-.28)	-.90***	(-.27)	-.86***	(-.26)	-.77**	(-.22)
%Gifted						(.10)	.13	(.09)	.11	(.08)	.09	(.06)
%ECE							-.51	(-.06)	-.49	(-.08)	.45	(-.07)
%Two Parents									.24	(.18)	.35	(.27)
%Black											.17	(.10)
%Female												
R <sup>2</sup>	.531		.582		.589		.594		.599		.603	
R <sup>2</sup> Change	.531***		.051***		.007		.005		.004		.004	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$

Table 42

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Elementary School NRT Index on School Composition Variables (N = 87)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	48.1***		62.9***		63.3***		65.8***		82.2***		80.8**	
%HiSES	.68***	(.75)	.55***	(.61)	.50***	(.55)	.49***	(.54)	.49***	(.54)	.46***	(.51)
Mobil Rate			-.91***	(-.24)	-.91***	(.23)	-.88**	(-.23)	-.89**	(-.23)	-.88**	(-.23)
%Gifted					.19	(.12)	.19	(.12)	.19	(.12)	.19	(.11)
%ECE							-.33	(-.05)	-.31	(-.04)	-.31	(-.04)
%Female									-.34	(-.04)	-.35	(-.04)
%Two Parents											.07	(.05)
%Black												
R <sup>2</sup>	.561		.595		.606		.608		.609		.610	
R <sup>2</sup> Change	.561***		.035***		.011		.002		.002		.000	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$

Table 43

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Elementary School Academic Index on School Composition Variables (N = 87)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	40.79***		48.36***		43.84***		30.92**		31.98**		32.46**	
%HiSES	.48***	(.85)	.41***	(.73)	.30***	(.54)	.30***	(.53)	.30***	(.52)	.29***	(.52)
Mobil Rate			-.46***	(-.19)	-.42**	(-.17)	-.41**	(-.17)	-.39**	(-.16)	-.39**	(-.16)
%Two Parents					.21	(.22)	.21	(.21)	.20	(.21)	.18	(.20)
%Female							.27	(.05)	.29	(.05)	.28	(.05)
%ECE									-.23	(-.05)	-.21	(-.05)
%Gifted											.05	(.05)
%Black												
R <sup>2</sup>	.719		.742		.749		.751		.753		.755	
R <sup>2</sup> Change	.719***		.023***		.006		.002		.002		.002	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$

The tolerance values at the middle school indicate more multicollinearity among the independent variables than at the elementary school level (Table 37). %HiSES and %Two Parents continued to have the most overlap with other variables. Mobility had a lower tolerance than it did at elementary school. %ECE, which had a very high tolerance at the elementary grades, was quite low at .195. The overlap of %Black and %Gifted with the other variables was little changed. %Female was lower at .733, indicating more collinearity with the other school composition variables than in elementary, although still less than most other variables

The school-level regressions for the sixth grade CTBS Reading and Mathematics tests had the same four variables entered first with similar coefficients. For the sixth grade CTBS Reading scores in Table 44, the optimum model included four variables. %HiSES had a positive effect with a standardized  $\beta$  of .77. As in regressions at the lower grades, Mobil Rate had a negative influence on test scores with a  $\beta$  of -.42. %Black was the third variable with a positive influence and %Female was the fourth with a negative coefficient ( $\beta = -.16$ ). In Table 45 for the sixth grade CTBS Mathematics regression, only three variables provide the optimum equation in Model 3. The coefficients were similar, but %Female with a  $\beta$  of -.13 was not included in the optimum model.

Table 44

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Sixth Grade CTBS Reading School Scores on School Composition Variables (N = 24)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	27.17***		37.33***		31.27***		43.98***		41.77***		41.63***	
%HiSES	.39***	(.91)	.27***	(.63)	.30***	(.69)	.33***	(.77)	.24**	(.56)	.24***	(.54)
Mobil Rate			-.34**	(-.33)	-.40**	(-.39)	-.42***	(-.42)	-.41***	(-.41)	-.40**	(-.40)
%Black					.16	(.17)	.23 **	(.24)	.25**	(.26)	.23*	(.24)
%Female							-.33*	(-.16)	-.36*	(-.17)	-.34*	(-.17)
%Two Parents									.17	(.25)	.16	(.24)
%Gifted											.04	(.06)
%ECE												
R <sup>2</sup>	.821		.856		.872		.893		.899		.902	
R <sup>2</sup> Change	.821***		.035**		.016		.020*		.007		.003	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$

Table 45

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Sixth Grade CTBS Mathematics School Scores on School Composition Variables (N = 24)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	24.37***		35.64***		28.37***		39.40***		36.52***		36.30***	
%HiSES	.43***	(.90)	.30***	(.63)	.33***	(.69)	.36***	(.75)	.24*	(.50)	.23 *	(.49)
Mobil Rate			-.37**	(-.33)	-.45**	(-.40)	-.47***	(-.42)	-.46**	(-.40)	-.44**	(-.39)
%Black					.19	(.18)	.25**	(.23)	.28**	(.26)	.25*	(.23)
%Female							-.29	(-.13)	-.32	(-.14)	-.29	(-.13)
%Two Parents									.21	(.29)	.20	(.28)
%Gifted											.06	(.08)
%ECE												
R <sup>2</sup>	.817		.852		.871		.883		.893		.899	
R <sup>2</sup> Change	.817***		.035**		.019		.013		.009		.006	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$

The effect sizes for the sixth grade CTBS Reading test ( $R^2 = .893$ , Table 44) and Mathematics test ( $R^2 = .871$ , Table 45) regressions were higher than those at elementary school as were all the middle and high school effect sizes for these regressions. The  $R^2$  values were all between .843 and .906 at the middle and high school levels. The CTBS Reading test regression  $R^2$  of .893 means that almost 90% of the total variance among schools was explained by the four variables in the optimum model.

Table 46 shows the regression coefficients for seventh grade KCCT Reading scores. The optimum model had two variables with an explained variance of .872. The first predictor entered was Mobil Rate with a standardized  $\beta$  of  $-.58$ ; the second variable was %Two Parents with a  $\beta$  of  $.40$ . Of note, %HiSES, which had been among the first variables entered in most of the previous equations, entered last in this regression.

For the regression of the eighth grade KCCT Mathematics scores in Table 47, %HiSES again entered first with a standardized  $\beta$  of  $.56$  and unstandardized  $B$  coefficient of  $.51$  in the optimum model. Two more variables were included in the optimum model, %Gifted and Mobil Rate. The effect size was  $.852$ . Interpreting this equation in terms of the unstandardized  $B$  coefficients, the optimum regression equation is as follows:

$$\hat{y} = 500.3 + .51(\%HiSES) + .26 (\%Gifted) - .73 (\text{Mobil Rate})$$

For one typical JCPS middle school, the calculations of the expected score would be as follows (based on the values of the three predictor variables in the optimum equation):

$$\hat{y} = 500.3 + .51(62.6) + .26 (18.7) - .73 (12.1)$$

$$\hat{y} = 528$$

Table 46

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Seventh Grade KCCT Reading School Scores on School Composition Variables (N = 24)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	526.3***		498.4***		487.2***		503.9***		514.6***		518.4***	
Mobil Rate	-1.65*** (-.90)		-1.06*** (-.58)		-1.16*** (-.63)		-1.19*** (-.65)		-1.09*** (-.59)		-1.00*** (-.55)	
%Two Parents			.48*** (.40)		.54*** (.45)		.62*** (.52)		.53** (.44)		.44** (.37)	
%Black					.27 (.16)		.37* (.21)		.41* (.23)		.35* (.20)	
%Female							-.46 (-.13)		-.52 (-.14)		-.48 (-.13)	
%ECE									-.54 (-.17)		-.76 (-.24)	
%Gifted											.13 (.12)	
%HiSES												
R <sup>2</sup>	.802		.859		.872		.884		.891		.902	
R <sup>2</sup> Change	.802***		.057***		.014		.012		.007		.011	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$



Table 47

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Eighth Grade KCCT Mathematics School Scores on School Composition Variables (N = 24)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	479.0***		478.3***		500.3***		492.3***		487.0***		484.7***	
%HiSES	.81***	(.88)	.76***	(.84)	.51***	(.56)	.54***	(.59)	.39	(.42)	.40	(.44)
%Gifted			.27*	(.20)	.26**	(.20)	.24*	(.18)	.23*	(.18)	.23	(.17)
Mobil Rate					-.73**	(-.34)	-.82**	(-.38)	-.80**	(-.37)	-.82**	(-.38)
%Black							.22	(.11)	.25	(.12)	.24	(.12)
%Two Parents									.28	(.20)	.29	(.21)
%ECE											.14	(.04)
%Female												
R <sup>2</sup>	.777		.815		.852		.858		.862		.863	
R <sup>2</sup> Change	.777***		.038**		.037**		.006		.004		.000	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$

The actual 2002 KCCT Mathematics score at the above school was 524 (which was within the bounds of the Standard Error of the Predicted Value). %HiSES accounted for an increase of 31.93 points to the score, %Gifted accounted for an increase of 4.86 points, and Mobility Rate accounted for a decrease of 8.83 points in the school score.

The regression equation for the Middle School NRT Index is shown in Table 48. The optimum model had three variables: %HiSES at standardized  $\beta = .68$ , Mobil Rate at  $\beta = -.40$ , and %Black at  $\beta = .16$ . The  $R^2$  was .887 and would account for almost 90% of the total variance on the Middle School NRT Index.

Table 49 shows the results of the multiple regression of the Middle School Academic Index on school composition variables. The optimum model includes the same three variables as with the NRT Index. Variables %HiSES and Mobile Rate were the first and second entered with similar standardized  $\beta$  coefficients. %Black had a larger positive  $\beta$  of .23. The  $R^2$  was the same as that for the NRT Index at .887.

The tolerance values among the high school composition variables were lower than in middle school, indicating a higher degree of multicollinearity among the variables (Table 37). The lowest were %HiSES, %Two Parents, and Mobil Rate, which share 90% or more of their variance. The tolerance of %Black was slightly lower than at earlier grades, but the tolerances of %Gifted and %Female were both quite a bit lower, indicating much more overlap with the other variables than in lower grades.

Table 48

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Middle School NRT Index on School Composition Variables (N = 24)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	9.29		38.50***		22.38		49.31*		42.08*		55.47*	
%HiSES	1.10***	(.91)	.76***	(.63)	.82***	(.68)	.89***	(.74)	.60**	(.50)	.55*	(.46)
Mobil Rate			-.97**	(-.34)	-1.14***	(-.40)	-1.19***	(-.42)	-1.15***	(-.41)	-1.04**	(-.37)
%Black					.43	(.16)	.57	(.21)	.64**	(.24)	.69**	(.26)
%Female							-.71	(-.12)	-.79	(-.14)	-.86*	(-.15)
%Two Parents									.54	(.29)	.40	(.27)
%ECE											-.68	(-.14)
%Gifted												
R <sup>2</sup>	.835		.872		.887		.899		.908		.912	
R <sup>2</sup> Change	.835***		.037**		.015		.012		.009		.004	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$

Table 49

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Middle School Academic Index on School Composition Variables (N = 24)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	26.48***		44.45***		31.54***		24.74**		26.29**		36.39**	
%HiSES	.62***	(.90)	.41***	(.60)	.46***	(.67)	.25	(.37)	.24	(.35)	.25	(.37)
Mobil Rate			-.60**	(-.37)	-.73***	(-.45)	.70***	(-.44)	-.68***	(-.42)	-.70***	(-.44)
%Black					.35**	(.23)	.38**	(.25)	.34**	(.22)	.40**	(.26)
%Two Parents							.34	(.34)	.35	(.33)	.38	(.36)
%Gifted									.10	(.10)	.09	(.09)
%Female											-.29	(-.09)
%ECE												
R <sup>2</sup>	.813		.857		.887		.899		.909		.915	
R <sup>2</sup> Change	.813***		.044**		.029**		.013		.010		.006	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$

The high school regression results were varied. The regression of ninth grade CTBS Reading scores on the school composition variables is in Table 50. The optimum equation includes %HiSES at standardized  $\beta = .78$ , Mobil Rate at  $\beta = -.28$ , and %Black at  $\beta = .17$ --a pattern similar to earlier equations. The optimum equation for the ninth grade CTBS Mathematics regression (Table 51) differed, including only two variables: %Two Parents at standardized  $\beta = .83$  and %Gifted at  $\beta = .17$ . Interestingly, %HiSES entered the mathematics regression sixth. The  $R^2$  values for the ninth grade CTBS regressions were .895 for the reading test and .843 for the mathematics test.

For the regression of the tenth grade KCCT Reading school scores shown in Table 52, the optimum equation is Model 2. It includes %HiSES with a standardized  $\beta$  of .34 and %ECE at  $\beta = -.22$ . This was the only instance in this set of regressions in which %ECE was included in an optimum equation. The  $R^2$  was .843, about 84% of the variance between schools.

The eleventh grade KCCT Mathematics regression also shows only two variables in the optimum equation in Table 53, %Two Parents (standardized  $\beta = .69$ ) and Mobil Rate ( $\beta = -.31$ ). %HiSES was the next variable entered, but its contribution ( $R^2$  change = .008) does not meet the decision rule for inclusion as a meaningful predictor. This suggests that a major portion of social class in this equation is subsumed within its collinearity with %Two Parents and Mobil Rate. The  $R^2$  is the highest in this study at .906. For the eleventh grade KCCT Mathematics test, the standard deviation was also the largest among the school-level KCCT tests (Table E3). With more than 90% of the variance explained and a standard deviation of 30.1, it is clear that most of the difference among schools was predicted by two demographic variables.

Table 50

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Ninth Grade CTBS Reading School Scores on School Composition Variables (N = 21)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	22.87***		35.35***		26.38**		21.79**		23.83**		30.20*	
%HiSES	.43***	(.89)	.31***	(.62)	.38***	(.78)	.21	(.44)	.17	(.35)	.15	(.30)
Mobil Rate			-.47**	(-.34)	-.38*	(-.28)	-.30	(-.22)	-.30	(-.22)	-.38	(-.28)
%Black					.11	(.17)	.16	(.23)	.15	(.23)	.17	(.26)
%Two Parents							.30	(.45)	.30	(.45)	.35	(.51)
%Gifted									.08	(.11)	.07	(.09)
%Female											-.13	(-.09)
%ECE												
R <sup>2</sup>	.788		.835		.850		.862		.868		.872	
R <sup>2</sup> Change	.788***		.048**		.015		.011		.007		.004	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$

Table 51

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Ninth Grade CTBS Mathematics School Scores on School Composition Variables (N = 21)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	14.70***		17.02***		24.44***		18.25*		24.45*		25.26*	
%Two Parents	.76***	(.94)	.67***	(.83)	.56***	(.69)	.64***	(.79)	.65***	(.81)	.70**	(.87)
%Gifted			.14*	(.17)	.14*	(.17)	.14	(.16)	.12	(.15)	.13	(.15)
Mobil Rate					-.29	(-.18)	-.22	(-.14)	-.30	(-.18)	-.30	(-.19)
%Black							.07	(.09)	.09	(.11)	.09	(.12)
%Female									-.14	(-.07)	-.15	(-.08)
%HiSES											-.04	(.07)
%ECE												
R <sup>2</sup>	.877		.895		.907		.911		.913		.914	
R <sup>2</sup> Change	.877***		.018*		.013		.003		.003		.000	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$

Table 52

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Tenth Grade KCCT Reading School Scores on School Composition Variables (N =21)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	406.8***		454.5***		428.3***		407.5***		407.6***		409.8***	
%HiSES	1.43***	(.89)	.97***	(.61)	1.18***	(.74)	.54	(.34)	.44	(.27)	.40	(.25)
%ECE			-1.70**	(-.36)	-1.46*	(-.31)	-1.03	(-.22)	-.90	(-.19)	-.42	(-.09)
%Black					.32	(.15)	.50	(.22)	.51	(.23)	.49	(.22)
%Two Parents							1.20	(.55)	1.27	(.58)	1.26	(.57)
%Gifted									.20	(.09)	.25	(.11)
Mobil Rate											-.52	(-.12)
%Female												
R <sup>2</sup>	.794		.843		.855		.871		.875		.877	
R <sup>2</sup> Change	.794***		.050**		.012		.016		.004		.001	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$



Table 53

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of Eleventh Grade KCCT Mathematics School Scores on School Composition Variables (N = 21)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	434.8***		471.4***		465.3***		424.1***		414.8***		414.8***	
%Two Parents	2.14***	(.93)	1.59***	(.69)	.85	(.37)	.76	(.33)	1.01	(.44)	1.02	(.44)
Mobil Rate			-1.44**	(-.31)	-1.35**	(-.29)	-.89	(-.19)	-.78	(-.17)	-.78	(-.17)
%HiSES					.58	(.35)	.72	(.43)	.67	(.40)	.67	(.40)
%Female							.65	(.12)	.54	(.10)	.53	(.10)
%Black									.20	(.09)	.20	(.09)
%Gifted											-.01	(-.00)
%ECE												
$R^2$	.868		.906		.915		.924		.927		.927	
$R^2$ Change	.868***		.038**		.008		.009		.003		.000	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$

Table 54 shows the regression of the high school NRT Index on school composition variables. The optimum equation, Model 2 with an  $R^2$  of .875, includes two variables: %Two Parents (standardized  $\beta = 1.09$ ) and %Black ( $\beta = .26$ ). The optimum equation for the high school Academic Index regression in Table 55 is similar with an  $R^2$  of .894 for Model 2 with the same two variables as the NRT Index: %Two Parents ( $\beta = 1.12$ ) and %Black ( $\beta = .30$ ). Interestingly the variable %Black was included as the second most influential predictor for both the high school level indices, even though it was not such an influential predictor for the tests from which the indices are calculated. On the CTBS Reading and Mathematics tests, %Black was the third and fourth entered variable respectively (Tables 49 and 50). For the high school KCCT Reading and Mathematics tests, which represent 40% of the Academic Index, %Black was the third and fifth entered variable respectively (Tables 52 and 53). It must be remembered that the CTBS was reported in NCE scores and the KCCT in scale scores, but the indices are calculated by a weighted formula based on performance level categories (explained in Chapter III).

#### *Trends for the School Composition Regressions*

Trends among the school composition variables are apparent in a comparison of effects of the regressions across the grades; on the CTBS, KCCT, and Indices; and on the reading and mathematics tests.

#### *Effect Sizes*

As in the individual level regressions, the effects at school level were much larger in the upper grades (Table 36). Also, the  $R^2$  values were higher at middle and high school than at elementary, as was the pattern at the individual level.

Table 54

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of High School NRT Index on School Composition Variables (N = 21)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	-9.70		-41.01**		-14.96		-11.05		-0.76		7.27	
%Two Parents	1.82***	(.92)	2.17***	(1.09)	1.81***	(.91)	1.69***	(.85)	1.72***	(.86)	1.70***	(.85)
%Black			.51**	(.26)	.39	(.20)	.37	(.19)	.41	(.21)	.41	(.21)
Mobil Rate					-.72	(-.18)	-.72	(-.18)	-.84	(-.21)	-.58	(-.15)
%Gifted							.17	(.09)	.15	(.07)	.10	(.05)
%Female									-.23	(-.05)	-.32	(-.07)
%ECE											-.43	(-.10)
%HiSES												
R <sup>2</sup>	.838		.875		.885		.889		.891		.891	
R <sup>2</sup> Change	.838***		.037**		.010		.004		.001		.001	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$

Table 55

*Standardized ( $\beta$ ) and Unstandardized B Coefficients from Forward Linear Regression of High School Academic Index on School Composition Variables (N = 21)*

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )	B	( $\beta$ )
Constant	13.55**		-9.26		5.81		7.77		6.46		4.43	
%Two Parents	1.15***	(.92)	1.41***	(1.12)	1.20***	(.96)	1.14***	(.91)	.98**	(.79)	1.02**	(.81)
%Black			.38**	(.30)	.30**	(.24)	.29*	(.23)	.29*	(.23)	.30*	(.24)
Mobil Rate					-.42	(-.16)	-.42	(-.17)	-.40	(-.16)	-.60	(-.23)
%Gifted							.09	(.07)	.07	(.06)	.09	(.07)
%HiSES									.13	(.14)	.11	(.12)
%ECE											.25	(.10)
%Female												
R <sup>2</sup>	.845		.894		.903		.905		.907		.908	
R <sup>2</sup> Change	.845***		.049**		.009		.003		.001		.001	

\* $p < .10$ . \*\* $p < .05$ . \*\*\* $p < .01$

### *Tolerance Trends at the School Level*

The tolerance values (Table 37) indicated that %HiSES and %Two Parents were extensively collinear with the other school composition variables at all grades. Mobil Rate was moderately collinear with the other variables at elementary grades but had a high degree of overlap in the middle school and even more in high school. Although %ECE had a high tolerance in elementary school, indicating the population of ECE identified children was distributed independently of the other school composition variables, in middle and high school, %ECE extensively shared variance with the other school composition variables. Tolerance values indicate that %Gifted and %Female were for the most part independently dispersed in elementary and middle school grades, but they became more collinear with the other variables in high school. %Black had a consistently medium level of collinearity with the other school composition variables, which does not vary much from grade to grade.

### *Influence of the Variables*

To compare the strength of the coefficients within a regression it is only necessary to look at the standardized  $\beta$  coefficients. These have been described in the presentation of the regression tables, but they are summarized in Table 56 for the CTBS and KCCT tests scores and in Table 57 for the school NRT and the Academic Indices

Table 56

*Standardized  $\beta$  Regression Coefficients for Optimum Equations for the School Level**CATS Tests*

CATS Test	Independent variables						
	%HiSES	%Two Parents	Mobil Rate	%Black	%Gifted	%Female	%ECE
	Elementary						
CTBS Reading	.58		-.28				
CTBS Mathematics	.49		-.26		.11		
KCCT Reading		.66			.15		
KCCT Mathematics	.56		-.28				
	Middle						
CTBS Reading	.77		-.42	.24			-.16
CTBS Mathematics	.69		-.40	.18			
KCCT Reading		.48	-.58				
KCCT Mathematics	.56		-.73		.26		
	High						
CTBS Reading	.78		-.28	.17			
CTBS Mathematics		.83		.17			
KCCT Reading	.61						-.36
KCCT Mathematics		.69	-.31				

In line with the low tolerance between %HiSES and %Two Parents at the school level, these two variables alternated in their position as part of the optimum regression equations (Tables 38-55). %HiSES was the strongest predictor for 12 of the 18

regressions at the school level. The standardized  $\beta$  coefficients for %HiSES in the school composition ranged from .56 (Table 39) to .63 (Table 48) when it entered the equation first at the school level. In five of the other six equations, the variable %Two Parents was the strongest predictor and entered into the equation first. When it entered first, the  $\beta$  coefficients for %Two Parents were larger than those for %HiSES and ranged from .66 (Table 40) to 1.12 (Table 55).

Table 57

*Standardized  $\beta$  Regression Coefficients for Optimum Equations for the School NRT and Academic Indices*

Index	Independent variables			
	%HiSES	%Two Parents	Mobil Rate	%Black
Elementary School				
NRT Index	.61		-.91	
Academic Index	.73		-.19	
Middle School				
NRT Index	.68		-.40	.16
Academic Index	.67		-.45	.23
High School				
NRT Index		1.09		.26
Academic Index		1.12		.30

The standardized  $\beta$  coefficient of 1.12 for %Two Parents on the high school Academic Index was the highest in the entire set of regressions. The magnitude of this

is such that in this instance a change in the standard deviation of %Two Parents of one  $z$  unit produces a change of 1.12 standard deviation  $z$  units in the Academic Index. Since at the high school level the standard deviation for %Two Parents was 13.1 (Table D3) and the standard deviation for the Academic Index was 16.38 (Table E3), a change of one standard deviation (13.1%) in %Two Parents would be equivalent to an 18.35 point increase ( $1.12 \times 16.38 = 18.35$ ) in the Academic Index.

In the elementary school regressions, %HiSES was the first entered variable for five of the six regressions. In those regressions %Two Parents entered much later except for the Academic Index regression in which %Two Parents entered third. Also, when either %HiSES or %Two Parents entered the regression, the value of the other variable simultaneously decreased in the regression equations. The same general pattern followed in middle school. %HiSES entered first for five of the six regressions. For the KCCT Reading test regression, %HiSES entered last, but %Two Parents entered second. In high school, %Two Parents was the variable that entered first most frequently for four of the six equations, and when it did not, %HiSES entered first.

The Mobility Rate (Mobil Rate) consistently had a negative coefficient in the school-level regressions. Mobility Rate was the first entered and strongest variable in only one regression, but it was the second entered eleven times and the third entered once. Thus, it was included in 13 of the 18 optimum equations. It had significant regression coefficients in those optimum equations. The entry of the other variables into the equation had little effect on the value of its coefficients. The coefficients for Mobil Rate were higher in middle school than in elementary or high school.

In the elementary school regressions, %Gifted entered into the equation third in



four of the six regressions and for KCCT Reading it entered second. The standardized  $\beta$  coefficient for %Gifted in those equations was between .10 and .16. In contrast, in middle school %Gifted entered early only for the KCCT Mathematics test regression, when it entered second. In high school, %Gifted entered second for the CTBS Mathematics test regression, but entered fourth or later for the other regressions. %Gifted had a positive coefficient, except for one regression when it entered next to last.

Interestingly, %Black always had positive coefficients in the school level regressions, which is contrary to the results at the student level. The magnitudes of the coefficients varied but were generally low in elementary grades. In middle and high school, %Black was more influential. It entered third in five of the six regressions at middle school. In the high school regressions for the NRT and Academic Indices, %Black entered second and had its highest coefficients with standardized  $\beta = .26$  and  $\beta = .30$ , accounting for almost 4% and 5% of the variance on those indices respectively (Tables 54 and 55).

The variables %ECE and %Female were usually the last entered into these school-level regression equations. The coefficients for the last entered variables had very high significance probabilities, which means the differences were very likely due to chance. %ECE was always negative, and %ECE was included in the optimum equation only once when it entered second for the tenth grade KCCT Reading test regression. In contrast, ECE was a major predictor at the student level. The coefficients for %Female were generally low and fluctuated between negative and positive. %Female was in the optimum equation only once, for the sixth grade CTBS

Reading school regression with a standardized  $\beta$  of -.16.

The statistical significance probabilities are useful in identifying those regression coefficients and  $R$  squares that were unlikely to be due to chance. Generally in the school-level analysis, only the first two regression models had an  $R^2$  that was significant, and only the variables entered first and second had statistically significant coefficients. Those significant variables were most frequently either %HiSES or %Two Parents and Mobil Rate. The highly correlated variables of %HiSES and %Two Parents were the first entered for 17 of the 18 school-level regressions. Mobil Rate was the second variable entered for 11 of the regressions and the first-entered variable once. %Gifted was the second entered three times, and %Black was the second entered two times in the high school NRT and Academic Indices.

#### *Points Accounted for by Demographic Factors*

The concrete effects of school composition factors on school CATS tests scores are demonstrated by calculating the points attributable to the school composition factors. These calculations may be done using the unstandardized  $B$  coefficients or the standardized  $\beta$  coefficients. Because the school composition variables were percentages, the unstandardized  $B$  regression coefficients must be multiplied by the percentage of the corresponding variable to produce the points attributable to the variables. For example, consider the optimal equation for the third grade CTBS Reading Test (Table 38, Model 2) for which  $B = .20$  for %HiSES and  $B = -.42$  for Mobil Rate.

$$\hat{y} = 46.35 + .20 (\%HiSES) - .42 (\text{Mobil Rate})$$

To find the points attributable to %HiSES when Mobil Rate is controlled in a

particular school, one must multiply .20 times the %HiSES of the school. For each unit change in the independent variable, in this case each percentage point, the dependent variable, the test score, would be changed by the value of the unstandardized  $B$  coefficient. For example the calculation of the predicted score,  $\hat{y}$ , for one particular elementary school is as follows:

$$\hat{y} = 46.35 + .20 (52.8\%) - .42 (10.5)$$

$$\hat{y} = 46.35 + 10.56 - 4.41$$

$$\hat{y} = 52.50$$

The actual points predicted by these regression equations are a function of various factors including the regression coefficients, the percentage of each variable in the school population, the variability of the different tests, and the construction of the test itself—concepts beyond the analysis of this study. However, the examples in Tables 56 and 57 provide an estimation of the relative importance of the variables to the results within each test and index. However, because the effects of the points on the different test scores and performance indices are determined by the standard deviations of the respective tests and indices, the magnitude of the regression coefficients cannot be compared across the different test or regressions.

### *Reading Versus Mathematics Results*

Unlike the trend at student level, there were no apparent patterns regarding the effects of demographic school composition factors on the different tests and indices with regard to mathematics and reading at the school level. The fluctuation among the significant variables that were included in the optimal regression equations appeared to be random and more likely a result of chance variations in the correlations due to the

small number of subjects and multicollinearity at the school level than to real differences in the influence of the variables.

### Summary

This chapter presented findings related to the core research questions regarding the effects of individual student background factors on student CATS tests results and the effects of demographic school composition factors at the school level on CATS school performance measurements. It was a secondary analysis of data for all students and schools from a large urban school district in Kentucky.

Population parameters for the background factors set the stage for the regression analysis. They indicated that the number of students qualifying for free and reduced meals was lower at higher grades, that student mobility is greater at the higher grades, and that students identified as disabled (ECE) represented about 10% in elementary and middle schools but were fewer in high school. Gifted students represented about 10% of the population. Less than half the students lived with both their mother and father at all grade levels.

At the student level, the correlations among background variables and between background variables and the test scores were not strong. Most were low; the highest were moderate. The strongest correlations among background variables were among Black, SES, and Two Parents. The highest correlations between test scores and background variables were with SES, Gifted, ECE, and Black.

Population parameters for the school-level factors revealed differences in the composition of the school populations. At elementary and middle school levels, the differences were greatest for proportions of students on free and reduced meals and

students identified as gifted or as having two parents in the home. At high school level, the difference in proportions of Blacks and females in the schools was more extreme than in the lower grades.

Generally, the correlations between school composition variables were much stronger than the correlations between comparable background variables at the individual level. The correlations of the CATS tests scores with the demographic variables were also much stronger at the school level though %Gifted was the exception. The strongest correlations with school scores were between the percent of students with a high SES (not on free and reduced meals), mobility rate, and the percent having two parents.

The regression analysis at the individual level revealed that Gifted, SES, and ECE consistently had the greatest influence on test scores, with Black and Female having fluctuating effects depending on the grade and test. The  $R^2$  values, which represent the proportion of the variance affected by the variables in the equation, ranged from .18 to .36. The effects were greater in middle and high school than in elementary school and greater on the mathematics tests than on the reading tests. Black was only included in the optimum equations of the CTBS tests, whereas Female was only included for the KCCT Reading tests at middle and high school.

The  $R^2$  values were much larger at the school level than for the student level regressions. They ranged from .59 to .91. Even though the variance was smaller between schools than between individuals, the school composition factors had a important effect on the scores. For the accountability performance measures, the NRT Index and the Academic Index, the school-level composition factors had substantial  $R^2$

values, ranging from .595 for the elementary NRT to .894 for the high school Academic Index.

At the school level, more than at the student level, some of the variables demonstrated high multicollinearity. The %HiSES and %Two Parents had the greatest overlap between each other and with the other variables. Mobility Rate also had a high degree of shared variance at the high school level. Generally, %HiSES or %Two Parents and Mobility Rate were the variables that significantly affected the regression equations. %HiSES and %Two Parents had positive effects on the scores while Mobility Rate had a negative effect. The strength of the influence of the other variables varied. %Gifted always had a positive effect on the scores as did %Black. This latter finding needs further investigation as there are conflicting possible explanations, such as high multicollinearity or the context of racial concentration in the school district. The current study cannot distinguish such explanations. The effects of %Black were only important and included in the optimum equations for the CTBS test and for the middle and high school NRT Index and Academic Index. %ECE was only important in one regression where it had a negative effect. The coefficients for %Female were generally very low and statistically insignificant. Finally, unlike at student level, there were no apparent trends in the effects on reading and mathematics scores at the school level.

Several statistical findings suggest that the school performance indices, which are ultimately from the school-level achievement tests are inflated and exaggerate the differences between schools. First, the range of the scores was much wider on the NRT Index than on the CTBS NCE scores from which the index is constructed. Second,

%Black appeared in the Academic Index despite not being a predictor for the KCCT tests from which the Academic Index is derived suggesting that the calculations are somehow transforming the relative influence of the various school composition factors. The size of the differences is particularly pertinent because most of the differences among schools were accounted for by socio-demographic factors as evidenced by the very high  $R^2$  values at the school level.

## CHAPTER V

### DISCUSSION AND CONCLUSIONS

#### Overview of the Study

Standards-based reform is based on the position that holding all students to high standards will enable all students to learn at higher levels and thereby increase educational equity (Fuhrman, 2001). Inherent in the arguments given for standards-based reform are several assumptions: (a) that all children can learn regardless of their backgrounds, (b) that being held accountable will motivate school practitioners to make the instructional changes necessary for all children to succeed academically regardless of personal background or school composition, and (c) that schools have the power and capacity to make the changes necessary for all children to be successful. However, critics have pointed out that having the same standards for all schools without providing the support and resources for disadvantaged students to reach those standards is inequitable (M. S. Smith, 1995). Other critics fear that a reliance on assessment and school accountability detracts from addressing the root causes of low achievement (Linn, 2000; Thomas & Bainbridge, 2001).

The Kentucky Education Reform Act (KERA) was launched within the national movement toward standards-based education reform that continues under the No Child Left Behind Act (NCLB) of 2001. In 1990, Kentucky adopted KERA and instituted an accountability system designed to raise student achievement by setting high standards



for all children and schools and by assessing the results. The principles of standards-based reform--assessment and local school accountability--form the core of the KERA curricular provisions. The current Commonwealth Accountability Testing System (CATS) requires schools to make constant progress toward a score of 100 on a 0-140 point scale in the year 2014 (Petrosko, 2000).

Overall, Kentucky students have raised their scores dramatically on the tests used in the KIRIS and CATS accountability systems. They have also improved on the national NAEP tests, although not to the same extent (Petrosko, 2000; Poggio, 2000). Kentucky results since KERA, however, indicate that poverty continues to have a substantial negative effect on individual and on school-level test scores. Overall in Kentucky, poverty and race seem to be influential in affecting student achievement scores and even more influential on the aggregate level in predicting school scores (Guskey, 1997; Roeder, 2000). Girls have higher scores on the CATS assessments than boys (Ennis, 2002; D. C. Smith et al., 1999), but the effects of school gender composition on aggregate school scores have not been widely reported. The evidence regarding mobility and disability is mixed at both the individual and school levels. Mobility and disability are frequently negatively correlated with achievement, but the correlations may be a function of poverty (Guskey, 1997; D. C. Smith et al.). Factors representing family structure have received minimal attention regarding their contribution to CATS scores (Muñoz & Dossett, 2000). There has been no analysis of the effects of giftedness on CATS assessments. Evans (2001) has developed an innovation component configuration map for use in evaluating gifted programs that may produce more systematic research on giftedness in the future.

Problems are evident in this accountability system that demands a diverse population of students reach the same high standards at the same time. In spite of demonstrated improvements in student achievement, achievement gaps (e.g., between rich and poor, black and white, male and female) persist both at the level of the individual student and among schools. These effects of demographic factors are not taken into account either in comparing individual student achievement or aggregate school-wide performance (Petrosko, 2000). Overall low-achieving students have not been able to catch up with the high-achieving students, and most disadvantaged, low-performing schools are still low performing (Pitts & Reeves, 1999; Roeder, 2001b) because educators have not been able to change substantially the effectiveness of those schools.

Studies of the effects of demographics on change scores have been contradictory. Several have found that demographic factors and change scores are not as strongly associated with demographic factors as are both individual and aggregate absolute scores (Ennis, 2002; Guskey, 1997; Luvisi, 2000). However, those studies did not control for previous achievement. When previous scores were included as a dependent variable in the regression, poverty had a significant negative effect on CTBS scores and a smaller negative effect on the CATS Accountability Index (Roeder, 2001b). These few studies indicate that the distinction between the effects of demographics on absolute (cross-sectional) scores and on change scores has not been adequately delineated.

Considering the implications of high-stakes testing and the effects of demographic factors on student achievement, several researchers have explored

alternative approaches to accountability including longitudinal models emphasizing change scores, statistical techniques to control for demographic variables, and a more comprehensive approach to accountability that includes inputs such as student demographics and instructional capacity as well as outcomes (Lee & Coladarci, 2001; Linn, 2001; Sanders, 1998a; Wheelock, 2000). However, in Kentucky the dilemma of maintaining high standards for all students while compensating for demographic influences on student achievement has not received adequate attention or debate.

### Study Methods

The over-arching purpose of this study was to contribute to an understanding of the interaction of demographic factors and student achievement that will inform evaluation of accountability systems and processes. The objectives were to examine (a) the relationships between student background factors and student achievement scores at the individual student level and (b) the relationship between demographic school composition factors and aggregated scores at the school level.

This study was a secondary analysis of data collected by the Jefferson County Public Schools (JCPS) and the Kentucky Department of Education (KDE). The measures of student and school achievement were the scores from the CATS achievement tests used for accountability in Kentucky. These CATS tests are comprised of subject-area-specific Kentucky Core Content Tests (KCCT), which are criterion-referenced tests aligned with the Kentucky standards, and the CTBS/5 Survey Edition, a national norm-referenced test.

The individual student predictors included in the analysis were gender, socioeconomic status, ethnicity, mobility, family structure, giftedness, and disability.

The school level demographic predictor variables addressed were parallel, representing the percentages of each category of student that compose the school population and Mobility Rate.

Each level of students and schools--elementary, middle and high--was analyzed independently for the two levels of data aggregation (individual and school). The data were analyzed by multiple regression using SPSS 11.5, which also provides descriptive statistics and tests of significance.

The study first addressed the relationship of individual student demographic variables with student achievement on the KCCT and CTBS reading and mathematics tests. Second, the study examined the extent to which school demographic composition affects aggregate school scores for the same tests. The study also examined the relationship of those demographic variables with the CATS measures of school performance used for accountability--the school Norm-Referenced Test Index and the school Academic Index.

This study adds important findings to the quest for understanding dynamics affecting both underachieving students and low-performing schools. The study provides a statistical analysis and comparison of the extent of the relationship between demographic factors and student achievement both at the level of the individual student and at the level of the school in a large urban/suburban district. It compares these effects across grade levels: elementary, middle, and high school and for reading versus mathematics. It demonstrates the extent of the differences in the effects of socio-demographic factors on the norm-referenced CTBS and the criterion-referenced KCCT. It examines the effects of school composition factors on the two school-performance

indices: The Norm-Referenced Test (NRT) Index and the Academic Index and explores the effects of additional factors that have received scant or no attention in previous Kentucky studies (percents mobility, gifted students, two parent families, and ECE students). The findings have implications for policy changes in regard to accountability systems and attempts to improve education.

### Discussion

This section addresses the ramifications and implications of the core research questions, which are

1. To what extent do individual student background factors--gender, SES, ethnicity, mobility, family structure, gifted and talented status, and disability --predict individual student achievement?
2. To what extent do school composition variables--proportions representing gender, SES, ethnicity, mobility, family structure, gifted and talented status, and disability--predict school performance?

In this section, descriptions of the students and schools in JCPS first provide a context for understanding the results of the multiple regression. Then the results of the regression respond to the core research questions regarding the relationship of student background factors to individual student achievement and the relationship of school composition factors to school performance. The discussion addresses implications regarding whether standards-based accountability is effective in ameliorating the effects of background and school composition on student achievement and whether schools have demonstrated the capacity to overcome those effects. Finally, recommendations for policy changes and future studies are offered.

## *Context*

### *Students*

The population parameters for the students in Jefferson County Public Schools as a district describe the average characteristics of the students and provide a measure by which to compare the different schools in the school analysis that follows. For example, 55% of elementary students qualified for free and reduced meals in the district, but certain schools had a much lower or higher percentage. Some trends in the student population parameters deserve mention because they may be pertinent to the discussion of the regression analysis as they apply to the research questions.

The number of students that qualified for free and reduced lunches was lower in the higher grades than in the early grades (Tables B1, B4, and B7). This does not necessarily mean that there were fewer students living in families in poverty at older ages; rather it is generally accepted that fewer students apply for the subsidy in the higher grades to avoid the stigma. Also, some high school students work and have their own income for lunch. Whether the factors that influence older students' decisions to apply for free meals would also affect students' choice of schools or their achievement is not apparent. It is just one of those many and complex phenomena that constitute the cultural and psychological influences on student achievement and school composition.

Mobility, the number of times a student changes schools, was greater in middle school than in elementary and greater in high school than in middle, except for the eleventh grade, which was almost as low as elementary (Table B10). The standard deviation also increased indicating a greater range of mobility among the students.

Students change schools for many reasons, which were not identified in this

study. The reasons for student mobility could be related to the child's background, the characteristics of the schools, or the child's relationship with the schools. At elementary school level, it seems logical that most changes would be due to families changing their residence and high mobility would be indicative of financial instability. However, parents may also transfer students for such reasons as discontent with the original school. At middle school, there was an increase in mean mobility and in the proportions of students who move frequently (Table B5), but there is no reason to expect that financial instability would be higher in families with middle school age children. The increase in mobility at higher grades may be due to students' changing schools more frequently because of personal issues such as discipline problems or discontent with the school and suggest a possible relationship to school climate, which bears investigation. For example, one strategy for dealing with students with behavior problems is to transfer them to a different school where the student can start afresh and associate with different peers.

The proportions of male and female students in JCPS was close to 50% with slightly more males than females (Tables B3, B6, and B9), which reflects national trends (U.S. Census Bureau, 2003b). The precipitous change in the eleventh grade where there were 4% fewer males than females suggest that possibly a higher proportion of males than females drop out of school.

A similar concern about selection factors affecting dropouts is prompted by the sharp decrease in percent of Black students in the eleventh grade (Table B3, B6, and B9). The gradual decrease in Black students at each grade mirrors the trend in the population at large (U.S. Census Bureau, 2003b), but the 4% decrease from the tenth to

the eleventh grade suggests a real phenomenon other than a decrease in their numbers in the population.

The proportion of students with disabilities was consistent through elementary and middle school, but substantially lower in high school (Tables B3, B6, and B9). That trend raises hopes that those students had progressed beyond needing special help, but it is also possible that disabled students disproportionately dropped out of school. The increase in the percentage of gifted students in the higher grades (Tables B3, B6, and B9) is accompanied by a decrease in the absolute number of gifted students and a greater proportional decrease in the number of regular students, which suggest disproportionate dropouts among the non-gifted. These trends in student demographics could significantly affect school scores but are areas that were beyond the scope of this study.

Less than half the students in JCPS lived with both their mother and father (Tables B3, B6, and B9). This is not consistent with national percentages according to the U.S. Census Bureau (2003a), which reports 69% of all children under 18 years old and 58% of central-city children live with two parents. However, the Census Bureau also reports that 32% of Black central-city children and 38% of all Black children live with two parents. Therefore, the discrepancy may be due in part to there being a higher proportion of Black children in JCPS than in the U.S. on average, as well as the possibility mentioned in Chapter III of inaccurate reporting and coding by parents and staff. A more extensive examination of the situation in Jefferson County in comparison to the balance of the state of Kentucky and to similar cities outside Kentucky would be useful in determining if the relatively low percentage of two parent families was an



accurate measure and useful for making state and national comparisons.

Although the highest correlations among the student background variables were only in the moderate range, they still indicate some degree of covariance that could be due to similar or related causes. The high correlations among SES, Two Parents, and Black suggest that Black and low income students were more likely also to have single parents but that could be due to many circumstances, including a decrease in income due to divorce or a young, never-married mother with few employment skills.

Distinguishing the multiple forms of family structure and the relationships of the different forms with SES could be important in identifying causality but are beyond this study. The correlations among some of the variables leads one to expect a degree of multicollinearity among those background variables, which could be caused by an overlapping factor or factors. The regression analysis was designed to provide some indication of the degree of multicollinearity among the variables.

### *Schools*

The population parameters of the school district are important in regard to the extent the socio-demographic composition of individual schools diverges from the district averages. The variables in this study represent a few of the many factors that create different environments in different schools. The differences were apparent in the large range of certain variables among different schools as shown in Appendix D. Most notable was the range of the percentage of students in different schools who qualify for subsidized meals and the percentage of students with two parents (Appendix D). This difference among schools was relevant to the examination of the effects of school composition factors on student achievement.

The high correlations of the school composition variables and the school-level test scores and performance indices suggested a strong relationship between the background variables and the CATS test scores. It remained for the multiple regression analysis and other more sophisticated procedures to provide evidence whether and to what extent those effects are independent or represent the same underlying factor.

### *Student Background Factors and Student Achievement*

At the student level, the impact of each background variable on student scores is apparent from the regression coefficients produced by the regression of the student CATS test scores on student background variables. In this study, income level (SES), identification as gifted (Gifted), and identification as a special needs student (ECE) had a substantial influence in predicting student scores. Those variables were consistently included in the optimal equations. Black and Female were included in the optimal equations only in special cases. The multicollinearity of the background factors at the student level was low and was demonstrated by the independent contributions of each additional variable in the successive  $R^2$  for each regression.

#### *SES*

SES (based on free and reduced lunch eligibility) usually had the highest correlation with the different tests and was most frequently the first variable entered into the equations. It was always one of the first three variables included in the equations. This result is in line with other results nationally (Coleman et al., 1966; Mayeske & Beaton, 1975; L. S. Miller, 1995) and in Kentucky (Ennis, 2002; Muñoz & Dossett, 2001; D. C. Smith et al., 1999). When SES alone was entered in the equations, it accounted for 10-18% of the variation. When the other variables were entered, it

contributed a smaller proportion. Although the effect of SES on achievement was moderate at the student level, as in White's (1982) meta-analysis, the small difference in scores attributable to SES by the regression analyses can be very important when assigning students to categories of achievement in which small differences in scores can result in differential classification. as in the CATS system.

### *Gifted and ECE*

Gifted or ECE--designations that apply to only about 10% of the student population--can have a great impact on predicting an individual student's score. This is not surprising since both identification as gifted and as special needs students serve as proxies for students' ability or previous achievement (Dika & Singh, 2002). In this study, both Gifted and ECE added considerably to the regression equation after SES was entered, which would indicate that their effects were to some degree independent of SES. However, the measures of SES were very limited, and did not identify the gradations of income higher than 185% of federal poverty line, the guideline for free and reduced lunch eligibility (California Food Policy Advocates, n.d.).

This study used only the qualification of students for free and reduced lunch subsidy as a measure of SES because the purpose of the study was to compare the effects of comparable factors at the student and school levels used in the Kentucky system. An investigation of the relationship of SES and student achievement that used more precise measurements of the complete range of family income and that included additional measures of SES such as parent's income and educational attainment would perhaps find a closer relationship between high SES and giftedness, representing the greater access to social networks and resources of high-income families described by

Dika and Singh (2002). In this study, it is possible that a higher correlation of Gifted with high income was not detected because of the truncated measurement of family income.

Although SES was always the first entered variable for the CTBS tests, Gifted was first entered for the KCCT Mathematics tests and for the high school (tenth grade) KCCT Reading test. Except for third grade CTBS tests (when few gifted have been identified) and the fourth grade KCCT Reading test, Gifted had the largest standardized beta regression coefficient in the optimum equations, which is not surprising since it is a measure of previous achievement and ability.

The effects of Gifted and ECE on student scores emphasize the need for taking previous achievement into account when attempting to assess student learning (Lee & Coladarci, 2001). If previous scores were not included, there is no way to know if the students were improving proportionately or not, since previously high-achieving students could lose ground and still be above grade level. An analysis that measured student improvement from year to year (Linn, 2001) and incorporated previous scores would provide a more direct measure of what individual students have actually learned in school, the value-added (S. K. Miller, 1992), rather than the social capital they bring with them to school (L. S. Miller, 1995).

### *Black*

The variable Black was included in the optimal equations for the CTBS tests but not for the KCCT tests. For the KCCT tests, Black has low negative coefficients that added little to the prediction of the regression equation. This suggests that the attribute of being Black has little effect on a student's KCCT scores when socioeconomic, gifted,

and special needs backgrounds are taken into account. However, on the CTBS tests, Black has a significant effect on the scores. This differential impact of predictors on the CTBS and KCCT tests prompts speculation about the causes of the discrepancy. For example, the differences could be caused by the different skills required for a norm-referenced vs. a criterion-referenced test, which may or may not be altered by classroom learning--the only learning over which schools and teachers have any control. Madaus and Clarke (2001) have pointed out the negative effects of testing on minority students because of influences outside of school on their achievement. That principle would seem to be operating with regard to the results on the CTBS tests.

### *Female*

Being female seemed to be an advantage only on the KCCT Reading tests at fourth, seventh, and eighth grades, which is supported by statewide KCCT test results (KDE, 2002b) and other studies (Ennis, 2002; D. C. Smith et al., 1999). Otherwise, Female did not add to the predictions. Although it had positive coefficients on the other reading tests, they were much smaller. On the mathematics tests, Female was usually insignificant as was found by Lubienski (2002) and Zhang and Manon (2000) when looking at the overall scores. However, the Lubienski study and Zhang and Manon as well found differences at the extremes when they looked at subgroups within their research subjects. Any differences between males and females at the top or low ends of the achievement spectrum or of the socioeconomic spectrum could not be identified in the current study.

The differences in achievement on the KCCT reading and mathematics tests of males and females have implications for instruction and for assessment practices.

Basically, differences may be due to something in the differential development or social roles based on gender, which would have implications for methods of instruction. The ramifications of gender differences and schooling are too many and too complex to review for this discussion. But the importance of context again is implicated because this research takes place within the culture of a large Kentucky city. Within that context, differences in achievement may be related to different experiences females and males have in school because of their relationships with teachers and peers (Sadker & Sadker, 1994). For example, females may have learned the Kentucky curriculum reflected on the KCCT Reading tests better because they have tried harder due to classroom experiences or because of gender expectations that are evident in attitudes and behaviors of themselves or others.

#### *Variables with Little Impact*

Neither Two Parents nor Mobil were included in any of the optimum equations. The Two Parent variable consistently had the lowest coefficients and thus the least influence on the scores, which corroborates the other studies that found family structure was not significant when other background variables such as income, previous achievement, and mother's education were controlled (Entwisle & Alexander, 1995; Mayeske & Beaton, 1975; Phillips, Brooks-Gunn, Duncan, Klebanov, & Crane, 1998; Wilson, 1987). In this study, income, race, gender, and previous achievement were controlled because they entered the equations before the Two Parent variable for family structure, which was found to be negligible at the student level.

However, when the equations that include Mobil are examined, the size of the coefficient for Mobil, which was consistently negative, would indicate that student

mobility could account for a considerable negative effect on an individual student's score, which is consistent with Medsker's findings (cited in Petrosko, 2000). For example, in Model 5, Mobil has an unstandardized  $B$  coefficient of -12.5 for the Tenth Grade KCCT Reading Test scores (Table 30). Yet the large negative coefficient indicates that for the very few students that have a mobility of 2 or 3 (mobility was ratio data), the impact on their expected scores could be substantial. This is a limited subset of students, too small to have an impact in the larger set of thousands of students, yet very likely associated with extreme poverty or unsocial behavior. Policy makers need to be aware of and accommodate the special needs for this unique group of students.

When Mobil entered the prediction equation, however, it added only negligibly to the  $R^2$  after the other variables were entered, indicating those other factors are more influential in predicting scores in the population, as also reported in the D.C. Smith et al. (1999) study. Thus, the current study does not support the contention that student mobility was a major factor overall in predicting average student achievement at the individual level as reported by Medsker (cited in Petrosko, 2000).

#### *The Influence of Student Background Factors on Test Scores*

The regression analysis demonstrated that student background factors had a larger impact on the scores at the middle and high school levels than at elementary (Table 32). If schooling counteracted the effects of background and social disadvantage, the influence of background factors on student scores would be less at the higher grades. In fact, it increased. That result would not logically be due to inherent student characteristics or to deprivation in the vulnerable preschool years. There are many possible explanations for this finding, which include the following. One is simply

cumulative effects of income, that older students have experienced the debilitating effects of poverty or the social and personal advantages of middle of upper class status in all aspects of their lives over a longer period (L. S. Miller, 1995; Wilson, 1987). A second possibility is differential parental input both in the degree to which parents may be involved in school and in the attitudes and values toward education imparted to their older children. A third possibility is that the schooling experience itself has contributed to the effects of demographic factors on student achievement. Solomon et al. (1996) demonstrate that student poverty affects teachers' attitudes and instructional practices. These attitudes and practices would have an increasingly negative effect on student achievement if compounded over time. As Brookover et al. (1979) demonstrate, schooling has a greater influence on low income and minority students than on more affluent and White students, and teacher and peer attitudes and teacher practices have the potential for greater effects the longer they are sustained. A fourth possibility is that the psychological and social development of children as they move into teen years makes them more susceptible to the effects of demographic factors, which may affect their self-concept and status in the society.

Some of the unique discoveries of this study were the differences between the effects of student background variables on the different CATS tests. The differential effects of Black and of female on the CTBS and KCCT at the individual student level were described in earlier sections. Black was an important negative predictor on the CTBS but not on the KCCT. In contrast, being female was only an important factor on the KCCT Reading tests at middle and high school. The reasons for a certain background variable having a proportionately weaker effect on KCCT than on CTBS



test scores could be crucial to decision-making on the choice and use of different kinds of tests as well as on instructional issues.

For example, an analysis of the CTBS by Marzano and Costa (1988) indicated that variance on CTBS tests was affected primarily by retrieval skills (remembering facts) while less than 4% of the variance in student achievement was accounted for by higher order cognitive skills such as comparing or inferring. The factual knowledge required for the CTBS is not purposefully aligned with what is taught in Kentucky (although, of course, it may overlap), but the KCCT is specifically aligned with Kentucky standards and academic expectations and requires higher level thinking skills in addition to recall. Therefore, one would expect a more direct connection between the KCCT tests and instruction in Kentucky schools, i.e., teachers would have a greater opportunity to mitigate the effects of student background. One would not expect the outcomes of instruction in higher level thinking skills to be detected by a test that relies primarily on recall. Since Black is only an important factor of CTBS test results, does it follow that race does not have an influence on students' achievement of the Kentucky Core Content, which is measured by the KCCT? Does it follow that female students have better learned the Reading Core Content because as a group they score higher on the KCCT Reading tests? Clearly, it is essential that the validity of this suggested relationship be pursued by additional analysis of the data.

Another interesting result was that at the student level, giftedness was a stronger influence on the KCCT scores at the high school level than at the other levels. Since the KCCT is a criterion referenced test, this finding suggest that the advantages of being gifted and of learning the tested curriculum are greater at the high school level than at

middle and elementary school. The context for this finding is the managed-choice student assignment plan in JCPS, which results in a concentration of gifted students at a few schools, especially at high school level. In light of the managed-choice plan, this finding could have implications in regard to the distribution of students, to equal access to knowledge, and to expectations of school performance.

### *School Composition and School-Level CATS Scores*

At the school level, there were far fewer subjects than at the individual level. There were only 87 elementary schools, 24 middle schools, and 21 high schools available in JCPS and included in the study. In addition, the correlations among certain of the school composition factors were much greater than the student-level correlations with student background factors. These conditions produce two major limitations of this study that must be kept in mind when interpreting these statistical findings: power and multicollinearity.

### *Power and Small Numbers of Schools*

Although all JCPS schools were included in the study, in effect the 2002 JCPS schools serve as samples of JCPS schools for all years. School scores for a given year are based upon the students in the grades assessed for that year's cohort. Comparisons are made from year to year with subsequent cohorts tested to determine progress at each grade level. Thus the entire population represents both the complete cross-sectional set of students for the year 2002 (the data used) and the data from cohorts of previous and subsequent years. The particular year analyzed is in effect a sampling of one year out of many. When considering the school data as a sample, however, the number of subjects (schools) was much smaller than would generally be appropriate to produce

regressions with sufficient power to make reliable statements about the regression coefficients--especially for the 24 middle schools and 21 high schools.

The problem of lack of power is mitigated by three considerations. First statistical significance is useful as an indication of the likelihood that the coefficient represents a real effect. Allison (1999) points out that even "in a small sample, statistically significant coefficients should be taken seriously" (p. 57). Second, the consistency of results across all grades and tests supports the position that the coefficients represent true effects for both the significant factors and some of the nonsignificant ones. As Allison also notes, "a nonsignificant coefficient is extremely weak evidence for the absence of an effect" (p.57). Considering that this was an exploratory study, the pattern or consistency of some of the nonsignificant coefficients suggest certain demographics are likely to produce significant findings in further research. Third, the use of forward stepwise procedures and the selection of optimum equation models (that represented a .02 increase in the  $R^2$  of the previous model) served to decrease the ratio of cases to variables. In the optimum equations, there were usually only two or three variables entered into the equation, which lowered the ratio to seven or eight cases per variable--still not the fifteen cases to be considered adequate (Stevens, 2002)--but better than the three cases per variable that would result if all the variables had been entered into the regressions simultaneously.

#### *Multicollinearity Among School-Level Factors*

The high correlations among certain of the demographic factors, their low tolerance levels (high proportion of shared variance), and the patterns in the regression equations indicate a high degree of multicollinearity among the school composition

variables--much higher than it was among the individual student variables. Several problems are presented by the multicollinearity of the variables. One problem is an increase in the standard error of the coefficients, which makes it difficult to get an accurate measure of the regression coefficients of those collinear variables (Allison, 1999). The large standard error can cause small differences to be magnified and the regression slope to be unstable because of the mathematics of the regression model. The unstable regression slope can produce a large, significant  $R^2$  without any of the coefficients being significant (Morrow-Howell, 1994). In the current research, that phenomenon was not apparent because the coefficients in the selected optimum equations were statistically significant. The problems of multicollinearity were somewhat mitigated by the use of forward regression, which produced equations that contained two or three significant coefficients as well as a significant  $R^2$ . Most variables in the optimum equations were significant even among the high school ( $N = 21$ ) regressions.

One possible downside of forward regression is that it capitalizes on chance in finding significant relationships (Tabachnick & Fidell, 2001). Chance may result in the “overfitting” (Tabachnick & Fidell) to a particular sample because of small fluctuations in the correlations that determine the order of the entry of variables into the regression equations. There is no problem if the data from 2002 is being considered independent of other years. In that case, it represents the total population of schools in JCPS and the fit of the equation to the population represents the real relationship. Chance “overfitting” may be an issue if the data serves as a sample to discover trends across the years. In that case the issue of overfitting exists in terms of comparing school scores

(cohort variation) from year to year.

The chance issue is also counteracted in the current study by the consistent finding that the same three variables--%HiSES, %Two Parents, and Mobility Rate--were important in the regressions across different grades. However, the fluctuation in the order of the entry among these three variables was likely due to chance and small differences in their correlations with the dependent variables, issues exacerbated by the high collinearity of the variables and the small number of schools.

Redundant or overlapping variables (low tolerance) cause the second problem with multicollinearity because the magnitude of the partial coefficients may be reduced and, therefore, the unique contribution of each variable may not be apparent. This can result in the regression coefficients being in the wrong direction or of inaccurate size (Allison, 1999; Morrow-Howell, 1994). The reduction in the value of the regression coefficients of highly correlated variables was evident in this study as additional predictor variables were entered in the subsequent models. The most extreme example was the relationship of %HiSES and %Two Parents, one of which usually entered the equation first and then the other was relegated to a later entry with a much reduced coefficient.

It is also possible that the direction of the coefficients was reversed with the variable %Black. The percentage of Black students in a school was strongly and negatively correlated with the income level and the percentage of two parent families in the school. Therefore, the positive coefficient for %Black when the other school composition variables were controlled could be a statistical anomaly, which requires further investigation. Issues related to multicollinearity are addressed in conjunction

with the discussion of the specific demographic school-composition factors in this study.

*Effects of School Composition Factors on CATS scores.*

The comparison of the effects of demographic variables on individual student scores versus the effects on school scores is a unique contribution of this study to the on-going research on the effects of Kentucky's accountability system. The differences between the influential variables at the school level and those at the student level were found to be sizeable and important.

In general, this study supports the results of other studies that have documented the relationship of student demographics to school performance on the Kentucky accountability assessments (Guskey, 1997; Luvisi, 2000; Muñoz & Dossett, 2001; Pitts & Reeves, 1999; Roeder, 1999, 2000). But, in addition, this study points out the much greater influence of school composition factors on school scores than the comparable student background factors have on student scores. The  $R^2$  for the school-level regressions was about three times larger than the  $R^2$  at the individual level (Tables 51 and 54). Although the variability at the school level was not as broad as the variability at the individual level, there are large differences in scores among schools, especially among middle and high schools (Appendix E).

School composition variables accounted for the major proportion of the variance in school scores--from 57% to 92%. Considering that school accountability decisions are based on small differences in accountability scores (which are primarily based on the Academic Index, Table 59), the effects of the school composition variables on the scores were crucial. Whether a school was considered to be reaching or exceeding its

goal or not and whether it was in rewards or considered in need of assistance was demonstrably affected by the demographics of the school.

This study also points out that the predictive socio-demographic factors at the school level were not simply the result of the aggregation of the factors that were influential at the student level. Instead, certain of the factors that were influential at the school level appear to represent different constructs that were more powerful in affecting the variance in student achievement than the simple sum of students' individual backgrounds. Put another way, the individual-level socio-demographic factors appear to operate in distinctive ways in the aggregate as a characteristic of the school rather than of the individual students. This finding is consistent with Wilson's (1987) theoretical notion of concentration effects in which increased concentration of dispossessed citizens in inner cities results in the loss of social buffers and exacerbates historic discrimination. Similar consequences may devolve from concentration effects in schools producing differences in the school learning climate (cf. Brookover et al., 1979).

Similar to the effects at the student level, at the school level the size of the effects of demographic variables was higher at the higher grades (Table 36). A large portion of the variance among schools was explained by school composition factors at the elementary grades--close to 60% on the tests and NRT Index; almost 75% on the Academic Index. In contrast, at the middle and high school levels, close to 90% of the variance on both the achievement tests and the performance indices was explained by school composition factors.

*Different Influential Factors at the School and Student Levels*

As described in the previous section, at the student level, only SES, Gifted, and ECE consistently and substantially contributed to the prediction equations. The percentage of high SES students (%HiSES) was also of primary importance at the school level, but the percentage of gifted students (%Gifted) and percentage of ECE students (%ECE) were less influential. Two Parents was the most insignificant variable in the student analysis, yet %Two Parents was a major predictor at the school level. Student mobility had little impact at the student level, but school Mobility Rate was a major predictor of school-level scores. At the student level, Black had a low, negative coefficient and was a substantial predictor on the CTBS scores. In contrast, %Black always had a positive coefficient at the school level, contributed to a minor extent to the CTBS results at only three grades, and was a substantial positive predictor on the NRT and Academic Indices at middle and high school. %Female was a weak predictor at both student and school levels and fluctuated between having a positive or negative coefficient.

#### *Influential Factors at the School Level.*

For the school regressions (Tables 33-50), the first two and sometimes the third models had significant effects, and the first two or three variables entered into the equations had significant coefficients. Three demographic factors stood out as most influential at the school level, being both statistically significant and frequently included in most of the optimal equations: the percentage of students not on free and reduced lunches (%HiSES), the percentage of two-parent families (%Two Parents), and the Mobility Rate. These factors typically accounted for about 60% of the variance at elementary level and 85% at middle and high school level.



*Socioeconomic status and family structure.* The highest degree of collinearity was evidenced between %HiSES and %Two Parents. Those variables consistently had the highest correlations and the lowest tolerance (shared variance) values. In addition, the way that %HiSES and %Two Parents seemed to substitute for each other in the regression models demonstrated their shared variance. At the school level, therefore, %Two Parents seemed to be representing a very different construct than it did at the individual level. The concentration of students with two-parent families in a school had the same effect as the concentration of higher income students.

Two-parent families are not restricted to higher income students nor single parents to low-income families, which was reflected in the low overlap between two parent families and higher SES at the individual level. At the school level, the concentration of two parent families and of children not on free and reduced meals (and conversely schools with high poverty and single parent families) was almost indistinguishable. This suggests that the percentage of students not on free and reduced meal programs and the percentage of two-parent families were both measurements of the same underlying factor or factors related to income, consistent with studies that have found the effects of non-intact families to be due primarily to the socioeconomic situation of the family (Entwisle & Alexander, 1995; Huston et al., 1994; L. S. Miller, 1995; Wilson, 1987). It is not clear whether one of these factors is more directly causative than the other or whether both are indicative of yet another underlying construct related to the uneven distribution of advantaged and disadvantaged students in the school populations.

At the lower grades, %HiSES was usually the most influential factor, but at the

high school level, %Two Parents was the most influential. The high multicollinearity and small number of schools increases the instability of the prediction equations. Therefore, chance due to the large standard error (Morrow-Howell, 1994) could be responsible for whether %HiSES or %Two Parents entered the equation first and consequently was attributed the greater predictor. A more substantive explanation for this finding is that at the upper grades, the percentage of single parent families may serve as a better proxy for concentrated poverty than free and reduced lunch status because many eligible high school students do not apply for the subsidized meal program. This explanation is supported by the observation in JCPS that among the high-poverty elementary schools, the percent of students on free and reduced lunches was greater than the percentage without two parents, whereas among high-poverty high schools, the percent without two parents was greater than those on free and reduced lunches.

Bankston and Caldas (1998) had similar findings, but a different interpretation from their study of tenth grade students in Louisiana. They found that the percentage of single parent families in a school consistently had the strongest effect on student achievement test scores. Percentage school poverty had a much lower negative coefficient, and race was insignificant when school proportion of single parents was included in the analysis. They interpreted their findings to mean that single parent status had the more direct affect on student achievement than SES or race; but they did not consider the effects of multicollinearity nor that high school students are less likely to apply for free and reduced meals. Their study looked at only tenth grade and did not reveal any patterns across grade levels. Another difference was that Bankston and

Caldas looked at single-parent, female headed families in comparison to all other arrangements, whereas, JCPS defines dual-parent (Two Parent) families as those that include the birth mother and father. All other arrangements--step parents, foster parents, other relatives, etc.--are lumped together in the single parent classification and are not included in %Two Parents. It is also important to note that there was a very high correlation in the Louisiana schools between the percentage of Black students and of single-parent families. In contrast in JCPS, the correlation was much stronger between SES and race than between family structure (%Two Parents) and race (%Black) at the school level.

The primary impact of school SES level (whether measured by percentage of students not on free and reduced lunch or by percentage of students with two parents) is in line with other Kentucky studies that have found poverty to be the primary predictor of school performance (Guskey, 1997; Roeder, 2000, 2001b). Guskey and Roeder did not include family structure in those studies, but school poverty level (% Free and Reduced Lunch) was included and had a strong, negative relationship with the school performance. Because of the high multicollinearity between SES and the number of parents in a family, when one of the variables is left out, the other has a greater influence on the equation (Allison, 1999). These conflicting results reinforce the points (Hoisch & Miller, 2003; Jennings et al., 2000) about the difficulty of maintaining fidelity of constructs from study to study and about differences in contextual settings in the social sciences.

*Mobility Rate.* At the school level, mobility was the second most important variable in predicting CATS scores, contrary to its impact at the student level where it

was among the least influential. At the school level, Mobility Rate always has a substantial, negative predictive effect, which is similar to the findings by Mao (1997) and Medsker (cited in Petrosko, 2000). The greater effect of Mobility Rate on school scores in contrast to the weak influence of mobility at the student level indicates that school Mobility Rate represents some additional forces beyond the simple aggregation of the individual student scores.

Mobility Rate showed evidence of shared variance with %Two Parents and %HiSES especially at the middle and high schools; nevertheless it added significantly (from 3.5% to 5.0% to the regression equation after %HiSES entered first. Guskey's study of 49 schools (Guskey, 1997) in another (smaller) Kentucky school district did not find the independent contribution of Mobility Rate to school scores when income level was entered first. Rather, only the percent of students on free and reduced lunch was significant, predicting 51-78% of the variance. With only 33 elementary, 11 middle, and 5 high schools, it is likely that Guskey's study was not sensitive enough to distinguish among multiple variables. However, it is also possible that the difference may be due to the different context in the two districts. For example, if the mobility rate is an indication of extreme poverty or of antisocial behaviors associated with certain types of poverty, the smaller school district may not have the concentration of poverty or the degree of extreme poverty that is found in the larger, more urban district.

Hypotheses about the negative relationship between school Mobility Rate and school scores include the possibility that Mobility Rate may be an indication of the concentration of a particularly extreme form of poverty, e.g., an indication of a concentration of children from unstable homes or who are homeless in a school

population.

Alternatively, the high predictability of Mobility Rate might represent something about school climate. For example, school discipline practices that caused some students to change schools might also affect the learning environment of the non-mobile students in a given school and thus lower student achievement. This hypothesis is congruent with research that found student-body characteristics affect school climate (Brookover et al., 1979) and that perceptions of discipline problems are associated with declining levels of achievement at the high school level (Lumsden, 2000). Also, student poverty level affects teacher attitudes and instructional approaches (Solomon et al., 1996), which could easily result in higher mobility in the student body.

When Mobility Rate entered after %Two Parents, it did not add significantly to the equation, demonstrating more overlap with %Two Parents than %HiSES did. This indicates that %Two Parents is a school composition factor that encompasses more than a measure of income. Perhaps mobility rate is one part of an array of factors subsumed by %Two Parents that describe the populations of schools--part of the culture of the students and families that select those schools or that reside in the student attendance zones for those schools in the JCPS student assignment system. Those schools with low proportions of two-parent families that instead were characterized by high proportions of single parent, female-headed families could be schools populated with high concentrations of children drawn from neighborhoods characterized by urban poverty, Wilson's (1987) "truly disadvantaged."

#### *Effects of the Percent of Black Student*

When controlling for other variables including SES or family structure and

mobility, the effect of %Black on school scores was positive in this study, a finding that is both intriguing and unusual. But it is supported by other studies on the CTBS in Jefferson County (Muñoz & Dossett, 2001) and on the Accountability Index in Jefferson and Fayette Counties (Roeder, 2001b). Roeder acknowledged that the correlation between race and poverty was higher than the correlation between race and the test scores and that his results, therefore, might represent multicollinearity.

Bankston and Caldas (1998) found a similar phenomenon on the tenth grade achievement test results in Louisiana. They first found that the percentages of African Americans and of families in poverty (Free and Reduced Lunch) in a school had negative regression coefficients in their multilevel linear modeling analysis. However, when the percentage of single-parent families was included in the model, the coefficients for both race and poverty become positive instead of negative. They considered this change in sign as a possible “consequence of partialling when two independent variables are more strongly correlated with each other than either is with the dependent variable” (p. 720). In contrast to this current study and Roeder’s findings (2001b), when percentage of poverty was included but not the percentage of single parent families in the Louisiana analysis, the coefficient for the percentage of African American students in a school remained negative. Again these differences may be related to the different context for the studies in the different locations. The Louisiana study was a statewide study with a lower poverty level overall and the percentages of female-headed families was very highly correlated with the percentages of African American students in a school, unlike in Jefferson County where most schools are desegregated by race.

The findings of Lee and Coladarci (2001) on NAEP scores contradict the current study. They found that race had no significant effect on school mean achievement in either Kentucky or Maine on the eighth grade NAEP. They also found that the racial and social gaps varied very little between schools, which they interpreted to mean that the gaps were not a result of school effects. However, Lee and Coladarci included all the school districts in Kentucky, and the proportions of Black students in Kentucky are substantially lower than the proportions in both this study (Jefferson County) and Roeder's (2001b, Jefferson and Fayette Counties). One would expect the effects of school racial percentages to have more influence (positive or negative) in schools with 50% Black students than in schools that average 5% Black, the percentage reported in the Lee and Coladarci study.

In the current study, the percentage Black students was the only independent variable that had a much higher correlation with the percentage of two parents and the percentage of non-subsidized lunch students than with the test scores. For each regression, %HiSES or %Two Parents entered the equation before %Black. Therefore, the positive coefficients for %Black could well be a result of multicollinearity.

On the other hand, the positive coefficients of %Black may be a true reflection of the effects of school race proportions on student achievement--that a higher percent of Black students in a school does result in higher test scores. The consistency of the positive coefficients for %Black, some of which were significant at the higher grades supports the possibility that in JCPS and Fayette County, the percentage Black in a school has a positive relationship to achievement.

Since all of the JCPS schools were racially desegregated, with most having

between 20 and 50% Black students, higher proportions of Black students would make the proportions of Black and White students in schools closer to being equal. It is conceivable that more equal proportions of students by race could have a beneficial effect on school climate, teacher expectations of all students, and student achievement. It is important to remember that the proportions of Black students in JCPS were very different from other districts and cities where residential segregation results in extreme differences in racial percentages among the schools.

The variations in the influence of race in Maine (Lee & Coladarci, 2001), Kentucky (Roeder, 2001b), Illinois (Sutton & Soderstrom, 1999), Washington (Abbot & Joireman, 2001), and Louisiana (Caldas & Bankston, , 1998), and nation-wide (Lubienski, 2002) must be considered in light of the different racial context in each region. For example, in Louisiana most Black students attend predominantly Black schools whereas in Jefferson County, no high school had greater than 71% Black students and most were less than 50%. In Maine, only about 5% of the students were Black.

In the current study, the positive effects of %Black were considerable at the middle and high school levels, but not at the elementary level, which is another finding that prompts questions about the dynamics of student achievement and school performance. A closer examination of the trends in %Black across schools in relationship to the school scores could reveal if the positive effect holds at the extremes of the %Black in schools, or if there is an optimal or minimum percentage Black that affects the scores. It is also possible that the stronger correlations among %HiSES, %Two Parents, and Mobility Rate at middle and high school increase the effects of



multicollinearity on the regressions. The findings regarding race in this study suggest many possibilities but at best only tentative conclusions.

#### *School Composition Factors with Little Impact*

The percent of gifted students and of ECE students in a school had only small effects on predicting school scores after the other variables were entered. The simplest explanation, which was not contradicted by this study, is that the aggregate scores of the few gifted and ECE students in the schools directly affect the school scores, as would be expected, but that the proportions of Gifted and ECE in a school do not have additional school-composition effects. As Guskey (1997) concluded, the apparent effects of concentration of ECE students are caused by the association of ECE students with the concentration of poverty in schools rather than by the concentration of ECE students per se.

Although the effects of percentages of gifted and ECE students were small relative to the other variables, they could still account for more than nine points on the KCCT (Table 56), which could be critical for a school's accountability rating. This presents an argument for an accountability system based on longitudinal measurements of student improvement that will not be affected by changing the composition of the student-body. As the Kentucky system is now constructed, school improvement is gauged by comparing the accountability scores of one grade at a school with the scores of the same grade the next year (a cross-sectional approach to measuring improvement over time; Linn, 2001) under the assumption that the populations of the different years will be comparable. However, that assumption does not always hold true; and especially in small schools, changes in the demographics of the student body can

change school scores independent of any changes produced by instruction. For example, adding gifted students to a school or eliminating ECE students can be expected to improve school scores. Therefore, it behooves the schools to attract as many gifted students as possible and to avoid as many ECE students as possible. Even if such actions do not actually occur, the perception that it occurs is detrimental to public attitudes toward the accuracy and fairness of the accountability system.

The regression coefficients for %Females were erratic and added little to the regression equations except for CTBS Reading Grade 6 when %Female predicted about 4-5 points. The coefficients were generally negative except for a positive coefficient for KCCT Math Grade 11. This direction was counter to expectations based on the results of the student regressions in this study and others (Ennis, 2002; D. C. Smith et al., 1999) in which female had a positive coefficient on CATS reading scores, where females consistently outscore males.

The inconsistency of the results in the case of %Female indicates the size and direction of its coefficients were likely due to chance. %Female was frequently a last entered variable in the regressions, and the standard errors were highest with the last entered variables. Therefore, unlike the evidence for the effects of gender at the individual level, there was inadequate evidence that the proportion of females in schools has any influence on student achievement.

#### *CTBS Versus KCCT*

The only contrast between the effects of demographic factors on the CTBS and on the KCCT tests was produced by the variable representing the percentage of Black students in a school. The percentage Black was an influential factor only on the CTBS

scores at the middle and high school levels and on the NRT Index and the Academic Index at the high school level. In all cases, the regression coefficients for the percentage Black were positive, which was addressed in an earlier section. It is quite plausible that the effects of the percentage Black on the CTBS and NRT scores are related to the CTBS being a norm-referenced test, which is less affected by schooling than a criterion-referenced test (Popham, 1981). However, considering the percentage Black had a positive regression coefficient, that interpretation would be inconsistent with the interpretation of the individual results, in which Black had a negative effect. The strong effects of the percentage Black on the school-level middle and high school Academic Indices is puzzling, considering that the percentage Black was not an important or significant factor any of on the KCCT Reading and Mathematics tests that are part of the Academic Index. The Academic Index includes five additional subject-area components, however, which may be more influenced by the percentage Black in the school.

### *Reading Versus Mathematics*

There were no apparent trends regarding the effects of demographic school composition factors on the different tests and indices in regard to mathematics and reading at the school level. With the limited size of  $N$  in the school level portion of this study, the fluctuation among the significant variables that were included in the optimal regression equations appears to be random and more likely a result of chance variations in the correlations than of real differences in the influence of the variables. Thus, although school composition factors were the major determinant of the differences among school scores on CTBS and KCCT Reading and Mathematics tests, school-level

demographic factors had no more apparent effect on school mathematics scores than on school reading scores or vice versa. Previous studies have shown that school demographics have a differential effect on individual student achievement in reading and mathematics (Betts, Zau, & Rice, 2003; Casserly & Council of the Great City Schools, 2001; Sutton & Soderstrom, 1999; Vanneman, 1998b); however, that trend does not extend to aggregate school scores, according to the current study.

### *Effects on Accountability*

The compelling results of this study indicate that demographic factors account for the major portion of the variance on school performance scores by which schools are held accountable. On the Academic Index, which constitutes the largest part of the Accountability Index (see Chapter III for details), school composition variables account for 60-74% of the variance at elementary school level and for 86-89% of the variance at the middle and high school levels. Although progress has undoubtedly been made in Kentucky educational achievement levels (Petrosko, 2000; Poggio, 2000), these findings indicate Kentucky education reform and the accountability system have not yet succeeded in eliminating the inequities among schools, as mandated by the Kentucky Supreme Court ruling in 1989.

The requirement under CATS that all schools reach Proficient by the year 2014 (Foster, 1999) is based on the assumption that having the same goal for all schools will result in educators making the changes to enable schools to reach the goal. When 86-89% of the variance among schools is accounted for by demographic school composition factors, however, it is evident that the changes in instruction have not overcome the impact of preexisting demographics of the schools. Further, instruction

itself is undoubtedly affected by school composition factors (Brookover et al., 1979). These results demonstrate that despite the dramatic improvements in Kentucky achievement results overall since KERA was enacted, the CATS accountability system (including the extensive professional development for standards-based instruction) has not been successful in reforming the relationship between demographics and school performance.

The effects of socio-demographic factors on the accountability scores in this study may have been compounded by a related problem. An examination between the aggregate CTBS NCE score range and the NRT Index range (Appendix E) suggests the differences among schools were magnified by the conversion of test scores into indices since the  $R^2$  represents the proportion of the variance in the dependent variable (scores) that is predicted by the independent variables (school composition). The NRT Index is calculated from the CTBS Reading and Mathematics scores--first into a complete battery and then into an index by a formula based on performance standards (explained in Chapter III). In light of that direct connection, one would expect the descriptive statistics for the CTBS test and the NRT Index to be similar, but they were not. The NRT Index range, representing the variability among the schools, is much wider than the aggregate CTBS Reading and Mathematics scores. Albeit the NRT Index scale at 140 is wider than the CTBS at 100, the differences in the ranges are much greater than that. At the middle and high school levels, the range of the NRT Index is more than twice as large as that of the CTBS.

Because the KCCT test scale is 475 and the Academic Index scale is only 140, it is not possible to ascertain if the same differences exist between the KCCT tests and the

Academic Index without further investigation, but the evidence from the CTBS and NRT Index scores suggests a follow-up exploration. The computation of the final Accountability Index includes the NRT Index (5%) and the Academic Index (85.5-90.25%). If a parallel inflation of the differences in the Academic Index scores and the KCCT Test do occur, the combination of the inflation of the two indices in their computation could contribute to the difficulties for disadvantaged schools in reaching their growth goals since it would exaggerate the differences between schools.

What is known is that the conversion of the CTBS scores inflates the ranges as compared to the reading and mathematics achievement results, accentuating the differences among schools--differences that are largely accounted for by factors that were determined by socioeconomic class position. Effectively, small differences are exaggerated, which may further demean and stigmatize schools that are struggling to catch up with socially advantaged schools.

Roeder's work (2001a) provides another perspective on the problem. From his analysis of the trends in accountability, he projects that less than half of all Kentucky schools and one third of Kentucky urban schools (Jefferson and Fayette Counties) can be expected to reach the minimum goal of 100 by 2014. If the accountability system imposes impossible goals and is frustrating rather than motivating to lower-scoring schools, it is functioning contrary to its espoused purposes.

In summary, at the school-level, SES, two-parent family structure, and Mobility Rate share much of their variance and account for a large proportion of the variability among school scores. These factors are not simply aggregates of the effects of background factors at the student level but represent the effects of concentration of

disadvantaged students in some schools. Thus the influence of the demographic factors was much stronger at the school level than at the student level. The effects of demographic school-composition factors were greater at middle and high school than at elementary and greater for the NRT Index than for the tests from which the Index is derived. The percentage of Black students in a school population was only influential on the middle and high school CTBS scores and the performance indices. All of these results can affect the classification of schools for accountability.

### Recommendations

The school-level demographic factors and the student-level background factors that had substantial effects on scores in this study are indicators of the larger socioeconomic stratification system in the United States. Recommendations for changes in policy offered below are based on the position that one purpose of American schooling in general and KERA in particular is to provide an equitable education for all students regardless of their position in the economic hierarchy, rather than establishing an accountability system that maintains and legitimates that preexisting hierarchy.

### *Policy Implications*

The findings in this study expose the necessity for policy changes in regard to instruction and resource allocation, school assignment, and accountability if education is truly to be equitable in Kentucky.

1. These findings, along with many others, indicate the Kentucky reform with its accountability component has not yet eliminated the effects of demographic factors on student achievement and that additional or alternative resources and approaches are needed. Recommendations for needed changes in teaching methods, resources, or

professional development are beyond the scope of this study, but promising strategies-- e.g., smaller class size in the early grades (Finn, 2002), school effectiveness research (Levine & Lezotte, 1990), culturally responsive instruction (Tharp & Gallimore, 1988), and early intervention--suggest that the gap can be closed (Muñoz, 2001; Thomas & Bainbridge, 2001). However, wide spread implementation of any of these strategies will be expensive. Smaller class sizes and early interventions require additional personnel as well as facilities. Curricular and instructional approaches just mentioned (as well as others) will likely lead to improvement, but they will not take place without extensive professional development to alter existing ineffective modes of instruction and to create school cultures more conducive to equitable achievement.

2. Regardless of the causes, the effects of the concentrations of disadvantaged students (as indicated in this study by percentages of free and reduced lunch, non two-parent families, or highly mobile students) on school performance are undeniable. Therefore, school policies that affect school composition should be analyzed carefully and adjusted to avoid high concentrations of disadvantaged students in any schools. Instead a school district should take positive steps to insure economic diversity among the student bodies of all schools. For example:

- a. School attendance zones should be drawn in ways to promote socioeconomic equity among schools as much as possible.
- b. In a choice system, guidelines for admission of students to magnets or other special programs should include provisions for economic diversity such as requirements for a certain percentage of students to be on the free and reduced lunch program.



- c. Magnet schools should be located in areas that will encourage economic diversity in the schools--usually in low-income areas--so that low-income families will have the same access as wealthier families.
- d. Since the concentration of disadvantaged students is more strongly related to school performance at middle and high school, special efforts must be made to create economic diversity at those levels.

3. This research supports exploration of alternative systems of accountability that will both accommodate the differences students bring to school and emphasize the continual improvement of schooling for all children. Examples of improvements include the following:

- a. Eliminate the policy of having the same goal for all schools; and replace it with a policy that requires reasonable, achievable goals that rise as the school performance improves.
- b. When setting school goals, take the total context of a school into consideration, including the demographics of the school and the resources that are required to enable a school to reach its goal.
- c. Base accountability on longitudinal measures of individual student achievement or on longitudinal comparisons of school scores taking demographics and school population changes into account, rather than reliance on changes in cross-sectional school-level scores.
- d. Remove CTBS test results from the accountability calculations. Use only tests that are aligned with academic expectations for accountability. Norm-referenced tests may appropriately be used for comparative purposes (in

recognition of the political pressures to confirm that Kentucky students are achieving at high levels relative to students in other states) without including such data in the accountability formula.

- e. Examine the calculations used for the accountability indices to determine whether they inflate the differences among schools from the achievement tests on which they are based and exaggerate the differences related to demographic factors.

4. Adopt an accountability system that uses multi-level analysis to distinguish class and school effects on student achievement as proposed by Lee and Coladarci (2001) and utilized by Sanders (1998b) to measure the value added to individual student achievement by the process of schooling.

#### *Future Studies*

The suggestions for research described below serve two general purposes. The first is to explicate the relationships among socio-demographic factors, student achievement, and school performance, with an ultimate goal of discerning causes, functional pathways, and remedies for barriers to learning. The second is to find more equitable and effective means of establishing and implementing educational accountability--if such is possible.

#### *To Explicate Relationships*

1. This study analyzed the scores for only one year. A similar analysis of several years would give a more complete assessment of the consistency of the effects of socio-demographic factors and determine if there are changes occurring from year to year that indicate trends in the progress of education reform in Kentucky. A follow-up

study similar to this one but examining changes in the regression equations of background variables on student scores over the five years the CATS tests have been in place would provide evidence of the capacity of teaching and assessment to overcome the effects of background factors. If standards-based teaching and assessment overcome effects of background factors, one would expect a decreasing influence of student background factors on CATS scores (decreasing regression coefficients) for each succeeding year as teachers instruct better and students learn better the standards-based curriculum. Further verification of the results could be investigated by applying the analysis to the KIRIS test scores when they were used in the 1990s, which could indicate if the effects of demographic variables on student achievement have waned during this period of educational reform in Kentucky.

2. At the student level, SES (free and reduced lunch eligibility) was found to be a significant predictor of student achievement and the effects of being identified as gifted or ECE seemed to be mostly independent of SES. However, the measurement of SES (whether students were eligible for the free and reduced meals program) was both crude and truncated, including only those near the poverty end of the economic spectrum. A study with more reliable measures of family income (census data or parent reports) that differentiated along the whole economic spectrum could reveal any correlations with gifted and ECE identifications, mobility, and race among the wide range of incomes above and below poverty level.

3. Gifted and ECE classifications of students are proxies for students' previous achievement or abilities. A study that controlled for previous CATS scores would address whether those students are learning at the same rate as other regular education

students or not. It could more directly address the value-added question about individual student achievement and reveal if the tests are just measuring previous knowledge and skills.

4. Although unique and important, the findings in this research were tentative at the school level because of the limited number of cases in the sample. The findings need to be tested with a larger sample of Kentucky schools.

5. At the school level, family structure (percentage of two-parent families) and family income (percentage of students not on the free and reduced lunch program) were extremely overlapping, redundant measures. Mobility Rate also showed some overlap with those variables. To shed light on these complex relationships among SES, family structure, mobility, and student achievement, further studies are needed to delineate concentration effects and to identify more accurately the proportions of the school scores that are accounted for by each of these factors.

One approach to discriminating among predictive factors (using a much larger sample of schools) would be to cluster schools by percent of students on free and reduced meals and then to compare the effects of single or two parent families and mobility within the clusters.

6. It would also seem useful, if practically possible, to delineate better among the types of families that have been lumped together in the JCPS data as non two-parent families, especially in trying to determine causative relationships. The culture of a school with high proportions of divorced and remarried mature adults might be very different from the culture of a school with high proportions of never-married, young mothers. And student achievement might be differently affected by the different

cultures. Yet the JCPS records collapse these and other combinations such as children being raised by relatives, step parents, or foster parents into the non two-parent category versus those children from an intact family, two-parent structure. Finding valid data for such a study would be a challenge, especially on a large scale. A researcher would probably have to depend on surveys (self-reported) and U.S. Census data (also self-reported).

7. The finding that the percentage of Black students composing a school has a positive regression coefficient deserves further investigation to ascertain if this was a true effect or a result of the multicollinearity with poverty and family structure. An approach that clusters schools of similar SES proportions and then regresses race against scores could demonstrate if the percentage of Black was a positive or negative predictor. Questions suggested by the positive coefficients for Black are: Do the scores of students in a school actually improve with a higher proportion of Black students in the school? If so, which students--White, Black, or both--have the higher scores? Is there an optimal range of percentage Black students which contributes to increased achievement?

8. The conflicting findings among studies in JCPS, Maine (Lee, 2002), Louisiana (Bankston & Caldas, 1998), and Washington State (Abbot & Joireman, 2001) suggest the importance of racial context. Comparative studies of the effects of race in such different racial environments would help educators to understand the complexity of the cultural and educational implications of race. For example, in Louisiana most Black students attend predominantly Black schools whereas in Jefferson County, no high school had greater than 71% Black students and most were less than 50%. Are the

findings from this current study limited to districts in which desegregation has eliminated the extreme racial segregation common in most urban areas and some states?

9. The identification of causal factors and pathways in the connections among SES, race, mobility, gender, ECE, giftedness, and family structure can only be suggested by quantitative studies such as this one. To further explicate the tentative relationships suggested by quantitative data, qualitative research that explores the attitudes, beliefs, and behaviors that function to connect those demographic factors with student achievement is necessary.

10. The many contributing and intervening factors that may connect school poverty level with student achievement have been and are being investigated by many researchers in many forms. JCPS offers a unique and useful site for research regarding urban poverty and its relationship to achievement because of its demographics and diversity. A more powerful approach to determine the connections between SES and student achievement would be to compare the extent of school composition and student background effects among schools in JCPS and in Kentucky using a multi-level analysis system such as hierarchical linear modeling to distinguish school effects on student achievement from student background factors. The current study provides preliminary data for constructing such research.

#### *To Improve Accountability*

Further research is needed to determine more accurately the effects of demographic factors on accountability in Kentucky. A first objective would be to untangle the relationships among family structure, mobility, race, and poverty that were concealed by the multicollinearity of those variables. Allison (1999) suggests several

solutions to the problems of multicollinearity. Those suggestions that are most applicable to this study are to (a) reduce the number of variables by deleting one or more, (b) reduce the number of variables by combining collinear variables into an index, and (c) increase the sample size, which will reduce the inflated standard errors. Another approach suggested by Stevens (2002) is first to perform a principal components analysis to identify major constructs that are then used in the regression.

1. Analysis of the effects of SES and family structure over several years on each test would perhaps demonstrate a trend in their relative contributions to school scores at different grades. If trends indicate one is the better predictor, the other could be eliminated and the number of variables could be reduced, which would ameliorate the multicollinearity that was present in this study.

2. Regression of test results over several years could provide data to reduce the number of variables by combining collinear variables into an index that would more comprehensively capture the effects on school scores.

3. At the school level, where there was high correlation among the independent variables, a principal components analysis could be used to identify major constructs as suggested by Stevens (2002).

4. Accountability systems that track individual student improvement from year to year (longitudinal approaches, Linn, 2000) should be explored to gauge more accurately both student learning and school performance. The affective results of such systems should also be examined to confirm whether they provide positive motivations for students and teachers.

5. In addition to using hierarchical linear modeling (HLM) or other multilevel

analysis for research purposes, such an approach could be used to provide more informative and equitable means of measuring performance for accountability. HLM can distinguish between class and school effects on student achievement, which could be used to identify schools' successes in overcoming the negative effects of salient school composition variables on student achievement.

6. As Lee and Coladarci (2001) have suggested, an estimation of true school effects would require a very explicit regression or HLM model that included additional school input variables including school practice and context. Such a comprehensive model that also includes background factors and prior achievement should be explored for use in creating a model that gives the most accurate estimate of the factors affecting school scores and best identifies value-added student achievement.

7. The exaggeration of the achievement differences between schools that occurs because of the conversion of the CTBS NCE (interval data) scores to performance indices has been introduced in this study. That topic needs careful analysis, followed by the necessary changes to eliminate that detrimental side effect of the accountability calculations.

### Conclusions

This study was unique in comparing the effects of socio-demographic variables on test scores and accountability indices in the Kentucky accountability system, CATS, (a) among seven socio-demographic factors, (b) at the student level and at the school level, (c) in the areas of reading and mathematics (d) across grades 3-11, and (e) on norm-referenced and criterion-referenced tests.

At the student level, the socio-demographic variables represented the



background of the students, which are generally considered unalterable variables: SES (income based on free and reduced lunch eligibility), race (Black), gender (female), family structure (two parents or not), disability (ECE), and gifted classifications of the students. At the school level, the socio-demographic variables were the percentages of students characterized by each of those individual background variables, thereby describing aspects of the composition of the schools. The multiple regression procedures applied at both student and school levels yielded results with important implications.

The results of this study provide powerful evidence that socio-demographic factors have substantial effects at the student and school levels. There were important differences and similarities in the effects of background on the CATS tests scores at the student level compared to the effects of school composition on aggregate scores at the school level.

1. *The effects of demographic variables on test scores at the school level were much stronger and more influential than at the student level.* At the student level, the effects sizes ranged from .15 to .36. But at the school level, the effect sizes on the aggregate tests ranged from .56 to .91. Although socio-demographic variables account for a relatively small portion of the variability in student scores, nevertheless they can make an important difference in the categorization of the students' performance levels.

The lower effect sizes at student level are reassuring in that most of the variance in student scores is not due to the personal socio-demographic background factors included in this study. Therefore, 60-85% of the variance in scores is unexplained. At least part of that unexplained variance may be open to influence by teaching and

learning that takes place in school, although other factors such as student aptitude may also play a part. But the greater effects of socio-demographic school composition on the aggregated school scores challenges that optimism in that schools with high concentrations of disadvantaged students have up to 91% of the variance in the school scores predicted by school composition. Although the variance among schools was much less than that among students, the effects were still sizeable (Appendices C and E).

The influence of socio-demographic factors is even more pronounced on the school accountability indices for the norm-referenced NRT Index. The differences in school aggregate scores, which are influenced to a great proportion by socio-demographic factors, are inflated even more by the conversion of the CTBS NCE scores into the NRT Index (Appendix E).

*2. Different variables were influential in predicting the scores at the different levels.* At the student level, SES, gifted and ECE were the most influential factors. SES alone explained 5% to 18% of the variance. Gifted and ECE are categories based on students' previous achievement (as measured by various tests and other observations). Not surprisingly, as proxies for previous achievement and ability, those factors account for a substantial portion of the variance in student scores.

SES as measured by the percent of students not on free and reduced lunches in the school population (%HiSES) was also most influential at the school level. The other influential factors were family structure proportions (%Two Parents) and Mobility Rate, both of which had little effect at the student level. Analysis of the collinearity data indicated those three variables (%HiSES, %Two Parents, and Mobility Rate) are

overlapping expressions of the underlying socioeconomic status of students who compose the school. Thus, the pervasiveness of economic advantage and disadvantage again was shown to be the determining factor for successful schools (cf. Bowles & Gintis, 1976; Coleman et al., 1966; L. S. Miller, 1995; Perkinson, 1977; Persell, 1977; Tyack, 1974; Wilson, 1987).

3. *Demographic factors predicted stronger for mathematics than for reading test scores at the individual level, but there was no difference between the effects of demographic factors on the aggregate reading and mathematics scores.* The differences in the effects of different demographic variables on the reading and mathematics scores at the student level--greater impact of SES and ECE on mathematics, positive effect of female on KCCT reading, negative effect of Black on CTBS mathematics--were not as dramatic as those involving SES and previous achievement. Even so they have implications for instruction and for classroom behaviors and experiences in regard to individual equity. These results draw attention to the different experiences males and females/Blacks and Whites have in classrooms. They require continued critical analysis of instruction and changes in teacher preparation and professional development to address gender and race differences in student achievement.

4. *At both the student and school levels, the effects of demographic factors were much stronger at the middle and high school levels than at elementary.* Effects of background factors (primarily SES, giftedness, and ECE) on student scores were stronger in the higher grades, resulting in a wider range of achievement among students. That trend contradicts the American premise that schooling provides equal educational

opportunity for students regardless of their backgrounds and that the more schooling one has, the better the results. Rather this finding supports the interpretation that the function of schools is to stratify student achievement and subsequent economic positions in the larger society (Bowles & Gintis, 1976; Meyer, 1977). Further, it suggests that the stratification is more successful or pervasive with older students or the longer students are in school.

5. *The effects of demographic factors were different for the criterion-referenced KCCT and the norm-referenced CTBS.* There were two credible findings in this study regarding differences in the effects of demographic factors. The first is that at both the student level and school level, Black is an important predictor on the CTBS tests but not on the KCCT. Being Black had a negative effect on CTBS scores at the individual student level, thus in regard to race, those scores were more influenced by background. This finding supports the position that criterion-referenced testing is a more valid measure of student learning than norm-referenced testing, which is more influenced by background factors and measures skills and knowledge not as affected by the classroom curriculum (cf. Popham, 1981). In effect, Black students were penalized by the CTBS. This finding challenges the equity of using a test on which race has a significant impact.

The second extant finding was that the range of the NRT Index was much greater than that of the CTBS tests from which it was calculated. This finding makes the calculation of the NRT Index questionable and invites exploration of the same phenomenon with the Academic Index.

### *Improving Instruction*

This study revealed the important role socio-demographic factors play in student

achievement and school performance. It was not the background of the students themselves that was the strongest factor predicting school level achievement; it was the composition of the school--the impact of concentrating students from poverty backgrounds into the same buildings. The precise mechanisms through which these concentration effects operate are beyond the scope of this study. Clearly the school learning climate (Brookover et al., 1979), peer effects (Cook & Ludwig, 1998), teacher attitudes (Solomon et al., 1996), instructional practices (Ennis, 2002), and community forces (Ogbu & Simons, 1998) are among the issues that must be sorted out.

These results do not imply that all students cannot learn or that the achievement gap cannot be closed. In fact, the few success stories for high-poverty schools indicate it is possible but requires extraordinary measures or personnel (Kannapel et al., 2001; Solomon et al., 1996). Rather, unless the limitations of the current system are acknowledged and alternative or supplementary actions implemented, widespread reduction in the achievement gap for haves versus have-nots is highly unlikely. Promising strategies suggest that the gap can be closed if resources are targeted effectively (Finn, 2002; Levine & Lezotte, 1990; Tharp & Gallimore, 1988). If Kentucky--or any state--is truly committed to an equitable education for all children, it must provide the funding that will in turn provide the resources necessary to eliminate the effects of economic disparities in education.

These findings, along with many others, clearly indicate that while overall great strides have been made in Kentucky, education reform has not eliminated the effects of socio-demographic factors on student achievement nor on school performance.

Although a few high-poverty schools have made great strides, the gains are not systemic

enough to overcome the historic differences between advantaged and disadvantaged schools. The equity goals of KERA remain elusive (Roeder, 2000). The premise that holding teachers responsible would ensure that all students perform at high levels has not been sufficient to eliminate inequality. New, additional, or different approaches are needed.

Results from studies such as this one make clear the role of schools in exacerbating the effects of social hierarchy in education. Data analysis can foster better understanding of the problems for educators as well as policy makers. Just as this study has revealed the relative contributions of the various demographic factors to student achievement versus school performance, school practitioners could apply regression findings to better understand the dynamics of social class in their own schools. For example, a better understanding of social class on teacher and school performance could combat the common tendency to attribute students' low achievement to family background, when the data demonstrate that school experiences are as much responsible.

Policy makers and educators must discuss these findings and issues openly and honestly if progress is to be made. The realization of productive instructional practices and allocation of resources will require more than research; it will require political resolve and governmental support, which can be initiated by acknowledgement of the problem.

#### *Student Assignment to Schools*

The implications from this study with regard to student assignment to schools are persuasive. The study shows that concentrations of non-two-parent and low-income

families in schools have a detrimental effect on school performance that extends beyond the effects of the individual students' backgrounds. Regardless of the causes of these relationships and the pathways that produce them--whether peer interactions, adult relationships with students, or other aspects of school climate--the concentration effects are undeniable. Therefore, it is imperative that school districts do everything in their power to disperse low-income students equally across schools rather than allowing high concentrations in some schools.

Busing for socioeconomic integration, of course, has political consequences that make it an unpopular solution. Many school districts exacerbate the problem with policies that increase economic isolation and segregation (e.g., ability grouping, tracking within and between schools). Other practices such as magnet schools, choice programs, and gerrymandering school attendance lines may contribute to further segregation or encourage socioeconomic desegregation depending on the ways they are designed and implemented. In a district such as JCPS that assigns students to schools by use of magnets and choice, there are several steps (other than busing) that can be taken to relieve socioeconomic stratification of the schools, including changes in the selection guidelines for choice and magnet schools.

The greater effects of demographic variables at the middle and high school levels especially direct attention to those sites. For example, district leaders could search for evidence of systemic policies or practices that contribute to socioeconomic segregation or high mobility in certain schools. Whatever the causative agents, altering the school populations to reflect the most favorable proportions of students by whatever means are politically feasible can be helpful in overcoming the negative effects of such

school composition factors as high poverty or high mobility.

### *Improving Accountability Systems*

This study demonstrates that because the current Kentucky accountability system does not take demographic factors into consideration, it is both unfair and unfulfilled vis-à-vis its purpose of improving educational equity. Kentucky holds schools accountable for improving their school scores each biennium. A regression line (the improvement line) is drawn from each school's initial biennium score to the goal of 100 in the year 2014. That line establishes the expectations for improvement for each school for each biennium. The slope of the improvement lines of initially lower performing schools is much steeper than that of initially high scoring schools. It is possibly made even steeper by the calculations used by the Commonwealth to convert the aggregate test scores to performance indices, which exacerbate the differences among schools.

The current study, Roeder's (2001a) projections, and other researchers (Ennis, 2002; Guskey, 1997; Luvisi, 2000; Muñoz & Dossett, 2001; Pitts & Reeves, 1999) have documented the effects of the various aspects of SES on school scores and the apparent inability of the vast majority of schools with high concentrations of disadvantaged populations to overcome those effects. With close to 90% of the variance at middle and high schools explained by socio-demographic variables, disadvantaged schools have virtually no hope for improving sufficiently to meet long-term improvement goals in the current conditions. Consequently, the most disadvantaged schools are presented with the impossible task of catching up with advantaged schools by 2014 (11 years from the time of this writing), whereas, a few low-poverty schools are hardly challenged to



improve, as they are close to the goal of 100 already. Instead of ensuring equity by mandating the same goals for all schools, the CATS system emphasizes the failure of high-poverty schools to meet those goals. The disadvantaged schools are demoralized, and the most advantaged schools are not much challenged. These studies provide strong arguments for the need for changes in the Kentucky accountability system that will avoid penalizing high-poverty schools and that will more equitably spur school improvement.

This is not a call for a return to the pre-KERA era. At that time, Kentucky was near the bottom of the 50 states on every outcome measured. The system represented a logic of confidence (Meyer & Rowan, 1978) in which schools were commonly thought to be good or bad depending on the mix of family backgrounds in the student composition (S. K. Miller, 1992). The lack of accountability and loose coupling of purpose and functions resulted in stasis with the poor and minorities permanently on the bottom of the hierarchy. A return to ignorance of student achievement outcomes cannot be expected to prompt any more attention to the actual improvement of education for disadvantaged students now than occurred in the past.

The alternative is to revise the accountability system so that both excellence and equity are expected and can be achieved. From the beginning KERA has been premised on the belief that all students can learn at high levels, and that to allow at-risk students to have lower goals is tantamount to having lower expectations for them--to formalize the expectation that they are not as capable as their affluent peers. Thus, the inequity of applying different standards, which will forever condemn high-poverty schools to low achievement, must be avoided.

Instead, this research supports exploration of alternative systems of accountability that will both accommodate the differences students bring to school and emphasize the continual improvement of schooling for all children. It is seldom acknowledged that as well as imposing impossible goals on disadvantaged schools, having the same school goal of 100 (on a 140 point scale) by the year 2014 *lowers* expectations for higher-performing schools in terms of value-added improvement. An at-risk school with a current score of 50 is expected to increase 50 points over the next 10 years, while an affluent school with a current score of 90 is expected to improve only by 10. High-achieving as well as under-achieving students are entitled to expectations that will stretch their capacities and expand their knowledge and skills.

Therefore, the current CATS accountability system does not provide optimal motivation to either the advantaged or the disadvantaged schools. An effective accountability system must set high but achievable standards for both students and teachers and then provide both guidance and resources for needed improvements. As Wheelock (2000) recommends, teachers must have effective incentives to change practices.

The strong relationships of socio-demographic factors to school performance recommend a system that focuses on how much students have improved rather than where the school is along the line toward 100. Even though Kentucky has a value-added system, the imposition of a common goal for all schools with a steeper improvement line slopes for the at-risk schools emphasizes the relative achievement of different schools and fails to provide schools with the tools for overcoming inequities.

Linn (2001) recommends tracking the longitudinal improvement of individual

student achievement from year to year but recognizes that is difficult to do and may exclude mobile students. However, that is the system developed by Sanders (1998b) and used in Tennessee. Alternatively, Linn recommends a quasi-longitudinal system in which all students in a grade are compared with all the children in the next grade the next year to determine how much that cohort of students have improved. Such a system would require much additional testing or a reduction in the subjects that are tested, since all students would have to be tested in all subjects every year.

Several authors (Ennis, 2002; Guskey, 1997; Luvisi, 2000) have found that change scores, the difference in school scores from year to year, were not correlated with socio-demographic variables; however, when Roeder (1999) included previous scores in the regression models, he found school poverty had a negative effect and that the change in accountability scores got smaller over time in high-poverty schools. A focus on change scores, rather than on progress on an improvement line, might be a reasonable alternative to the current system, but previous scores must be included in the calculations so that the effects of demographics are not undetected.

Alternatively, an accountability system that uses multi-level analysis to distinguish class and school effects on student achievement (as proposed by Lee and Coladarci, 2001; and by Sanders, 2001) could be helpful in measuring value-added student achievement. Lee and Coladarci suggested that other background factors, such as prior achievement and mobility, be included in the regression in order to create a comprehensive model.

The selection of tests that are aligned with academic standards is also essential for valid measures of student achievement and school improvement. This study has

revealed that race is a significant factor affecting the CTBS scores. Also, the CTBS is not aligned with the Kentucky standards and does not encourage the teaching of higher level thinking skills (Popham, 1981). Fortunately, the CTBS only counts for 5% in the school accountability formulas; thus it is not a major detriment to the Accountability Index. However, it receives much attention from policy-makers and the media. Policy makers need to examine critically their purposes in including the CTBS test in an accountability system. This issue is important for consideration of tests that will be used to implement the No Child Left Behind Act in light of the serious consequences it entails for schools and school districts

The current system in which the state does not take demographics into consideration in its accountability system is patently unfair. Unless the state is willing to make changes that lead to challenging, equitable, and achievable standards, education will continue to serve as a means of social control (Bowles & Gintis, 1976).

Legitimation of social stratification (Della Fave, 1980) will have been once again confirmed. As Sutton and Soderstrom (1999) have pointed out, implications of using measures that do not take different conditions into account result in reinforcement of societal and personal prejudices against low-performing schools, their students, and teachers.

In the case of Kentucky, the CATS system of goals and regulations that determine success and failure currently serve as mechanisms for emphasizing the apparent inadequacy of high-poverty schools. By labeling these schools as failures, CATS demeans and devalues the education of low-income students, thereby rationalizing the advantages and economic capital accruing to the affluent public and

private schools. It solidifies the expectations and obligations of teachers and students in the system according to the place the testing system assigns them in the social class hierarchy (cf. Dika & Singh, 2002).

Kentucky instituted an extensive educational reform with KERA, perhaps the most comprehensive among the entire 50 states. The philosophy and major tenets regarding high standards and additional resources that will prepare students for all aspects of life are sound. The accountability provisions promised a higher level of education for all children in the state--that children would no longer be allowed to pass through 12 years of schooling and graduate without the level of skills and knowledge needed in today's society. But within the current accountability system there are contradictions that undercut the fulfillment of KERA's promise.

The Kentucky educational accountability system needs both conceptual and programmatic revisions so that it will serve the purpose for which it was intended--the improvement of teaching and learning for all students. To accomplish those purposes, the accountability system must be equitable for students and schools, teachers must be held accountable only for student learning over which they have some control, and the achievement goals must be attainable and motivating. The evidence from this study, as well as from the broader research literature as well, provide the information to initiate the needed changes in the accountability system. It is the responsibility of educators, the media, and the public to demand changes, and for policy-makers to implement those changes in order to continue the significant progress that has been achieved under KERA. Only then will all of Kentucky's children have the opportunity to achieve excellence no matter their family background.

## REFERENCES

- Abbot, M. L., & Joireman, J. (2001). *The relationships among achievement, low income, and ethnicity across six groups of Washington State students* (No. WSRC-TR-1). Lynnwood, WA: Washington School Research Center. (ERIC Document Reproduction Service No. ED454356)
- Airasian, P. W. (1997). Oregon Teacher Work Sample Methodology: Potential and problems. In J. Millman (Ed.), *Grading teachers, grading schools. Is student achievement a valid evaluation measure?* (pp. 46-52). Thousand Oaks, CA: Corwin Press.
- Allison, P. D. (1999). *Multiple regression: A primer*. Thousand Oaks, CA: Pine Forge Press.
- Alston, E., Brinly, B., Carr, A., Deaton, S., Dutton, P., Little, D., et al. (1999). *Kentucky Education Reform Act: A citizen's handbook*. Frankfort, KY: Legislative Research Commission.
- Anyon, J. (1981). Social class and school knowledge. *Curriculum Inquiry*, 11, 3-42.
- Appalachia Educational Laboratory. (1994). Instruction and assessment in accountable and nonaccountable grades. *Notes from the Field: Education Reform in Rural Kentucky*, 4(1).
- Apple, M. W. (1985). *Education and power*. Massachusetts: Routledge & Kegan Paul.

- Babbie, E. (1986). *The practice of social research* (4th ed.). Belmont, CA: Wadsworth Publishing.
- Bankston, C. L., & Caldas, S. J. (1998). Family structure, schoolmates, and racial inequalities in school achievement. *Journal of Marriage and the Family*, 60, 715-723.
- Best, J. W., & Kahn, J. V. (1989). *Research in education* (Sixth ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Betts, J. R., Zau, A. C., & Rice, L. A. (2003). *Determinants of student achievement: New evidence from San Diego* (No. LC205.5.C2B48). San Francisco: Public Policy Institute of California.
- Bloom, B. S. (1974). Time and learning. *American Psychologist*, 29, 682-688.
- Bloom, B. S. (1980). The new direction in educational research: Alterable variables. *Phi Delta Kappan*, 61, 382-385.
- Bolon, C. (2000). School-based standard testing, *Education Policy Analysis Archives* (Vol. 8).
- Borland, M. V., & Howsen, R. M. (1999). A note on student academic performance. *The American Journal of Economics and Sociology*, 58, 537-546.
- Bowles, S., & Gintis, H. (1976). *Schooling in capitalist America: Educational reform and the contradictions of economic life*. New York: Basic Books.
- Boysen, T. C. (1992). Irreconcilable differences: Effective urban schools versus restructuring. *Education and Urban Society*, 25, 85-95.

- Braddock, J. H., II, & Slavin, R. E. (1993). Why ability grouping must end: Achieving excellence and equity in American education. *Journal of Intergroup Relations*, 20, 51-64.
- Brookover, W., Beady, C., Flood, P., Schweitzer, J., & Wisenbaker, J. (1979). *School social systems and student achievement: Schools can make a difference*. New York: Praeger Publishers.
- Caldas, S. J., & Bankston, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research*, 90, 269-278.
- California Food Policy Advocates. (n.d.). *California National School Lunch Program facts*. Retrieved August 18, 2003, from [http://www.cfpa.net/School\\_Food/Lunch/School%20Lunch.htm](http://www.cfpa.net/School_Food/Lunch/School%20Lunch.htm)
- Calvert, J., Gaus, D., & Ruscoe, G. (2000). The tortuous journey toward school reform. In B. L. Whitford & K. Jones (Eds.), *Accountability, assessment, and teacher commitment* (pp. 109-125). Albany, NY: State University of New York.
- Casserly, M., & Council of the Great City Schools. (2001). *Beating the odds: A city-by-city analysis of student performance and achievement gaps on state assessments*. Washington, DC: Council of the Great City Schools. (ERIC Document Reproduction Service No. ED459280)
- Clements, S. K. (2000). Linking curriculum and instruction to performance standards. In R. S. Pankratz & J. M. Petrosko (Eds.), *All children can learn: Lessons from the Kentucky reform experience* (pp. 98-115). San Francisco: Jossey-Bass.



- Clinchy, B. M. (1997). The standardization of the student. In E. Clinchy (Ed.), *Transforming public education: A new course for America's future* (pp. 66-78). New York: Teachers College Press.
- Cohen, D. K. (1995). What is the system in systemic reform? *Educational Researcher*, 24(9), 11-17, 31.
- Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of educational opportunity*. Washington, DC: National Center for Educational Statistics.
- Coley, R. J. (2001). *Differences in the gender gap: Comparisons across racial/ethnic groups in education and work*. Princeton, NJ: Educational Testing Service, Policy Information Center. (ERIC Document Reproduction Service No. ED451222)
- Cook, P. J., & Ludwig, J. (1998). The burden of "Acting White": Do Black adolescents disparage academic achievement? In C. Jencks & M. Phillips (Eds.), *The black-white test score gap* (pp. 375-400). Washington, DC: Brookings Institution Press.
- Cuban, L. (1990). Reforming again, again, and again. *Educational Researcher*, 19(1), 3-13.

- Darling-Hammond, L. (1994). Performance-based assessment and educational equity. *Harvard Educational Review, 64*, 5-30.
- David, J. (1999). *Creating successful schools: A continuous commitment* [Electronic version]. Louisville: Partnership for Kentucky Schools and Prichard Committee for Academic Excellence.
- Della Fave, L. R. (1980). The meek shall not inherit the earth: Self-evaluation and the legitimacy of stratification. *American Sociological Review, 45*, 955-971.
- Delpit, L. (1995). *Other people's children: Cultural conflict in the classroom*. New York: New Press.
- Dika, S. L., & Singh, K. (2002). Applications of social capital in educational literature: A critical synthesis. *Review of Educational Research, 72*, 31-60.
- Education Trust. (2001). *State summary of Kentucky*. *Ed Watch Online*. Washington, DC: Education Trust. (ERIC Document Reproduction Service No. ED459238)
- Elmore, R. F., & Fuhrman, S. H. (2001). Holding schools accountable: Is it working? *Phi Delta Kappan, 63*, 67-72.
- Ennis, L. S. (2002). *Effects of instructional strategies in seventh grade science achievement as perceived by Kentucky students*. Unpublished doctoral dissertation, University of Louisville.
- Entwisle, D. R., & Alexander, K. L. (1995). A parent's economic shadow: Family structure versus family resources as influences on early school achievement. *Journal of Marriage and the Family, 57*, 399-409.

- Evans, M. (2001). *Developing and testing an innovation component configuration map for gifted education in the elementary school*. Unpublished doctoral dissertation, University of Louisville, Louisville.
- Fayette County District Report Card, 2000-2001. (2000). Retrieved September 20, 2002, from [http://apps.kde.state.ky.us/report\\_card/](http://apps.kde.state.ky.us/report_card/)
- Ferguson, R. E. (1998). Can schools narrow the black-white test score gap? In C. Jencks & M. Phillips (Eds.), *The black-white test score gap* (pp. 318-374). Washington, DC: Brookings Institution Press.
- Finn, J. D. (2002). Small classes in American schools: Research, practice, and politics. *Phi Delta Kappan*, 83, 551-560.
- Foster, J. D. (1999). *Redesigning public education: The Kentucky experience*. Lexington, KY: Diversified Services Inc.
- Fuhrman, S. H. (2001). Introduction. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states* (pp. 1-12). Chicago: University of Chicago Press.
- Gardner, H. (1985). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Gnadinger, C., McIntyre, E., Chitwood-Smith, T., & Kyle, D. (2000). Primary Program [Electronic version]. In J. Petrosko & J. C. Lindle (Eds.), *2000 review of research on the Kentucky education reform*. Lexington: Kentucky Institute for Education Research.
- Goodlad, J. I. (1984). *A place called school. Prospects for the future*. New York: McGraw-Hill Book Company.

- Guskey, T. R. (1997, March). *The relationship between socioeconomic characteristics and school-level performance assessment results*. Paper presented at the American Educational Research Association Annual Meeting, Chicago.
- Hofferth, S. L. (1987). Implications of family trends for children: A research perspective. *Educational Leadership*, 44(5), 78-84.
- Hoisch, M. C., & Miller, S. K. (2003). Developing a theoretical model of employee relations. *Manuscript submitted for publication*.
- Hopper, J. A. (1997). *The effects of divorce on children: A review of the literature*. Unpublished manuscript, Loyola College in Maryland.
- Hornbeck, D. W. (1990). Recommendations related to curriculum (pp. 66). Frankfort, KY: Legislative Research Commission.
- Huston, A. C., McLoyd, V. C., & Coll, C. G. (1994). Children and poverty: Issues in contemporary research. *Child Development*, 65, 275-282.
- Jefferson County District Report Card, 2000-2001. (2001). Retrieved September 20, 2002, from [http://apps.kde.state.ky.us/report\\_card/](http://apps.kde.state.ky.us/report_card/)
- Jefferson County Public Schools. (1999a). Consolidated plan and funding application, Jefferson County 1998/2000. Frankfort, KY: Kentucky Department of Education.
- Jefferson County Public Schools. (1999b). *High School Principal's Retreat: Lighting the way for instruction*. Louisville: Author.
- Jefferson County Public Schools. (1999c). *The Individual Success Plan: A safety net for student success*. Louisville: Author.

- Jefferson County Public Schools. (2000). *Advance Program and Primary Talent Pool: Guidelines, information, and procedures*. Louisville: Author.
- Jefferson County Public Schools. (2001a). *Elementary schools 2001-2002: Data book*. Louisville: Author.
- Jefferson County Public Schools. (2001b). *High schools 2001-2002: Data book*. Louisville: Author.
- Jefferson County Public Schools. (2001c). *Middle schools 2001-2002: Data book*. Louisville: Author.
- Jefferson County Public Schools. (2002a). *A guide to the Jefferson County Public Schools: Food for thought*. Louisville: Author.
- Jefferson County Public Schools. (2002b). *Minority Student Achievement Initiative report* (pp. 51). Louisville: Author.
- Jencks, C., & Phillips, M. (Eds.). (1998). *The black-white test score gap*. Washington, DC: Brookings Institution Press.
- Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., et al. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books.
- Jennings, T. A., Kovalski, T. M., & Behrens, J. T. (2000, April 24-28). *Predicting academic achievement using archival mobility data*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.
- Jeynes, W. H. (1999). Effects of remarriage following divorce on the academic achievement of children. *Journal of Youth and Adolescence*, 28, 385-393.

- Kannapel, P. J., Aagaard, L., Coe, P., & Reeves, C. A. (2001). What can be said about reform progress. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states* (pp. 242-262). Chicago: University of Chicago Press.
- Kannapel, P. J., & Coe, P. (2000). Improving schools and school leaders. In R. S. Pankratz & J. M. Petrosko (Eds.), *All children can learn: Lessons from the Kentucky reform experience* (pp. 159-176). San Francisco: Jossey-Bass.
- Kannapel, P. J., Coe, P., Aargaard, L., & Moore, B. D. (1996, April). *"I don't give a hoot if somebody is going to pay me \$3,600": Local school district reactions to Kentucky's high stakes accountability program*. Paper presented at the American Educational Research Association Annual Meeting, New York.
- Kentucky Association of School Councils. (2002). *Top 2002 results for African-American students*. Retrieved January 22, 2003, from <http://www.kasc.net/downloads/2002topdisag.pdf>
- Kentucky Board of Education. (2001). *The Kentucky Board of Education strategic plan*. Retrieved June 12, 2002, from [http://www.kde.state.ky.us/commiss/kbe/strategic\\_plan.pdf](http://www.kde.state.ky.us/commiss/kbe/strategic_plan.pdf)
- Kentucky Department of Education. (2001a). *Commonwealth Accountability Testing System spring 2001: Interpretive guide, Kentucky evaluator's edition* [Electronic version]. Frankfort, KY: Author.
- Kentucky Department of Education. (2001b). *2001-2002 District assessment coordinator implementation guide for the Commonwealth Accountability Testing System* [Electronic version]. Frankfort, KY: Author.

Kentucky Department of Education. (2002). *2002 CATS Interpretive Guide, Version 1.02*

Kentucky Department of Education. (2002a). *Kentucky Core Content Tests: 2000 technical report* [Electronic version]. Frankfort, KY: Author.

Kentucky Department of Education. (2002b). *Spring 2002 Kentucky performance report*. Retrieved September 20, 2002, from [http://www.kde.state.ky.us/oaa/implement/0102/KPR02\\_275.pdf](http://www.kde.state.ky.us/oaa/implement/0102/KPR02_275.pdf)

Kulik, J. A. (1993). *An analysis of the research on ability grouping*. Storrs, CT: National Research Center on the Gifted and Talented. (ERIC Document Reproduction Service No. ED367095)

Leder, G. C. (1990). Gender differences in mathematics: An overview. In E. Fennema & G. C. Leder (Eds.), *Mathematics and gender* (pp. 10-26). New York: Teachers College Press.

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational Researcher*, 31(1), 3-12.

Lee, J., & Coladarci, T. (2001, April). *Imperative or choice? Multi-level and multi-measure analysis of student assessment data for evaluation of systemic school reform*. Paper presented at the American Educational Research Association Annual Meeting, Seattle.

Levine, D. U., & Lezotte, L. W. (1990). *Unusually effective schools: A review and analysis of research and practice*. Madison, WI: National Center for Effective Schools Research and Development. (ERIC Document Reproduction Service No. ED330032)

- Linn, R. L. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4-16.
- Linn, R. L. (2001). *Reporting school quality in standards-based accountability systems*. Los Angeles: Center for Research on Evaluation Standards and Student Testing, University of California at Los Angeles. (ERIC Document Reproduction Service No. ED455251)
- Lubienski, S. T. (2001, April). *A second look at mathematics achievement gaps: Intersections of race, class, and gender in NAEP data*. Paper presented at the American Educational Research Association Annual Meeting, Seattle.
- Lubienski, S. T. (2002, April). *Are we achieving "mathematical power for all?" A decade of national data on instruction and achievement*. Paper presented at the American Educational Research Association Annual Meeting, New Orleans.
- Lumsden, L. (2000). *Early intervention to prevent violence*. Oregon: NAESP National Principals Resource Center. (ERIC Document Reproduction Service No. ED448498)
- Lunenburg, F. C. (1992). Introduction: The current educational reform movement--History, progress to date, and the future. *Education and Urban Society*, 25, 3-17.
- Luvisi, C. L. (2000). *Implementing Kentucky's Primary School Program: Its relationship to measures of academic achievement*. Unpublished doctoral dissertation, University of Louisville.
- Madaus, G. F., & Clarke, M. (2001). *The adverse impact of high stakes testing on minority students: Evidence from 100 years of test data*. New York: The Century Foundation. (ERIC Document Reproduction Service No. ED450183)



- Maker, C. J. (1982). *Curriculum development for the gifted*. Rockville, MD: Aspen Systems Corp.
- Mao, M. X., Whitsett, M. D., & Mellor, L. T. (1997, March). *Student mobility, academic performance, and school accountability*. Paper presented at the American Educational Research Association Annual Meeting, Chicago.
- Marzano, R. J., & Costa, A. L. (1988). Question: Do standardized tests measure general cognitive skills? Answer: No [Electronic version]. *Educational Leadership*, 45(8), 66-71.
- Mayeske, G. S., & Beaton, A. E. (1975). *Special studies of our nation's students*. Washington, DC: Office of Education, U.S. Printing Office. (ERIC Document Reproduction Service No. ED118701)
- McDermott, M. (1983, April 21-24). *The impact of classroom interaction patterns on students' achievement-related beliefs and behaviors*. Paper presented at the biennial meeting of the Society for Research in Child Development, Detroit.
- Meyer, J. W. (1977). The effects of education as an institution. *American Journal of Sociology*, 83, 55-77.
- Meyer, J. W., & Rowan, B. (1978). The structure of educational organizations. In M. W. Meyer & Associates (Eds.), *Environments and organizations* (pp. 78-109). San Francisco: Jossey-Bass Publishers.
- Miller, L. S. (1995). *An American imperative: Accelerating minority educational advancement*. New Haven: Yale University.

- Miller, S. K. (1985). Research on exemplary schools: An historical perspective. In G. R. Austin & H. Garber (Eds.), *Research on exemplary schools* (pp. 3-30). Orlando: Academic Press.
- Miller, S. K. (1992). Changing conceptions of good schools: Implications for reforming urban education. *Education and Urban Society*, 25, 71-84.
- Molfese, V. J., DiLalla, L. F., & Bunce, D. (1997). Prediction of the intelligence test scores of 3- to 8-year-old children by home environment, socioeconomic status, and biomedical risks. *Merrill-Palmer Quarterly*, 43, 219-234.
- Morrow-Howell, N. (1994). The M word: Multicollinearity in multiple regression. [Electronic version]. *Social Work Research*, 18, 247-251.
- Muñoz, M. A. (2000). *Mathematics achievement in a reform environment: The effect of teacher, student, and parental characteristics on student testing performance* [Electronic version]. Louisville: Jefferson County Public Schools. (ERIC Document Reproduction Service No. ED463168)
- Muñoz, M. A. (2001). *The critical years of education for at-risk students: The impact of an early childhood program on student learning*. Louisville: Jefferson County Public Schools. (ERIC Document Reproduction Service No. ED456913)
- Muñoz, M. A., & Dossett, D. (2000). *School district equity and accountability: Toward a comprehensive model*. Louisville: Jefferson County Public Schools. (ERIC Document Reproduction Service No. ED467036)
- Muñoz, M. A., & Dossett, D. (2001). Equity and excellence: The effect of school and sociodemographic variables on student achievement. *Journal of School Leadership*, 11, 120-134.

- National Assessment of Educational Progress. (2003, January 3). *The nation's report card, NAEP data*. Retrieved March 4, 2003, from <http://nces.ed.gov/nationsreportcard/naepdata/>
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform. An open letter to the American people. A report to the Nation and the Secretary of Education*. Washington, DC: Superintendent of Documents, Government Printing Office. (ERIC Document Reproduction Service No. ED226006)
- The No Child Left Behind Act of 2001, Executive summary*. (2001). Retrieved Nov. 30, 2002, from <http://www.ed.gov/offices/OESE/esea/exec-summ.html>
- Oakes, J. (1985). *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.
- Oakes, J., & Wells, A. S. (1998). Detracking for high student achievement. *Educational Leadership*, 55(6), 38-41.
- Ogbu, J. U., & Simons, H. D. (1998). Voluntary and involuntary minorities: A cultural-ecological theory of school performance with some implications for education. *Anthropology & Education Quarterly*, 29, 155-188.
- Orfield, G., & Wald, J. (2000). *Testing, testing, 2002*.
- Pankratz, R. S. (2000). The legal and legislative battles. In R. S. Pankratz & J. M. Petrosko (Eds.), *All children can learn: Lessons from the Kentucky reform experience* (pp. 11-28). San Francisco: Jossey-Bass.
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (Third ed.). Fort Worth: Harcourt Brace College Publishers.

- Perkinson, H. J. (1977). *The imperfect panacea: American faith in education, 1865-1976* (2nd ed.). New York: Random House.
- Persell, C. H. (1977). *Education and inequality: The roots and results of stratification in America's schools*. New York: The Free Press.
- Petrosko, J. (2000). Assessment and accountability [Electronic version]. In J. Petrosko & J. C. Lindle (Eds.), *2000 review of research on the Kentucky Education Reform Act* (pp. 1-81). Lexington: Kentucky Institute for Education Research.
- Phillips, M., Brooks-Gunn, J., Duncan, G. J., Klebanov, P., & Crane, J. (1998). Family background, parenting practices, and the black-white test score gap. In C. Jencks & M. Phillips (Eds.), *The black-white test score gap* (pp. 103-145). Washington, DC: Brookings Institution Press.
- Phillips, M., Crouse, J., & Ralph, J. (1998). Does the black-white test score gap widen after children enter school? In C. Jencks & M. Phillips (Eds.), *The black-white test score gap* (pp. 229-272). Washington, DC: Brookings Institution Press.
- Pitts, T. C., & Reeves, E. B. (1999). *A spatial analysis of contextual effects on educational accountability in Kentucky*. Morehead, KY: Morehead State University. (ERIC Document Reproduction Service No. ED433976)
- Poggio, J. P. (2000). Statewide performance assessment and school accountability. In R. S. Pankratz & J. M. Petrosko (Eds.), *All children can learn: Lessons from the Kentucky reform experience* (pp. 75-97). San Francisco: Jossey-Bass.
- Popham, J. W. (1981). *Teacher evaluation--The wrong tests for the right job*. Paper presented at the Annual Meetings of the National Council of Measurement in Education and the American Educational Research Association, Los Angeles.

- Purkey, W. W., & Novak, J. M. (1984). *Inviting school success: A self-concept approach to teaching and learning* (2nd ed.). Belmont, CA: Wadsworth Publishing.
- A report from the Kids Mobility Project.* (1998). Minneapolis: Family Housing Fund. (ERIC Document Reproduction Service No. ED453326)
- Roeder, P. W. (1999). *Education reform and equitable excellence: The Kentucky experiment* [Electronic version]. Lexington, KY: RDS Publishing.
- Roeder, P. W. (2000). *Education reform and equitable excellence: The Kentucky experiment (Update-June 2000), Race, poverty, and urban schools in Kentucky.* Retrieved March 10, 2002, from <http://www.uky.edu/~proeder/keraweb.htm>
- Roeder, P. W. (2001a). *The KERA endgame: Which Kentucky schools will achieve proficiency by 2014?* Retrieved March 10, 2002, from <http://www.uky.edu/~proeder/keraweb.htm>
- Roeder, P. W. (2001b). *A report on KERA Report Cards.* Retrieved March 10, 2002, from <http://www.uky.edu/~proeder/keraweb.htm>
- Roeder, P. W. (2002, May). *Resisting the urge to merge: Does school size matter?* Retrieved August 8, 2002, from <http://www.uky.edu/~proeder/urbschlsz.pdf>
- Rotberg, I. C. (2001). A self-fulfilling prophecy. *Phi Delta Kappan*, 83, 170-171.
- Rothstein, R. (1998). *The way we were? The myths and realities of America's student achievement, Chapter 1.* Retrieved February 20, 2002, from [http://www.tcf.org/Publications/Education/Way\\_We\\_Were/Chapter1.html](http://www.tcf.org/Publications/Education/Way_We_Were/Chapter1.html)
- Sadker, M., & Sadker, D. (1994). *Failing at fairness: How America's schools cheat girls.* Washington, DC: Macmillan Publishing Company.

- Sanders, W. L. (1998a). Value-added assessment: A method for measuring the effects of the system, school and teacher on the rate of student academic progress. *The School Administrator Web Edition*, 11(55), 24-27.
- Sanders, W. L. (1998b). Value-added assessment: A method for measuring the effects of the system, school and teacher on the rate of student academic progress, *The School Administrator Web Edition* (December 1998 ed.).
- Sanders, W. L., Saxton, A. M., & Horn, S. P. (1997). The Tennessee Value-Added Assessment System: A quantitative, outcomes-based approach to educational assessment. In J. Millman (Ed.), *Grading teachers, grading schools. Is student achievement a valid evaluation measure?* (pp. 137-162). Thousand Oaks, CA: Corwin Press.
- Sarason, S. B. (1996). *Revisiting "The culture of the school and the problem of change."* New York: Teachers College Press.
- Schallock, H. D., Schallock, M., & Girod, G. (1997). Teacher Work Sample Methodology as used at Western Oregon State College. In J. Millman (Ed.), *Grading teachers, grading schools. Is student achievement a valid evaluation measure?* (pp. 15-45). Thousand Oaks, CA: Corwin Press.
- Shanker, A. (1995). The case for high stakes and real consequences. In D. Ravitch (Ed.), *Debating the future of American education* (pp. 45-153). Washington, DC: Brookings Institution.
- Shannon, D. M., & Davenport, M. A. (2001). *Using SPSS to solve statistical problems: A self-instruction guide*. Upper Saddle River, NJ: Merrill Prentice Hall.

- Smith, D. C., Neff, D. O., & Nemes, J. M. (1999). *Assessing race and gender subgroup performance differences in KIRIS accountability cycle 2 and cycle 3 results: A report for the Kentucky Department of Education Office of Assessment and Accountability*. Retrieved May 15, 2002, from [http://kde.state.ky.us/comm/commrel/kiris\\_report\\_gender\\_race/Report\\_KIRIS.doc](http://kde.state.ky.us/comm/commrel/kiris_report_gender_race/Report_KIRIS.doc)
- Smith, M. S. (1995). Education reform in America's public schools: The Clinton agenda. In D. Ravitch (Ed.), *Debating the future of American education* (pp. 9-32). Washington, DC: Brookings Institution.
- Smith, M. S., & O'Day, J. (1991). Systemic school reform. In S. H. Fuhman & B. Malen (Eds.), *The politics of curriculum and testing: The 1990 yearbook of the Politics of Education Association* (pp. 233-267). New York: Farmer Press.
- Solomon, D., Battistich, V., & Hom, A. (1996). Teacher beliefs and practices in schools serving communities that differ in socioeconomic level. *The Journal of Experimental Education*, 64, 327-347.
- Spievak, E., Snyder, E., Miles, S., & Burns, B. (2001). *Achievement trajectories: An analysis of the impact of family and school factors on children's patterns of achievement between the third and sixth grade* (Technical Report). Louisville: University of Louisville, Department of Psychological and Brain Sciences, Cognition and Development Lab.
- Steele, C. M., & Aronson, J. (1998). Stereotype threat and the test performance of academically successful African Americans. In M. Phillips (Ed.), *The black-white test score gap* (pp. 401-427). Washington, DC: Brookings Institution Press.

- Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Stipek, D. J. (2002). *Motivation to learn: Integrating theory and practice* (4th ed.). Boston: Allyn and Bacon.
- Stufflebeam, D. L. (1997). Teacher Work Sample Methodology: Educational policy review. In J. Millman (Ed.), *Grading teachers, grading schools. Is student achievement a valid evaluation measure?* (pp. 53-61). Thousand Oaks, CA: Corwin Press.
- Sutton, A., & Soderstrom, I. (1999). Predicting elementary and secondary school achievement with school-related demographic factors. *Journal of Educational Research, 92*, 330-340.
- Swanson, C. B., & Stevenson, D. L. (2002). Standards-based reform in practice: Evidence on state policy and classroom instruction from the NAEP state assessments. *Educational Evaluation and Policy Analysis, 24*, 1-27.
- Tabachnick, B. G., & Fidell, L. S. (1983). *Using multivariate statistics*. New York: Harper & Row.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn & Bacon.
- Tharp, R. G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context*. New York: Cambridge University Press.
- Thomas, M. D., & Bainbridge, W. L. (2001). 'All children can learn': Facts and fallacies. *Phi Delta Kappan, 82*, 660-662.



- Top and bottom schools. (2001). *The Courier-Journal* (Courier-Journal.com News ed., Vol. 2002). Louisville, KY: The Courier-Journal.
- Tyack, D. B. (1974). *The one best system: A history of American urban education*. Cambridge, MA: Harvard University Press.
- U.S. Census Bureau. (2003a, June). *Children's living arrangements and characteristics: March 2002*. Retrieved July 23, 2003, from <http://www.census.gov/population/www/socdemo/hh-fam/cps2002.html>
- U.S. Census Bureau. (2003b, April 25). *Current population survey, March 2002*. Retrieved July 22, 2003, from <http://www.census.gov/population/socdemo/race/black/ppl-164/tab01.xls>
- Vanneman, A. (1998a). *Long-term trends in student mathematics performance* (No. NCES-98-462). Washington, DC: National Center for Education Statistics. (ERIC Document Reproduction Service No. ED424115)
- Vanneman, A. (1998b). *Long-term trends in student reading performance [and] long-term trends in student mathematics performance [and] long-term trends in student science performance [and] long-term trends in student writing performance* (No. NCES-98-464, NCES-98-462, NCES-98-468, NCES-98-465). Washington, DC: National Center for Education Statistics, U.S. Department of Education. (ERIC Document Reproduction Service No. ED451199)
- Vanneman, A., & White, S. (1998). *Long-term trends in student reading performance* (No. NCES-98-464). Washington, DC: National Center for Education Statistics. (ERIC Document Reproduction Service No. ED415494)

- Watts, J. A., Gaines, G. F., & Creech, J. D. (1998). *Getting results: A fresh look at school accountability*. Atlanta: Southern Regional Education Board. (ERIC Document Reproduction Service No. ED426510)
- Webster, W. J., Mendro, R. L., Orsak, T. H., & Weerasinghe, D. (1998, April 13-17). *An application of hierarchical linear modeling to the estimation of school and teacher effect*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego.
- Weisner, T. S., & Garnier, H. (1992). Nonconventional family life-styles and school achievement: A 12-year longitudinal study. *American Educational Research Journal*, 29, 605-632.
- Wheelock, A. (2000). A new look at school accountability. In B. L. Whitford & K. Jones (Eds.), *Accountability, assessment, and teacher commitment* (pp. 179-209). Albany: State University of New York.
- White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91, 461-481.
- Whitford, B. L., & Jones, K. (2000). The next generation of school accountability. In B. L. Whitford & K. Jones (Eds.), *Accountability, assessment, and teacher commitment* (pp. 233-246). Albany, NY: State University of New York.
- Wilson, W. J. (1987). *The truly disadvantaged: The inner city, the underclass, and public policy*. Chicago: University of Chicago Press.
- Zhang, L., & Manon, J. (2000, April). *Gender and achievement--Understanding gender differences and similarities in mathematics assessment*. Paper presented at the

Annual Meeting of the American Educational Research Association, New Orleans.

APPENDIX A

LETTER OF APPROVAL FROM THE HUMAN STUDIES COMMITTEE



UNIVERSITY of LOUISVILLE

Health Sciences Center

January 28, 2003

■ HUMAN STUDIES COMMITTEES

Instructional Building  
Room 230  
500 South Preston Street  
University of Louisville  
Louisville, Kentucky 40292  
Office: 502-852-5188  
Fax: 502-852-2164  
research.louisville.edu

Stephen K. Miller, PhD  
LFHRE  
343 Education

**RE: 22-03 The Influence of Selected Demographic Factors on Measurements of Student Achievement in One Kentucky School District**

Dear Doctor Miller:

The Human Studies Committees office has received the above study. When inquiring about your study, please refer to the assigned number.

This study has been approved through the Expedited Review Procedure, according to 45 CFR 46.110(b), category 7, since it is research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation or quality assurance methodologies.

The purpose of the study is to investigate the relationship among demographic variables and student achievement scores by means of multiple regression analyses. The information collected will expand the knowledge base regarding the extent of the effects of demographic background factors on individual student achievement and on school-level performance scores in order to support continuing efforts at improving education.

This study was also approved through 45 CFR 46.116 (D), which means that it has been granted a waiver of informed consent because it meets the following criteria:

- > The research involves no more than minimal risk to the subjects.
- > The waiver or alteration will not adversely affect the rights and welfare of the subjects.
- > The research could not practicably be carried out without the waiver or alteration.
- > Whenever appropriate, the subjects will be provided with the additional pertinent information after participation.

**The study has approval through January 27, 2004. You should complete and return a Progress Report/Continuation Request Form eight weeks prior to this date in order to ensure that no lapse in approval occurs.** It will be necessary to send the completed form promptly for Committee review and re-approval of the study. Federal regulatory agencies have indicated that studies must be re-approved by the Committee by the expiration date otherwise the approval will expire. Regulatory agencies have indicated that no further subjects may be entered until the study is re-approved by the

**Please note: new forms are located on the web and downloadable in 2 formats for easier use.  
You must use the most current forms for all submissions.  
<http://research.louisville.edu/UHSC/index.htm>**

Best wishes for a successful study.

Sincerely,

*Edward R. Leist*

Edward R. Leist, Pharm.D.  
Vice Chair, Human Studies Committees

ERL/rsh

**APPENDIX B**

**POPULATION PARAMETERS FOR**  
**STUDENT LEVEL INDEPENDENT VARIABLES**

## Appendix B

### Population Parameters for Student-Level Independent Variables

Table B1

*Number and Percentage of Elementary School Students Qualifying for Subsidized Meals*

Subsidy	Grade 3		Grade 3		Grade 4		Grade 5	
	Reading		Math					
	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>
Free	3405	49.0	3404	49.0	3306	46.3	3415	46.4
Reduced	609	8.8	609	8.8	699	9.4	642	8.7
None	2934	42.2	2934	42.2	3168	44.4	3300	44.9
Totals	6946	100.0	6947	100.0	7133	100.1	7439	100.0

*Note.* All percentages do not add up to 100 because of rounding.



Table B2

*Number and Percentage of Elementary School Students' Mobility*

Mobility	Grade 3		Grade 3		Grade 4		Grade 5	
	Read		Math		N	P	N	P
	N	P	N	P				
0	6499	93.5	6498	93.5	6691	93.7	6968	94.7
1	385	5.5	385	5.5	376	5.3	303	4.1
2	51	0.7	51	0.7	60	0.8	81	1.1
3	8	0.1	8	0.1	10	0.1	3	0.0
4+	5	0.1	5	0.1	6	0.1	2	0.0
Total	6948	99.9	6947	99.9	7143	100.0	7357	99.9

*Note.* Mobility is the number of times a student changes schools during a school year.

All percentages do not add up to 100 because of rounding.

Table B3

*Number and Percentage of Elementary School Students for Nominal-Level Independent*

Variable	Grade 3		Grade 3		Grade 4		Grade 5	
	Read		Math					
	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>
<b>Gender</b>								
Female	3311	47.7	3311	47.1	3502	49.0	3662	49.8
Male	3637	52.3	3636	52.3	3641	51.0	3695	50.2
<b>Ethnicity</b>								
Black	2576	37.1	2576	37.1	2640	37.0	2704	36.8
Other	4372	62.9	4371	62.9	4503	63.0	4653	63.2
<b>Disability</b>								
ECE	685	9.9	685	9.9	788	11.0	828	11.3
Other	6263	90.1	6262	90.1	6355	89.0	6529	88.7
<b>Giftedness</b>								
G&T	212	3.1	212	3.1	574	8.0	709	9.6
Other	6736	96.9	6735	96.9	6569	92.0	6648	90.4
<b>Families</b>								
Two Parents <sup>a</sup>	2935	42.2	2934	42.2	3250	45.5	3183	43.3
Other	4013	57.8	4013	57.8	3893	54.5	4174	56.7
Totals	6948	100.0	6947	100.0	7143	100.0	7357	100.0
<i>Variables</i>								

<sup>a</sup>Two Parents means the student lives with both mother and father rather than one parent, step-parents, foster parents, or other guardians. See Chapter III for more detail.

Table B4

*Number and Percentage of Middle School Students Qualifying for Subsidized Meals*

Subsidy	Grade 6 Reading		Grade 6 Math		Grade 7		Grade 8	
	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>
	Free	3048	44.1	3047	44.1	2652	41.5	2267
Reduced	618	8.9	618	8.9	566	8.9	512	8.4
None	3248	47.0	3248	47.0	3171	49.6	3310	54.4
Totals	6913	100.0	6913	100. 0	6389	100.0	6089	100.0

Table B5

*Number and Percentage of Middle School Students' Mobility*

Mobility	Grade 6		Grade 6		Grade 7		Grade 8	
	Read		Math		N	P	N	P
	N	P	N	P				
0	6355	91.9	6347	91.8	5933	92.9	5693	93.5
1	442	6.4	444	6.4	333	5.2	269	4.4
2	102	1.5	104	1.5	92	1.4	92	1.5
3	14	0.2	14	0.2	18	0.3	21	0.3
4+	4	0.1	4	0.1	13	0.2	14	0.2
Total	6913	100.1	6913	100.0	6913	100.0	6913	99.9

*Note.* Mobility is the number of times a student changes schools during a school year.

All percentages do not add up to 100 because of rounding.

Table B6

*Number and Percentage of Middle School Students for Nominal-Level Independent Variables*

Variable	Grade 6 Read		Grade 6 Math		Grade 7		Grade 8	
	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>
	<b>Gender</b>							
Female	3399	49.1	3395	49.1	3168	49.6	3065	50.3
Male	3518	50.9	3518	50.9	3221	50.4	3024	49.7
<b>Ethnicity</b>								
Black	2509	36.3	2509	36.3	2250	35.2	2091	34.3
Other	4408	63.7	4404	63.7	4139	64.8	3998	65.7
<b>Disability</b>								
ECE	694	10.0	695	10.1	635	9.9	624	10.2
Other	6223	90.0	6218	89.9	5754	90.1	5465	89.8
<b>Giftedness</b>								
G&T	605	8.7	604	8.7	685	10.7	795	13.1
Other	6312	91.3	6309	91.3	5704	89.3	5294	86.9
<b>Families</b>								
Two Parents <sup>a</sup>	2997	43.3	2997	43.4	2749	43.0	2677	44.0
Other	3920	56.7	3916	56.6	3640	57.0	3412	56.0
Totals	6917	100.0	6913	100.0	6389	100.0	6089	100.0

<sup>a</sup>Two Parents means the student lives with both mother and father rather than one parent, step-parents, foster parents, or other guardians. See Chapter III for more detail.

Table B7

*Number and Percentage of High School Students Qualifying for Subsidized Meals*

Subsidy	Grade 9		Grade 9		Grade 10		Grade 11	
	Reading		Math					
	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>
Free	1886	31.4	1881	31.4	1781	28.2	1179	22.3
Reduced	414	6.9	413	6.9	453	7.2	356	6.7
None	3714	61.8	3706	61.8	4082	64.6	3747	70.9
Totals	6014	100.1	6000	100. 1	6316	100.0	5282	99.9

*Note.* All percentages do not add up to 100 because of rounding.

Table B8

*Number and Percentage of High School Students' Mobility*

Mobility	Grade 9		Grade 9		Grade 10		Grade 11	
	Read		Math					
	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>
0	5500	91.5	5489	91.5	5790	91.7	4975	94.2
1	322	5.4	321	5.4	330	5.2	165	3.1
2	123	2.0	122	2.0	130	2.1	107	2.0
3	37	0.6	38	0.6	37	0.6	23	0.4
4+	32	0.5	30	0.5	29	0.5	12	0.2
Total	6014	100.0	6000	100.0	6316	100.1	5282	99.9

*Note.* Mobility is the number of times a student changes schools during a school year.

All percentages do not add up to 100 because of rounding.

Table B9

*Number and Percentage of High School Students for Nominal-Level Independent**Variables*

Variable	Grade 9		Grade 9		Grade 10		Grade 11	
	Read		Math					
	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>	<i>N</i>	<i>P</i>
<b>Gender</b>								
Female	3059	50.9	3051	50.9	3112	49.3	2744	52.0
Male	2955	49.1	2949	49.2	3204	50.7	2538	48.0
<b>Ethnicity</b>								
Black	1839	30.6	1832	30.5	1894	30.0	1393	26.4
Other	4175	69.4	4168	69.5	4422	70.0	3889	73.6
<b>Disability</b>								
ECE	391	6.5	388	6.5	425	6.7	303	5.7
Other	5623	93.5	5612	93.5	5891	93.3	4979	94.3
<b>Giftedness</b>								
G&T	607	10.1	605	10.1	820	13.0	782	14.8
Other	5407	89.9	5395	89.9	5496	87.0	4500	85.2
<b>Families</b>								
Two Parents <sup>a</sup>	2730	45.4	2723	45.4	2729	43.2	2357	44.6
Other	3284	54.6	3277	54.6	3587	56.8	2925	55.4
Totals	6014	100.0	6000	100.0 <sup>b</sup>	6316	100.0	5282	100.0

<sup>a</sup>Two Parents means the student lives with both mother and father rather than one parent, step-parents, foster parents, or other guardians. See Chapter III for more detail.

<sup>b</sup>All percentages do not add up to 100 because of rounding.



Table B10

*Mean and Standard Deviation for Student Mobility at Each Grade*

Grade	3 <sup>a</sup>	4	5	6 <sup>a</sup>	7	8	9 <sup>a</sup>	10	11
Mean	.08	.08	.07	.10	.10	.10	.14	.13	.09
Standard Deviation	.33	.36	.30	.38	.44	.45	.54	.55	.43

*Note.* Mobility is the number of times a student changes schools during a school year.

<sup>a</sup>Only one mean and standard deviation is given for each of grades 3, 6, and 9 because the values were essentially the same for the students taking the CTBS Reading and Mathematics tests.

**APPENDIX C**

**STUDENT LEVEL DEPENDENT VARIABLES:**

**CATS ACHIEVEMENT TEST RESULTS**

## Appendix C

### Student Level Dependent Variables: CATS Achievement Test Results

The KCCT Reading and Mathematics tests are reported as scale scores that range from 325 to 800. For the first year, the scale scores had a mean of about 500 and a standard deviation of about 50 (KDE, 2002a). The CTBS Reading and Mathematics tests are expressed in Normal Curve Equivalents ranging from 1 to 99.

Table C1

#### *CATS Tests Parameters for Elementary Students*

Test	<i>N</i>	Range	<i>M</i>	<i>SD</i>
CTBS Reading (Grade 3)	6948	98	50.87	21.41
CTBS Mathematics (Grade 3)	6947	98	51.28	21.00
KCCT Reading (Grade 4)	7143	380	539.62	47.79
KCCT Mathematics (Grade 5)	7357	475	554.44	54.47

Table C2

*CATS Tests Parameters for Middle School Students*

Test	<i>N</i>	Range	<i>M</i>	<i>SD</i>
CTBS Reading (Grade 6)	6917	98	47.41	19.95
CTBS Mathematics (Grade 6)	6913	98	46.91	22.17
KCCT Reading (Grade 7)	6389	475	504.93	42.54
KCCT Mathematics (Grade 8)	6089	475	519.82	52.13

Table C3

*CATS Tests Parameters for High School Students*

Test	<i>N</i>	Range	<i>M</i>	<i>SD</i>
CTBS Reading (Grade 9)	6014	98	52.55	18.89
CTBS Mathematics (Grade 9)	6000	98	49.67	22.25
KCCT Reading (Grade 10)	6316	475	500.22	64.69
KCCT Mathematics (Grade 11)	5282	475	529.20	65.31

**APPENDIX D**

**POPULATION PARAMETERS FOR  
SCHOOL LEVEL INDEPENDENT VARIABLES**

## Appendix D

### Population Parameters for School Level Independent Variables

All parameters in Appendix D are reported in percentages.

Table D1

*Population Parameters for Elementary School Composition Variables (N = 87)*

Variable	Lowest percent	Highest percent	Range	<i>M</i>	<i>SD</i>	<i>Mdn</i>
%Females	43.0	53.7	11.7	48.2	2.3	48.0
%Black	17.0	50.2	34.2	36.8	9.7	40.1
%ECE	1.0	13.9	13.9	7.4	2.7	7.4
%Gifted	0.0 <sup>a</sup>	45.4	46.4	8.9	11.9	2.1
%HiSES	6.8	89.8	84.0	41.3	21.6	39.1
Mobil Rate	0.4	25.1	25.7	10.5	5.0	10.1
%Two Parents	20.5	75.6	56.1	41.8	13.1	40.1

*Note.* These independent variables represent the school composition of the elementary schools for the year 2002 and were used for the analysis of all CATS tests and indices at the elementary level.

<sup>a</sup>34 elementary schools were 0.0% Gifted.

Table D2

*Population Parameters for Middle School Composition Variables (N = 24)*

Variable	Lowest percent	Highest percent	Range	<i>M</i>	<i>SD</i>	<i>Mdn</i>
%Females	44.0	58.0	15.0	49.5	3.9	48.1
%Black	24.1	56.0	33.1	35.1	8.3	34.0
%ECE	1.1	22.3	22.2	12.1	4.6	12.1
%Gifted	0.0 <sup>a</sup>	35.8	36.8	10.8	12.9	5.7
%HiSES	12.0	84.6	73.6	51.5	18.6	51.6
Mobil Rate	1.0	31.9	31.9	12.2	7.9	11.4
%Two Parents	21.1	75.3	55.2	43.5	12.2	41.1

*Note.* These independent variables represent the school composition of the middle schools for the year 2002 and were used for the analysis of all CATS tests and indices at the middle school level.

<sup>a</sup>10 middle schools were 0.0% Gifted.

Table D3

*Population Parameters for High School Composition Variables (N = 21)*

Variable	Lowest percent	Highest percent	Range	<i>M</i>	<i>SD</i>	<i>Mdn</i>
%Females	40.1	64.2	25.1	49.2	5.6	48.4
%Black	19.9	71.4	52.5	32.1	13.2	25.8
%ECE	1.5	24.4	23.9	11.2	6.1	11.1
%Gifted	0.0 <sup>a</sup>	45.0	46.0	10.0	12.7	5.3
%HiSES	31.8	90.1	59.3	62.6	18.0	59.8
Mobil Rate	0.7	22.4	22.7	9.4	6.4	7.9
%Two Parents	20.8	66.2	46.4	42.0	13.1	40.9

*Note.* These independent variables represent the school composition for the high schools for the year 2002 and were used for the analysis of all of the CATS tests and indices analyzed at the high school level.

<sup>a</sup>9 high schools were 0.0% Gifted.



**APPENDIX E**

**SCHOOL LEVEL DEPENDENT VARIABLES:**

**CATS ACHIEVEMENT TESTS RESULTS**

## Appendix E

### School-Level Dependent Variables: CATS Achievement Test Results

In Appendix E the KCCT Reading and Mathematics tests are reported as scale scores that range from 325 to 800. The CTBS Reading and Mathematics tests are expressed in Normal Curve Equivalents that range from 1 to 99. The Academic Index and Norm Referenced Test (NRT) Index are computed based on weighted formulas (See Chapter III). Both indices range from 1 to 140. A score of 100 is considered proficient and is the year 2014 goal for all Kentucky schools.

Table E1

*Parameters for Elementary School CATS Tests and Indices (N = 87)*

Test	Range	<i>M</i>	<i>SD</i>
CTBS Reading (Grade 3)	31.9	50.2	7.4
CTBS Mathematics (Grade 3)	34.1	50.7	7.9
KCCT Reading (Grade 4)	67.0	542.4	14.6
KCCT Mathematics (Grade 5)	93.0	557.0	17.0
NRT Index (school)	83.5	76.1	19.6
Acad. Index (school)	57.3	60.6	12.2

Table E2

*Parameters for Middle School CATS Tests and Indices (N = 24)*

Test	Range	<i>M</i>	<i>SD</i>
CTBS Reading (Grade 6)	28.2	47.3	8.0
CTBS Mathematics (Grade 6)	32.5	46.6	8.9
KCCT Reading (Grade 7)	56.0	506.3	14.5
KCCT Mathematics (Grade 8)	63.0	520.5	17.0
NRT Index (school)	79.8	65.9	22.4
Acad. Index (school)	49.5	58.2	12.7

Table E3

*Parameters for High Schools CATS Tests and Indices (N = 21)*

Test	Range	<i>M</i>	<i>SD</i>
CTBS Reading (Grade 9)	35.8	50.1	8.8
CTBS Mathematics (Grade 9)	38.6	46.6	10.6
KCCT Reading (Grade 10)	107.0	496.3	28.9
KCCT Mathematics (Grade 11)	110.0	524.6	30.1
NRT Index (school)	92.4	66.6	26.0
Acad. Index (school)	55.3	61.8	16.4

## CURRICULUM VITAE

NAME: Beverly Derington Moore

ADDRESS: 2424 Dundee Road  
Louisville, KY 40205

DOB: Fort Myers, Florida - November 9, 1938

EDUCATION: A.B., Biology Major/English Minor  
Emory University, Atlanta, GA  
1956-1960

M.S., Biology  
Emory University, Atlanta, GA  
1961-1964

M.Ed., Science/Gifted Education  
University of Louisville  
1984-1985

AWARDS: French Government Award 1960  
League of Women Voters Social Justice Award  
University of Louisville Fellowship 1985, 1986  
Grawemeyer Fellowship 2001, 2002

PROFESSIONAL SOCIETIES: American Education Research Association  
Phi Delta Kappa

### PUBLICATIONS:

Moore, B. D. (1989). Improved effectiveness or increased control: The Kentucky Beginning Teacher Program. In *Philosophical Studies in Education: Proceedings of the Annual Meeting of the Ohio Valley Philosophy of Education Society 1987-89* (pp. 344-358).

Coe, P., Kannapel, P. J., Moore, B. D., & Aagaard, L. (1993). School-based decisionmaking after two years. *Notes from the Field: Education Reform in Rural Kentucky*, 3(2).

- Coe, P., Kannapel, P. J., Moore, B. D., & Aagaard, L. (1993). Kentucky's Primary Program. Notes from the Field: Education Reform in Rural Kentucky, 3(1).
- Coe, P., Kannapel, P. J., Aagaard, L., & Moore, B. D. (1994). Instruction and assessment in accountable and nonaccountable grades. Notes from the Field: Education Reform in Rural Kentucky, 4(1).
- Appalachia Educational Laboratory [Coe, P., Kannapel, P. J., Aagaard, L., & Moore, B. D.] (1995). AEL study of KERA implementation in four rural Kentucky school districts 1993-1994 (Annual). Charleston, WV: Author
- Appalachia Educational Laboratory [Coe, P., Kannapel, P. J., Aagaard, L., & Moore, B. D.] (1995). 1991-1995 Final report of a study of KERA implementation in four rural Kentucky school districts 1993-1994 (Final). Charleston, WV: Author
- Kannapel, P. J., Moore, B. D., Coe, P., & Aagaard, L. (1995). Six heads are better than one? School-based decision making in rural Kentucky. *Journal of Research in Rural Education*, 11(1), 15-23.
- Kannapel, P. J., Coe, P., Aagaard, L., Moore, B. D., & Reeves, C. A. (2000). Teacher responses to rewards and sanctions. In B. L. Whitford & K. Jones (Eds.), *Accountability, assessment, and teacher commitment* (pp. 127-146). Albany, NY: State University of New York. (pp. 127-146). Albany, NY: State University of New York.
- Moore, B. D. (2001, November 9). Continuity and success. *The Courier-Journal*, p. A9.

#### NATIONAL MEETING PRESENTATIONS:

- Miller, S. K., & Moore, B. D. (1990, April). Whither tracking? The unexamined element in the educational reform movement. Paper presented at the American Educational Research Association Annual Meeting, Boston.
- Kannapel, P. J., Moore, B. D., Coe, P., & Aagaard, L. (1994, April). School-based decision making in rural Kentucky schools: Interim findings of a five-year longitudinal study. Paper presented at the Annual Meeting of AERA, New Orleans.
- Kannapel, P. J., Moore, B. D., Coe, P., & Aagaard, L. (1995, April). Opposition to outcome-based education in Kentucky. Paper presented at the American Educational Research Association Annual Meeting, San Francisco.

Kannapel, P. J., Coe, P., Aargaard, L., & Moore, B. D. (1995, April). Nonlinear evolution of school-based decisionmaking in Kentucky. Paper presented at the American Educational Research Association Annual Meeting, San Francisco.

Kannapel, P. J., Moore, B. D., & Coe, P. (1996, April). Vulnerability and invulnerability to change in two rural Kentucky school districts. Paper presented at the American Educational Research Association Annual Meeting, New York.

Kannapel, P. J., Coe, P., Aargaard, L., & Moore, B. D. (1996, April). "I don't give a hoot if somebody is going to pay me \$3600". Paper presented at the American Educational Research Association Annual Meeting, New York.

#### INVITED PRESENTATIONS:

Moore, B. D. (2000). The implementation of reform in a large Kentucky district. San Francisco: National School Board Association.