2007

Third grade standardized tests as indicators and estimates of future performance on the ACT

Greg Wieman

Third Grade Standardized Tests as Indicators and Estimates of Future Performance on the ACT

by

Greg Wieman

Dissertation

Submitted to the Department of Leadership and Counseling

Eastern Michigan University

in partial fulfillment of the requirements for the degree of

DOCTOR OF EDUCATION

Dissertation Committee:

Ronald Williamson, EdD, Chair

Nelson Maylone, EdD

Jaclynn Tracy, PhD

Gary Marx, EdD

2007

Ypsilanti, Michigan

ACKNOWLEDGEMENTS

I wish to express deep appreciation to my committee chair, Dr. Ronald Williamson. His patience, guidance, and encouragement typify the tremendous support provided by the Eastern Michigan University faculty to its students.

I wish to thank Dr. Nelson Maylone, Dr. Jaclynn Tracy, Dr. Gary Marx, Dr. Ella Burton, and Dr. Helen Ditzhazy for serving on my dissertation committee. Thank you to Dr. William Price for serving as my academic advisor.

Thank you to Dr. Soon Hong for his advice on statistical analysis and Dr. Norma Ross for her suggestions as to the final edits of this document.

The love and support of a family is an important asset for educational success. I have been blessed to have positive support throughout my life from my parents, Roger and Carolyn, and brother, Scott. It has been magnified by my wife Kris, my daughter Jill, and my son Grant.

To all educators: Continue the quest of making it a better world for our children.

ABSTRACT

The purpose of this study was to determine the relationship between 3rd grade standardized test scores and ACT Composite scores in high school. A second purpose of the study was to determine the consistency of rank within a cohort from 3rd grade to high school as measured by standardized tests.

Data consisted of standardized test scores from three Michigan public school districts analyzed as distinct cohorts. Correlations were determined using various factors from three different commonly used 3rd grade standardized tests and ACT Composite scores. Multiple regressions of 3rd grade test factors were generated to yield the strongest correlations with ACT Composite scores.

Strong positive correlations were found between 3rd grade standardized test scores and high school ACT Composite scores. The study also indicated consistency of individual student rank within a cohort from 3rd grade to high school. Variability of ACT Composite scores explained by 3rd grade standardized tests ranged from 55 to 90 percent. Variability of cohort rank from 3rd grade to high school ranged from 58 to 90 percent.

The results of this study cast doubt on the value of using standardized tests to measure instructional effectiveness or the quality of schools.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction and Background

The federal legislation labeled the "No Child Left Behind" (NCLB) Act of 2001 (Public Law 107-100) was signed into law by President George W. Bush on January 8, 2002. The law mandates that by January 8, 2014, all students in the United States will attain full achievement of the levels of academic proficiency as established by each state. The main goal of this legislation is to bring all students to these state levels of achievement as measured by standardized tests (McLeod, S., D'Amico, J. J., & Protheroe, N., 2003). Despite the negative consequences of using standardized tests in our educational system (Popham, 2001), standardized test scores are used to determine the effectiveness of our schools (McLeod, S., D'Amico, J. J., & Protheroe, N., 2003; Standard & Poor, 2003).

The pressure to fulfill the mandate of NCLB rests upon the educational system. One of the accountability standards of NCLB is Adequate Yearly Progress (AYP). AYP requires an incremental increase in the number of students meeting state levels of achievement as measured by state standardized tests for various subgroups. As the standards for AYP continue to rise, the failure to meet these standards will result in sanctions for schools, districts, and each state (U.S. Department of Education, 2005).

Standardized test scores were not designed to measure a school's effectiveness (Ravitch, 2002). The requirement to

fulfill AYP continues in spite of a plethora of evidence showing that the results on standardized tests are contingent on factors outside of the educational system (Adams, 1994; Barton, 2004; Jencks & Phillips, 1998; Roscigno, 1998; Rothstein, 2004; Thorndyke, 1951). The implementation of NCLB has not had a significant impact on reading and mathematics achievement and has not helped to significantly narrow the racial and socioeconomic achievement gaps (Lee, 2006).

In the spring of 2007, Michigan will join Illinois and Colorado in using the ACT as a portion of its state assessment, called the Michigan Merit Exam (MME), for high school students to meet the requirements of NCLB. The MME will be used to measure AYP for high schools in Michigan. The ACT is a college entrance exam that assesses high school students' general educational development and their ability to complete college-level work. The ACT has shown consistent score gaps between different racial and ethnic groups (ACT, 2006). The switch to the ACT from the Michigan Educational Assessment Program (MEAP) test gives little hope of reducing score gaps. A direct relationship between the demographics of a school district and its MEAP test scores was shown in Michigan (Maylone, 2002).

This study investigated the consistency of standardized test scores of individual students from third grade to high school. Confirmation that individual student performance on standardized tests was determined during the elementary years and resistant to change by the end of high school would help to illustrate the difficulty of closing achievement gaps of various

subgroups by high school and assert the ineffectiveness of using standardized tests scores as a measurement of school and teacher performance.

Statement of the Problem

Why do achievement gaps as measured by standardized test scores continue despite the call to eliminate them? Consistent achievement gaps as measured by standardized tests have remained between different racial/ethnic groups and by socioeconomic level (National Center for Educational Statistics, 2006). If performance on standardized tests is consistent throughout a student's academic career, the mandate of NCLB to close these achievement gaps as measured by standardized tests will not be fulfilled.

Purpose of the Study/Justification and Significance

The purpose of this study was to examine the consistency of individual student standardized test scores from third grade to high school in two ways: first, by determining the relationship of third grade standardized test scores and high school ACT Composite scores; and second, by determining the consistency of individual student rank within a cohort from third grade to high school as measured by standardized test scores. Consistency of standardized test scores and student rank within a cohort from third grade through high school may demonstrate that standardized tests as an entity are not easily influenced by our

educational system, thus representing the difficulty of closing achievement gaps.

Delimitations

In an attempt to account for a more consistent academic experience, although students may not have had identical teachers or identical curriculums, data were collected from students who took both a third grade standardized test and the ACT while enrolled at the same district. All third grade data were drawn from the following standardized tests: Stanford Achievement Test Series (9th Edition) with Otis-Lennon School Ability Test (7th Edition), Comprehensive Tests of Basic Skills (4th Edition), and TerraNova CTBS Complete Battery (1996). Third grade standardized test scores were compared with high school ACT composite scores. If the ACT was taken multiple times by an individual student, the composite score from the first test taken (chronologically) in high school was recorded.

The ACT was used as an admission tool for college and not required for all students in the class of 2006. In Michigan, the ACT was administered to self-selected students who were deemed college-bound. Therefore, the demographics of the sample data may not match the demographics of the school districts analyzed in this study.

Limitations

Uncontrolled variables such as family support, chemical usage, and individual motivation may have an effect on academic achievement (Barton, 2004). Variance of student schedules, teacher assignment, student transience, and student turnover were not accounted for in this study because of the difficulty of measuring and comparing these variables. The study did not account for practice or special preparation by students for the third grade standardized tests (SRA/McGraw-Hill, 2003) or the ACT (ACT, 2006; Michigan Department of Education, 2006B).

Methodology

A statistical analysis was conducted to indicate the extent of relationships between third grade standardized tests scores and high school ACT scores. Pearson Product-Moment correlation coefficients were determined to measure the strength of those relationships. A correlation was computed for each of the identified variables and for all of the samples combined. Multiple regressions were used to find the best combination of predictor variables. The SPSS software program was used for all statistical analyses.

Individual student performance on standardized tests taken in third grade was compared statistically with the same student's ACT Composite scores at the high school level. Normal Curve Equivalent (NCE) scores (range 1-99) on standardized tests

were used to determine statistical relationships and find the best predictive variables.

The population in this study consisted of students from the Class of 2006 at three Michigan high schools. Data from the standardized test scores of those students were used in the study. Third grade standardized test scores and a high school ACT composite score were recorded for each student. The selected districts varied in size and demographics. The data obtained for the study were secondary data. All data obtained for the study were individual test scores recorded in rows. Any information that provided individual student identification was not recorded to assure the anonymity of all subjects.

Preliminary administrative approval was granted in writing from three public school districts before obtaining data for the study. A formal written document of approval was granted by the Human Subjects review process at Eastern Michigan University and the dissertation committee. Data were available by cohort (students who were in the same school system in third grade and again in twelfth grade) for each school district.

Data were recorded on a tally sheet in rows by individual student and then entered into a spreadsheet for statistical analysis. Name, gender, identification number, and all other individual information were removed from the tally sheets, leaving no possibility for the recognition of individual student identities. It was an unobtrusive procedure, as only student academic records were used in the study. There was no attempt to influence the variables.

Third grade standardized test scores were compared with high school ACT Composite scores to determine if there were any statistical relationships. Third grade standardized tests that were used in this study included TerraNova Comprehensive Tests of Basic Skills (CTBS), TerraNova California Achievement Tests (CAT), and Stanford Achievement Test (SAT) with the Otis-Lennon School Ability Index. The third grade standardized tests were norm-referenced and considered reliable, commonly used, national standardized tests (Popham, 2001). Michigan will begin using the ACT in the spring of 2007 as its state-required high school assessment, replacing the Michigan Educational Assessment Program (MEAP) test (Roeber, 2006). The ACT is presently used in Colorado and Illinois at the high school level to meet federal standards as required by NCLB.

Definition of Terms

The following definitions were used to clarify the study. The definitions include descriptions of the standardized tests identified in the study.

ACT Composite score

A curriculum-based test intended to measure the skills and knowledge that students have acquired in high school and need to be successful in college. The ACT Composite score is the arithmetic average of the scores on the four academic subject areas of the test: English, Mathematics, Reading, and Science. Scores are reported on a scale of 1 to 36 (Noble, J. P., Roberts, W. L., & Sawyer, R. L., 2006).

Comprehensive Tests of Basic Skills (CTBS)

A norm-referenced test that measures student achievement in reading/language arts, mathematics, science, and social studies (McGraw-Hill Education, 2006).

Michigan Educational Assessment Program (MEAP)

The statewide assessment program used in Michigan to test and report student achievement in the core academic subjects at certain grade levels (MDE, 2006).

Norm-referenced tests

Standardized tests designed to measure how a student's performance compares with that of other students. Scores on norm-referenced tests are often reported in terms of grade-level equivalencies or percentiles derived from the scores of the original students (ASCD, 2005).

Otis-Lennon School Ability Index

A test administered with the Stanford Achievement Test (SAT) that assesses the abilities that relate to success in school (Harcourt Assessment, 2006).

Pearson Product-Moment Test

An index of correlation appropriate when the data represent either interval or ratio scales; it takes into account each and every pair of scores and produces a coefficient between .00 and either +/- 1.00 (Fraenkel & Wallen, 1996).

Standardized tests

Tests that are administered and scored under uniform (standardized) conditions (ASCD, 2005).

Stanford Achievement Test (SAT)

A norm-referenced test that measures student achievement in reading/language arts, mathematics, science, and social studies (Harcourt Assessment, 2006).

Organization

This chapter discussed the need to question standardized tests as a measurement of academic achievement of our schools. Particular emphasis was given to closing the achievement gaps as required by NCLB. Included in this chapter were the purpose of the study/justification and significance, delimitations, limitations, methodology, and definition of terms.

Subsequent chapters provide a review of related literature, the research design and methodology, a presentation and analysis of data, and a discussion of the conclusions, inferences, implications, and recommendations for further research.

CHAPTER 2

Review of Related Literature

This chapter provides a review of literature relevant to the questions and concerns in Chapter One. This review addresses the use of standardized tests as predictors, consistent gaps of achievement as measured by standardized tests, potential issues with the use of the ACT in Michigan, and specific parts of NCLB.

Prediction

The predictive value of standardized tests has been used as early as preschool to predict the success of students when they reach second grade (Funk, Sturner, & Green, 1986). Parental motivational practices (Gottfried, Fleming, & Gottfried, 1994) and parental attributes (Georgiou, 1999) have been studied to predict student achievement. Predictions of gender differences in later achievement (Witt, Dunbar, & Hoover, 1994) and specific (math) predictions for gender (Olszewski-Kubilius, Kulieke, & Shaw, 1990) and race (Powers, Thompson, & Azevedo, 1983) have been made. The predictive abilities of standardized tests have been used to compare Asian and American students (Chen, Lee, & Stevenson, 1996; Stone, 1992). Standardized tests have been used to determine placement in kindergarten (West, Denton, & Germino-Hausken, 2000), for placement in special education programs (Frey, 2002), and to identify talented students for inclusion into programs for accelerated instruction (Olszewski-Kubilius, Kulieke, & Shaw, 1990).

Studies have shown a predictive validity for both the Wechsler Intelligence Scale for Children (Prewett & Fowler, 1992) and the Metropolitan Achievement Test (Weller, Schnittjer, & Tuten, 1992). IQ has been shown to be both a consistent (Rosenbach & Rusch, 1991) and accurate (Antonak, 1988, and Grossman & Johnson, 1983) predictor of academic achievement. Preschool children (mean age 3 years, 9.8 months) were assessed with the McCarthy Scales of Children's Abilities. The study concluded that the McCarthy Scales were valid predictors of school achievement at the second grade level (Fuchs & Migdail, 1977). Kindergarten test scores were highly correlated with third grade test scores (Rock & Stenner, 2005).

ACT, Inc., features the Educational Planning and Assessment System (EPAS), consisting of three testing programs: EXPLORE, PLAN, and the ACT Assessment Program (ACT). EXPLORE is administered in the eighth grade, PLAN in the tenth grade, and ACT in the eleventh or twelfth grade. The ACT is intended to measure the skills and knowledge that students have acquired in high school and need to be successful in college (Noble, Roberts, & Sawyer, 2006). A positive relationship was shown between ACT scores and college grades (Allen & Sconing, 2005).

The composite scores of the three testing programs are strongly correlated when compared by grade level, making the earlier tests highly predictive of the ACT (Woodruff, 2003). There is a consistent high correlation between EXPLORE and PLAN (Roberts & Noble, 2004).

Test Score Gaps

NCLB requires states to set standards for grade-level achievement and to develop a system to measure the progress of all students (U.S. Department of Education, 2004). Academic achievement is measured by proficiency on statewide tests disaggregated by student subgroups such as gender, major racial and ethnic groups, English proficiency status, learning disabilities, and economic disadvantage (McLeod, S., D'Amico, J. J., & Protheroe, N., 2003). Mandating that schools produce level scores on standardized tests for certain subgroups may prove difficult as there have been persistent and large score gaps in the past (Barton, 2005).

There are apparent differences in cognitive skills and knowledge (West, Denton, & Germino-Hausken, 2000) and substantial differences by race and ethnicity in children's test scores as they begin kindergarten (Lee & Burkam, 2002; Rock & Stenner, 2005). Cognitive skills are more closely related to socioeconomic status (SES) than race/ethnicity (Lee & Burkam, 2002).

Achievement gaps have been shown during preschool and kindergarten when accounting for race/ethnicity (Kober, 2001). The Nation's Report Card (Institute of Educational Sciences National Center for Educational Statistics, 2005) is based on the National Assessment of Educational Progress (NAEP) tests given each year across the schools of the United States. Consistent test score gaps on the NAEP are shown in both reading and mathematics when accounting for race/ethnicity and

eligibility for free/reduced lunch. Thorndike (1951) recognized that demographic and community variables explained test score variance fifty years before the implementation of NCLB. Socioeconomic status has been consistently linked to standardized tests scores (Adams, 1994; Maylone, 2002).

Consistent sociocultural factors appear when analyzing academic achievement data. The *Trends in International Mathematics and Science Study* (TIMSS) in 1999 ranked countries by their achievement scores. The top five in mathematics and four of the top five in science were Asian countries. The TIMSS data is similar to the academic achievement in the United States. From 1990 to 2003, Asian and White scores have been significantly higher than Black and Hispanic scores (Institute of Educational Sciences National Center for Education Statistics, 2005).

Achievement-level results by race/ethnicity for mathematics in grades 4 and 8 in the United States (The Nation's Report Card: Mathematics Highlights 2003) have consistently shown Asian/Pacific Islanders (and Whites) with higher scores than other racial/ethnic groups for thirteen years. The same pattern has occurred for reading (The Nation's Report Card: Reading Highlights 2003) for eleven years. Ethnic groups have shown consistent gaps of performance on the ACT (Noble, J., Davenport, M., Schiel, J., & Pommerich, M., 1999). Standardized tests have been noted to contain biased items that contribute to test score gaps (Jencks & Phillips, 1998; Popham, 2001).

Predictions of academic achievement by gender, race or ethnicity, or socioeconomic status run the risk of stereotyping individual students. Carlson (2004) stated, "There is no sound argument for disaggregating scores by race and ethnicity, but serious danger in doing so. The difference itself explains nothing, but it can reinforce for many the notion that some groups are 'naturally' inferior to others in cognitive ability" (p. 379). He listed several factors that may affect group scores including poverty, family dysfunction, poor parenting skills, transiency, substance abuse, and the devaluing of academic achievement. The analysis of achievement data by the scores of individual students may not be as simple but may reflect a more accurate measure of accountability. Carlson acknowledged that, "To break out test scores by easily obtainable but misleading data on race and ethnicity has the advantage only of convenience" (p. 379). High stakes testing increases social stratification (Bracey, 2000), and problems associated with high-stakes testing disproportionately affect minority students (Nichols, Glass, & Berliner, 2005).

A prior concern with the racial disparity of MEAP scores in Michigan will not be remedied with a switch to the ACT (Putnam, 2006). Colorado (Colorado Department of Education, 2006) and Illinois (Illinois State Board of Education, 2006) have shown consistent gaps between various ethnic and racial groups as measured by ACT Composite scores. National ACT Composite score averages from 1997 to 2004 demonstrate stable disparities when

compared by race/ethnicity (Institute of Educational Sciences National Center for Educational Statistics, 2005).

Norm-Referenced Standardized Tests and the ACT

Norm-referenced achievement tests are based on a normal distribution and designed to produce score-spread. Better score-spread increases the test's reliability (Popham, 2001). The NCLB mandate to close the score gap between subgroups of students on standardized tests cannot be accomplished if the instrument itself is designed to produce score-spread.

If ACT scores can be increased by specific test preparation available only with sufficient economic resources, the results of the test may become more contingent on social class. Several informational links for the Michigan Merit Exam (MME) were available on the Michigan Department of Education (2006B) website. Links for parents and students included a letter from the executive director of the Michigan Department of Education Office of Educational Assessment and Accountability, suggesting that test-prep courses have some benefit and a direct link to the ACT guide for preparing for the ACT test. The ACT guide recommended additional test prep materials, including (for a fee) an online practice program (ACT, 2006).

Education YES!

Education YES! is the name given the state accreditation system for schools in Michigan. It has been revised, making student progress on the Michigan Merit Exam (MME) the measure of

AYP for high schools. Results of the MME will also help determine the letter grade assigned to each school, thus labeling the perceived quality of that school (Michigan Department of Education, 2006A). The ACT is a major component of the MME.

The ACT is not directly aligned to Michigan's articulated curriculum, making it less receptive to instruction than the MEAP tests that were based upon the Michigan Curriculum Framework. Popham (2005) stated a concern that schools will fail to achieve AYP because tests chosen by states to meet NCLB standards are instructionally insensitive because the tests are contingent on the socioeconomic status of the students tested.

Pillars of NCLB

The implementation of No Child Left Behind was based upon the four pillars of NCLB (U.S. Department of Education, 2005). The names of the first two pillars, much like the title "No Child Left Behind," made promises that will not be kept. The first pillar, Stronger Accountability for Results, required states to describe how they will close the achievement gap and make sure all students, including those who are disadvantaged, achieve academic proficiency. Continuing to use standardized tests as a measurement will not allow for this gap to close. Research has continually shown that socioeconomic status is consistently linked with standardized test scores.

Rothstein (2004) found an association of social and economic disadvantage with student achievement gaps. Maylone

(2002) showed a direct correlation between socioeconomic status (SES) factors and school district scores on Michigan educational assessment tests. Studies have shown positive relationships between per capita income and SAT scores (Adams, 1994) and household income and ACT scores (Fair Test, 2005).

In their report on "priority schools," Lee and Reimann (2003) showed that 77% of the student body in failing schools received free and reduced lunch, as opposed to 34% of the students in non-failing schools. The average percentage of minority students in failing schools was 86%, while the minority population in non-failing schools was 23%. Priority schools were defined as those 216 schools in Michigan that missed their AYP targets at the end of the 2002-03 school year.

The second pillar, More Freedom for States and Communities, allowed for the transfer and consolidation of federal formula grant funds received by school districts in an attempt to increase revenue for schools. This pillar had little effect because federal funding accounts for about 7% of overall school funding (Austin, 2005), and state funds have been increasingly strained by the economic downturn since 2001.

Resources that affect student achievement are not equally allocated among schools. There is an inequity of per pupil spending between schools (Dively, 2004), and poor and minority children in the U.S. are often forced to attend poorly funded schools (Biddle & Berliner, 2002). Affluent and largely white districts are much more likely to have veteran and certified

teachers than those schools with significant numbers of poor and minority students (Navarette, 2003).

This chapter presented a review of literature related to the use of standardized tests as predictors, consistent gaps of achievement as measured by standardized tests, potential issues with the use of the ACT in Michigan, and specific parts of NCLB and Education YES! Chapter Three provides the research design, methodology for data collection, and procedures for data analysis.

CHAPTER 3

Research Design and Methodology

The requirement of NCLB to close the gaps of achievement as measured by standardized tests will not be attained if scores of standardized tests are resistant to change. The review of literature in Chapter Two showed a consistent gap of achievement as measured by standardized tests. This study was designed to determine the relationship between third grade standardized test scores and high school ACT Composite scores of individual students and the consistency of individual rank within a cohort as measured by these tests.

Methodology

Two questions were under consideration in this study. The intent of the first question was to determine if there was a statistically significant relationship between third grade standardized test scores and high school ACT Composite scores of individual students. The second question would determine the variability of the rank of individual student scores on third grade standardized tests relative to the high school ACT Composite scores within their class cohort group.

Design

Correlation research serves two purposes. The first purpose is to identify relationships among variables. The second purpose is prediction; if the relationship between variables is of

sufficient magnitude, it becomes possible to predict a score on either variable if the score of one variable is known (Fraenkel & Wallen, 1996).

The questions posed in this study suggested a quantitative, correlational, multivariate study design. Thus, third grade standardized test scores and high school ACT Composite scores of individual students were compared to determine if a statistically significant relationship existed between the scores. An additional goal was to determine the relationship of individual student rank within a cohort group, hence demonstrating the consistency of individual test score rank within cohort groups from third grade to high school. The review of literature supported the concept that gaps in individual student standardized test scores exist at the elementary level and remain consistent through high school.

Methods

Step One: Collection of Data.

Data were collected with permission from three Michigan public schools. The selected districts had relatively high student retention rates to facilitate the cohort comparison and were accessible to the researcher. Data for the Class of 2006 for each of the three schools were analyzed in distinct cohorts. Each school administered a norm-referenced third grade standardized test in the spring of 1997.

The school districts in the study were located in the lower peninsula of Michigan, met AYP standards through 2005, and had ACT Composite scores above the state average. The following summations of the districts and Table 1 describe demographic and academic information as obtained from the Standard & Poor's (2006) SchoolMatters website.

District A is predominantly Caucasian and located in western Michigan. The percentage of students receiving free/reduced lunch and special education services was well below the state average in 2005. The adult education levels were significantly above the state averages for 2005. MEAP reading and mathematics scores for grades 3-8 were above the state average for 2006. The ACT participation rate for 2005 was 71.2 percent.

District B is predominantly Caucasian and located in the central part of Michigan. The percentage of students receiving free/reduced lunch was below the state average in 2005. The percentage of students receiving special education services was above the state average in 2005. The adult education levels were slightly above the state average for 2005. MEAP reading and mathematics scores for grades 3-8 were above the state average for 2006 except eighth grade reading. The ACT participation rate for 2005 was 39.8 percent.

District C is about 60 percent Caucasian and located in western Michigan. The percentage of students receiving free/reduced lunch and special education services was slightly below the state averages in 2005. The adult education levels were near the state average for 2005. MEAP reading and

mathematics scores for grades 3-8 were above the state average for 2006 except third grade reading and sixth grade reading. The ACT participation rate for 2005 was 55.0 percent.

Table 1

School District Demographics

School District	A	B	C
Total Student Enrollment	878	983	8157
White Student Population (%)	90.8	95.9	59.9
Black Student Population (%)	2.4	0.8	3.4
Hispanic Student Population (%)	4.0	1.6	22.7
Asian/Pacific Islander Student Population (%)	1.7	1.2	9.6
Economically Disadvantaged	14.0	26.0	33.0
Students with Disabilities	9.7	15.3	12.4
ACT Composite - Average Score	23.4	22.7	22.4

Third grade standardized test scores and high school ACT scores were collected from test rosters and/or CA-60 files for individual students of the class of 2006 from each school. Individual student scores were then transferred to a spreadsheet for statistical analysis. All data obtained for the study were individual test scores recorded in rows. Any information that

identified individual students was not recorded to assure the anonymity of all subjects. Secondary data were used for the study.

Step Two: Correlational Analysis.

Individual coefficients for third grade standardized test scores/ACT Composite scores and cohort rank for third grade and high school were generated using the Pearson Product-Moment Test within the SPSS computer statistical program. These coefficients indicated the level or strength of the relationship between test scores or rank. All correlational coefficients computed were significant at the 0.01 level (2-tailed).

Step Three: Examination of Scatterplots.

The SPSS computer statistical program allowed for the production of scatterplots of paired factors. Scatterplots were produced for each of the third grade standardized test scores used for statistical analysis, paired with the ACT Composite scores. Scatterplots were also produced for third grade standardized test cohort rank paired with ACT Composite cohort rank. The scatterplot of each relationship was examined to confirm the computed correlations. All relationships examined were significant at the 0.01 level.

Step Four: Multiple Regression.

The SPSS program allowed for the examination of several independent variables in relation to an identified dependent variable at once. Multiple regressions were generated using various third grade standardized test scores as independent

variables and the ACT Composite score as the dependent variable for each school.

Step Five: Drawing Conclusions.

Conclusions were drawn following a statistical analysis of the data. Inferences and policy implications are discussed in Chapter Five.

Methodology Notes

This study discovered correlational coefficients between third grade standardized test scores and high school ACT Composite scores, and for test score rank within a cohort. The chosen methodological approach provided definitive answers to the questions posed in Chapter One.

All samples used in the study consisted of test data derived from students who had taken the third grade test and the ACT at the same school district. Each school's data were analyzed as a distinct cohort for two reasons: The different third grade tests administered in each district, and the ability to make a statistical comparison of the consistency of rank within each cohort. The following third grade standardized tests used by school districts in the study are commonly used norm-referenced achievement tests (Fair Test: The National Center for Fair Open Testing, 2006).

School A - The Stanford Achievement Test Series (9th Edition) with Otis-Lennon School Ability Test (7th Edition).

School B - Comprehensive Tests of Basic Skills (4th Edition).

School C - TerraNova CTBS Complete Battery (1996).

Third grade test score data were available in the areas of Reading, Language, and Mathematics from each of the three schools. Additional scores for School A were available in these areas: Using Information, Thinking Skills, Complete Battery, and School Ability Index. Other School B scores used in the study were Word Analysis, Spelling, and Total Battery.

Normal Curve Equivalent scores were used to assure a common measurement of achievement on all third grade standardized tests. All third grade tests were administered during the spring of 1997.

Summary

Correlations for third grade subtest scores and ACT Composite scores were computed for all available data. Subtest scores from third grade tests that were found to have significant correlational coefficients in relation to ACT Composite scores were combined to generate higher correlations than individual subtest scores for each school. Correlations for rank within a cohort were also computed. Scatterplots of each relationship were examined and supported the initial correlational analysis.

The chosen methodological approach allowed the researcher to provide definitive answers to the questions posed in Chapter

One. Presentation and analysis of the data continues in Chapter Four.

CHAPTER FOUR

Presentation and Analysis of Data

Introduction

The results of the statistical analyses outlined in Chapter Three are presented, interpreted, and analyzed. Data from three Michigan public schools were analyzed in three distinct cohorts. Each school administered a norm-referenced third grade standardized test in the spring of 1997.

The purposes of this study were two-fold: first, to determine if there was a significant statistical relationship between third grade standardized test scores and high school ACT Composite scores; and second, to determine if individual students (within a cohort) maintained their relative rank as measured by their scores on third grade standardized tests and ACT Composite.

Spreadsheets containing the collected data for each cohort were analyzed. The spreadsheets contained relevant third grade standardized test score factors as expressed in Normal Curve Equivalent (NCE) scores (range 1-99) and ACT Composite scores (range 1-36). Each row represented the scores of an individual student.

Pearson product-moment correlation coefficients (r) were computed using the SPSS program to determine statistical relationships. Correlational coefficients were expressed in decimal numbers ranging from -1.00 to 1.00. Correlations

significant at the 0.01 level (2-tailed) indicate a very low probability that the result is due to chance.

The statistical software (SPSS) had the ability to generate multiple regressions, which used the best combination of predictor variables to generate the highest predictive power as expressed by a correlational coefficient. The square of the correlation coefficient (r^2), known as the coefficient of determination, indicated the degree (generally expressed by percentage) of the relationship between two variables. Multiple regressions allowed for the production of prediction equations. However, it was not the goal of this study to predict individual scores; thus, predictive equations were not generated.

Relationship of Third Grade Tests and ACT Composite

The data were analyzed to determine the relationship between third grade standardized test scores and high school ACT Composite scores. Correlation coefficients were computed for each of the third grade test factors and ACT Composite scores, and for individual rank within each cohort. Multiple regressions combining third grade standardized test score factors were generated. Coefficients of determination were generated for each multiple regression.

Data for each cohort included a spreadsheet of test scores, tables showing the correlation of test variables, corresponding scatterplots, and multiple regressions given in table form, including r^2.

School A Test Relationships

School A administered the Stanford Achievement Test Series, Ninth Edition with Otis-Lennon Ability Test, Seventh Edition (SAT/9) in May, 1997. Definitions for the columns of the spreadsheet of test data for School A (Appendix A) were given these titles: Reading3, Mathematics3, Language3, Using Information3, Thinking3, Complete Battery3, School Ability Index3, and ACT Composite.

Positive correlations for all third grade SAT/9 test factors as related to the ACT Composite score for School A are shown in Table 1. The correlations ranged from .656 to .835, indicating a strong relationship between the third grade SAT/9 and the ACT Composite score. Complete Battery3 (.835) and Thinking3 (.818) were extremely strong individual predictor variables of ACT Composite. All correlations were significant at the 0.01 level (2-tailed).

Table 2

School A Correlations: ACT Composite vs. 3rd Grade SAT/9

Correlations - School A

		READ3	MATH3	LANG3	USEINFO3	THINK3	COMPBAT3	SAI3	ACTCOMP
READ3	Pearson Correlati	1	.618**	.728**	.761**	.899**	.886**	.592**	.754**
	Sig. (2-tailed)	.	.000	.000	.000	.000	.000	.000	.000
	N	45	45	45	45	45	45	45	45
MATH3	Pearson Correlati	.618**	1	.707**	.837**	.780**	.840**	.771**	.765**
	Sig. (2-tailed)	.000	.	.000	.000	.000	.000	.000	.000
	N	45	45	45	45	45	45	45	45
LANG3	Pearson Correlati	.728**	.707**	1	.707**	.867**	.834**	.651**	.656**
	Sig. (2-tailed)	.000	.000	.	.000	.000	.000	.000	.000
	N	45	45	45	45	45	45	45	45
USEINFO3	Pearson Correlati	.761**	.837**	.707**	1	.874**	.907**	.780**	.779**
	Sig. (2-tailed)	.000	.000	.000	.	.000	.000	.000	.000
	N	45	45	45	45	45	45	45	45
THINK3	Pearson Correlati	.899**	.780**	.867**	.874**	1	.964**	.749**	.818**
	Sig. (2-tailed)	.000	.000	.000	.000	.	.000	.000	.000
	N	45	45	45	45	45	45	45	45
COMPBAT	Pearson Correlati	.886**	.840**	.834**	.907**	.964**	1	.725**	.835**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.	.000	.000
	N	45	45	45	45	45	45	45	45
SAI3	Pearson Correlati	.592**	.771**	.651**	.780**	.749**	.725**	1	.743**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.	.000
	N	45	45	45	45	45	45	45	45
ACTCOMP	Pearson Correlati	.754**	.765**	.656**	.779**	.818**	.835**	.743**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.
	N	45	45	45	45	45	45	45	45

**. Correlation is significant at the 0.01 level (2-tailed).

Scatterplots shown in Figures 1-7 confirmed the test correlations for School A.

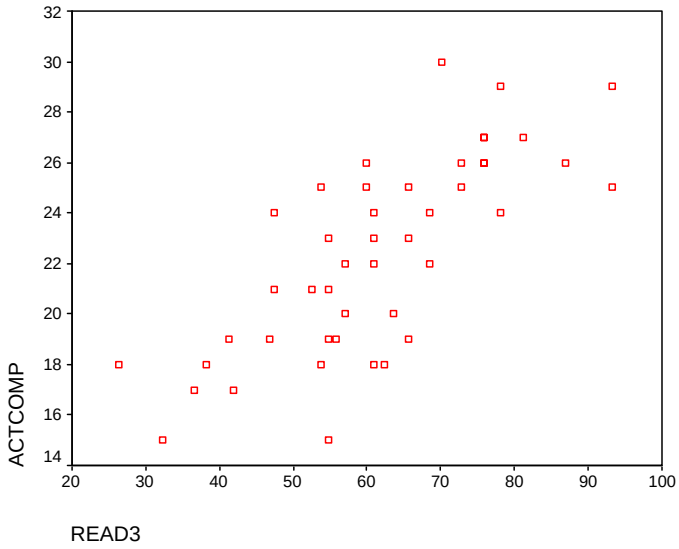

Figure 1. ACT Composite vs. Reading3 (School A)

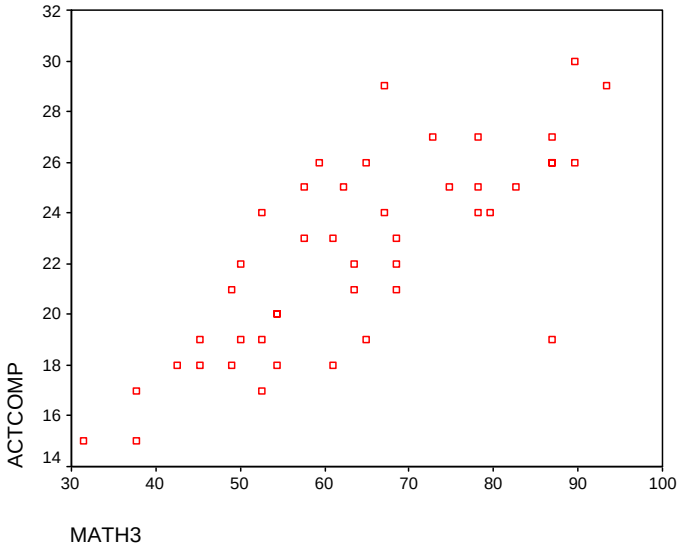

Figure 2. ACT Composite vs. Mathematics3 (School A)

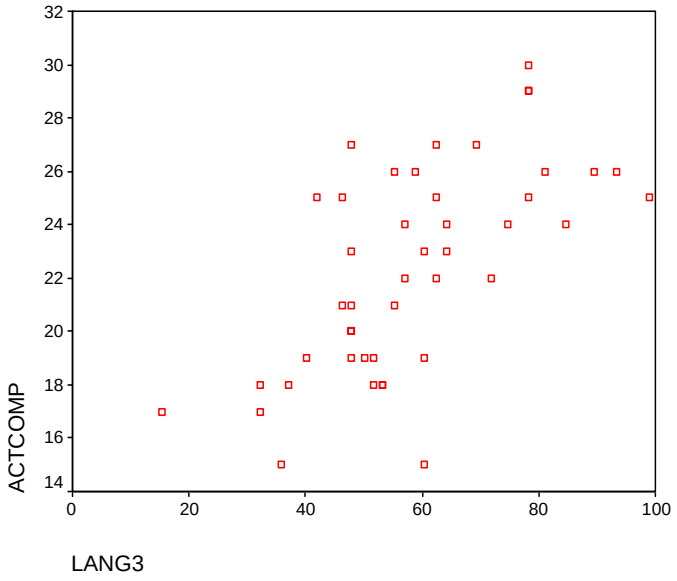

Figure 3. ACT Composite vs. Language3 (School A)

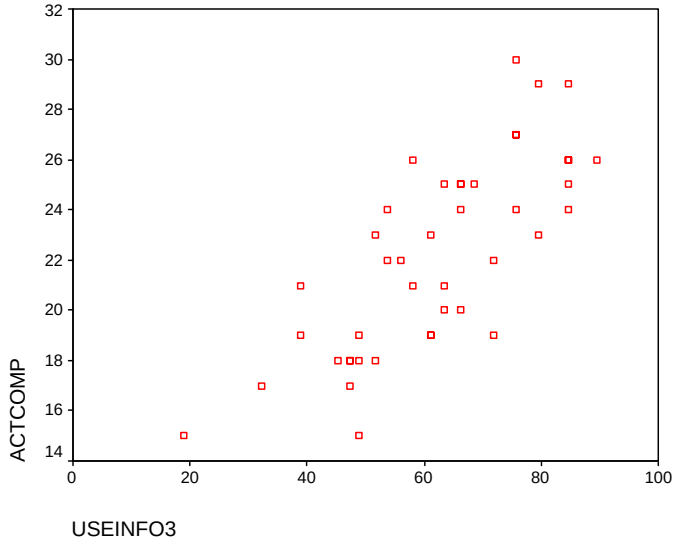

Figure 4. ACT Composite vs. Using Information3 (School A)

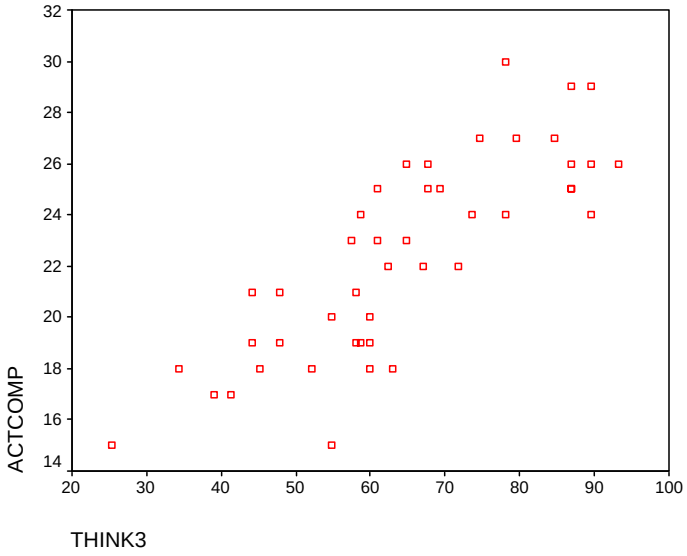

Figure 5. ACT Composite vs. Thinking3 (School A)

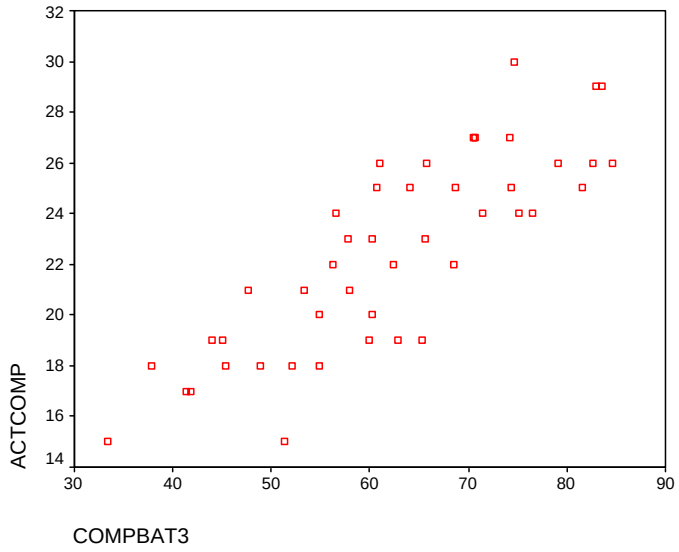

Figure 6. ACT Composite vs. Complete Battery3 (School A)

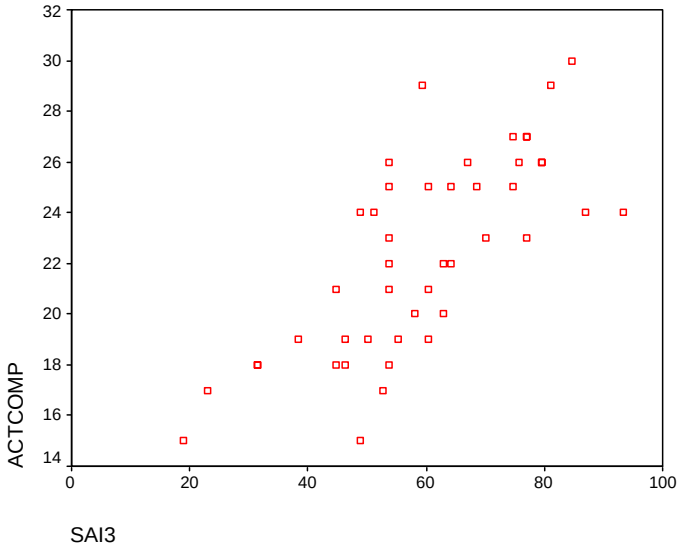

Figure 7. ACT Composite vs. School Ability Index3 (School A)

A multiple regression for School A, shown in Table 2, combining all third grade SAT/9 test factors, produced a correlation of .872, which, when squared, produced a coefficient of determination of .761. Thus, 76.1% of variability of ACT Composite scores could be attributed to the third grade SAT/9 for School A.

Table 3

School A Regression: ACT Composite vs. 3rd Grade SAT/9

Model Summary[b]

Model	r	r²	Adjusted r²	Std. Error of the Estimate
1	.872[a]	.761	.716	2.06354

a Predictors: (Constant), SAI3, READ3, LANG3, MATH3, USEINFO3,THINK3,COMPBAT3

b Dependent Variable: ACTCOMP

School B Test Relationships

School B administered the Comprehensive Tests of Basic Skills, Fourth Edition (CTBS/4) in May, 1997. Definitions for the columns of the spreadsheet of test data for School B (Appendix B) were given these titles: Reading3, Language3, Mathematics3, Total Battery3, Word Analysis3, Spelling3, and ACT Composite.

Positive correlations for all third grade CTBS/4 test factors as related to the ACT Composite score for School B are shown in Table 3. The correlations ranged from .387 to .888, with four of six factors indicating a strong relationship

between the third grade CTBS/4 and ACT Composite. Reading3 (.888) and Total Battery3 (.858) were extremely strong individual predictor variables of ACT Composite. All correlations were significant at the 0.01 level (2-tailed) except Word Analysis3.

Table 4

School B Correlations: ACT Composite vs. 3rd Grade CTBS/4

Correlations - School B

		READ3	LANG3	MATH3	TOTBATT3	WORDANA3	SPELL3	ACTCOMP
READ3	Pearson Correlati	1	.579**	.621**	.820**	.326	.603**	.888**
	Sig. (2-tailed)	.	.004	.002	.000	.129	.002	.000
	N	23	23	23	23	23	23	23
LANG3	Pearson Correlati	.579**	1	.738**	.856**	.509*	.693**	.643**
	Sig. (2-tailed)	.004	.	.000	.000	.013	.000	.001
	N	23	23	23	23	23	23	23
MATH3	Pearson Correlati	.621**	.738**	1	.932**	.424*	.700**	.706**
	Sig. (2-tailed)	.002	.000	.	.000	.044	.000	.000
	N	23	23	23	23	23	23	23
TOTBATT3	Pearson Correlati	.820**	.856**	.932**	1	.477*	.767**	.858**
	Sig. (2-tailed)	.000	.000	.000	.	.021	.000	.000
	N	23	23	23	23	23	23	23
WORDANA	Pearson Correlati	.326	.509*	.424*	.477*	1	.753**	.387
	Sig. (2-tailed)	.129	.013	.044	.021	.	.000	.068
	N	23	23	23	23	23	23	23
SPELL3	Pearson Correlati	.603**	.693**	.700**	.767**	.753**	1	.597**
	Sig. (2-tailed)	.002	.000	.000	.000	.000	.	.003
	N	23	23	23	23	23	23	23
ACTCOMP	Pearson Correlati	.888**	.643**	.706**	.858**	.387	.597**	1
	Sig. (2-tailed)	.000	.001	.000	.000	.068	.003	.
	N	23	23	23	23	23	23	23

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Scatterplots shown in Figures 8-13 confirmed the test correlations for School B.

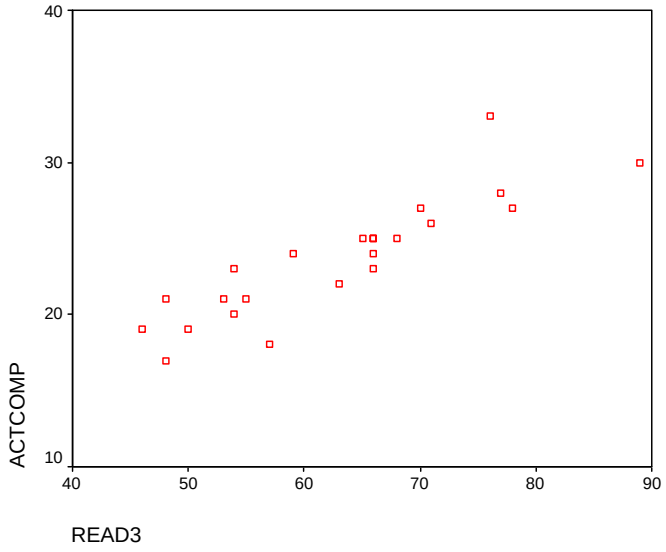

Figure 8. ACT Composite vs. Reading3 (School B)

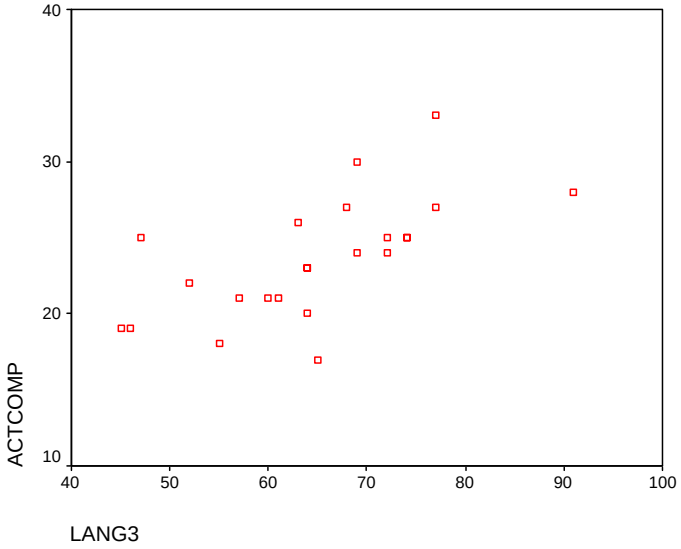

Figure 9. ACT Composite vs. Language3 (School B)

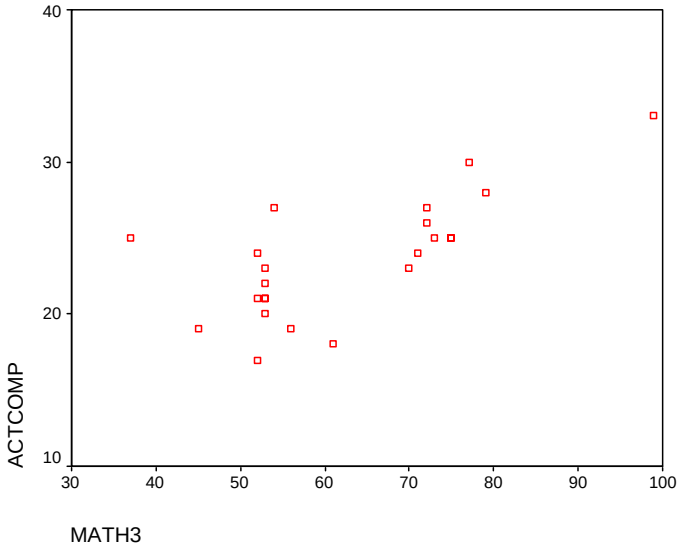

Figure 10. ACT Composite vs. Mathematics3 (School B)

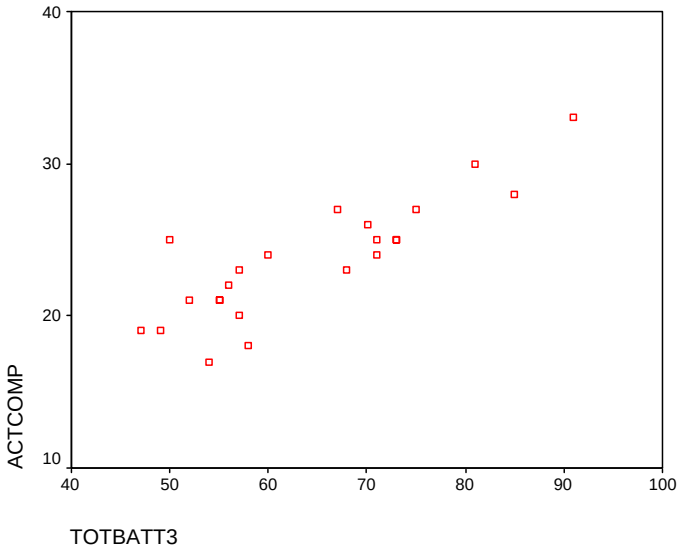

Figure 11. ACT Composite vs. Total Battery3 (School B)

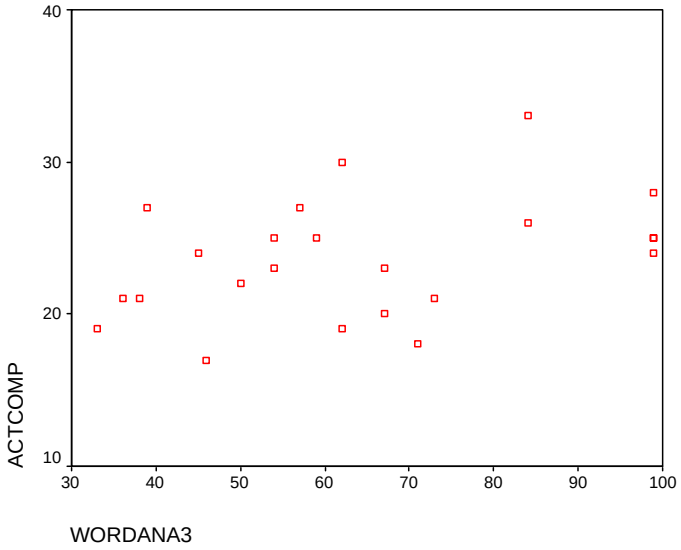

Figure 12. ACT Composite vs. Word Analysis3 (School B)

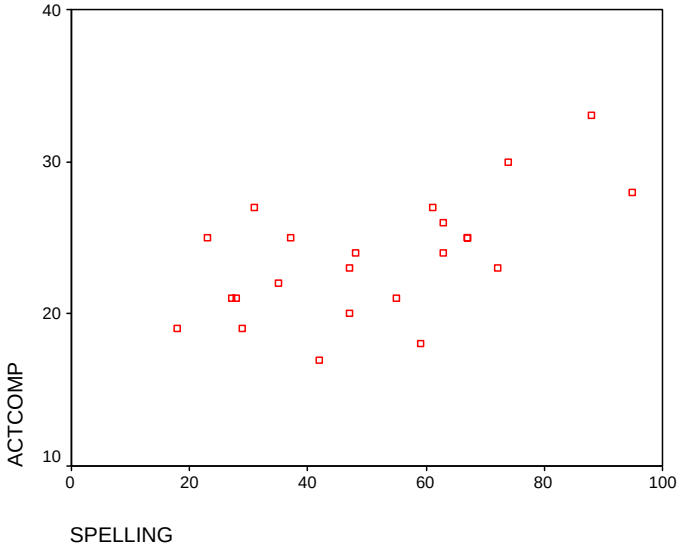

Figure 13. ACT Composite vs. Spelling3 (School B)

A multiple regression for School B, shown in Table 4, combining all third grade CTBS/4 test factors, produced a correlation of .953, which, when squared, produced a coefficient of determination of .908. Thus, 90.8% of variability of ACT

Composite scores could be attributed to the third grade CTBS/4 for School B.

Table 5

School B Regression: ACT Composite vs. 3rd Grade CTBS/4

Model Summary[b]

Model	R	r²	Adjusted r²	Std Error of the Estimate
1	.953[a]	.908	.874	1.39438

a Predictors: (Constant), SPELL3, READ3, LANG3, WORDANA3, MATH3, TOTBATT3

b Dependent Variable: ACTCOMP

School C Test Relationships

School C administered the TerraNova CTBS Complete Battery (TN-CTBS) in April, 1997. Definitions for the columns of the spreadsheet of test data for School C (Appendix C) were given these titles: Reading3, Language3, Mathematics3, and ACT Composite. School C test data for the TN-CTBS contained only three third grade test factors and lacked any overall cumulative test measurement.

Positive correlations for all third grade TN-CTBS test factors as related to the ACT Composite score for School C are shown in Table 5. The correlations for each test factor, .697, .647, and .654, indicated a firm relationship between the third grade TN-CTBS and ACT Composite score. All correlations were significant at the 0.01 level (2-tailed).

Table 6

School C Correlations: ACT Composite vs. 3rd Grade TN-CTBS

Correlations - School C

		READ3	LANG3	MATH3	ACTCOMP
READ3	Pearson Correlation	1	.700**	.719**	.697**
	Sig. (2-tailed)	.	.000	.000	.000
	N	182	182	182	182
LANG3	Pearson Correlation	.700**	1	.682**	.647**
	Sig. (2-tailed)	.000	.	.000	.000
	N	182	182	182	182
MATH3	Pearson Correlation	.719**	.682**	1	.654**
	Sig. (2-tailed)	.000	.000	.	.000
	N	182	182	182	182
ACTCOMP	Pearson Correlation	.697**	.647**	.654**	1
	Sig. (2-tailed)	.000	.000	.000	.
	N	182	182	182	182

**. Correlation is significant at the 0.01 level (2-tailed).

Scatterplots shown in Figures 14-16 confirmed the test correlations for School C.

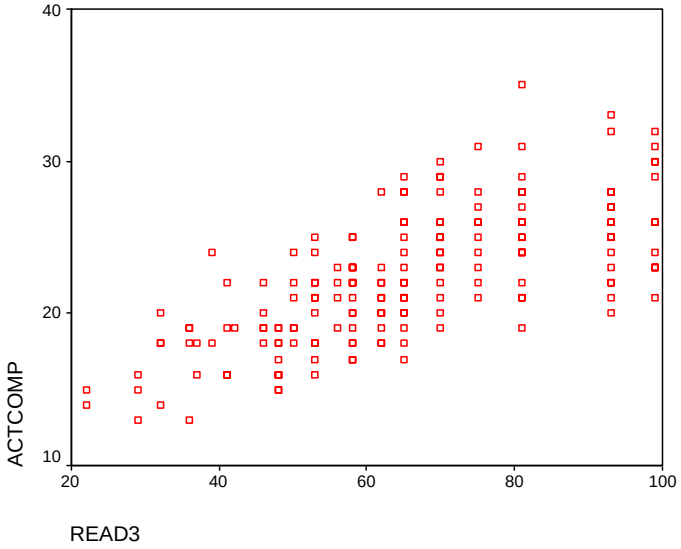

Figure 14. ACT Composite vs. Reading3 (School C)

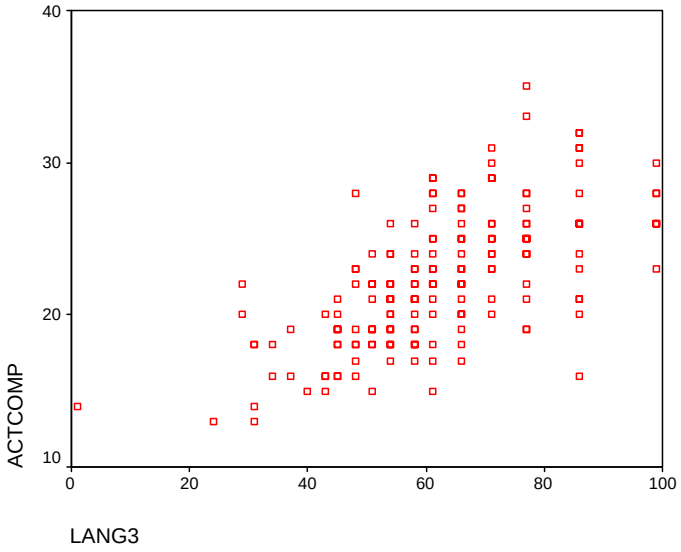

Figure 15. ACT Composite vs. Language3 (School C)

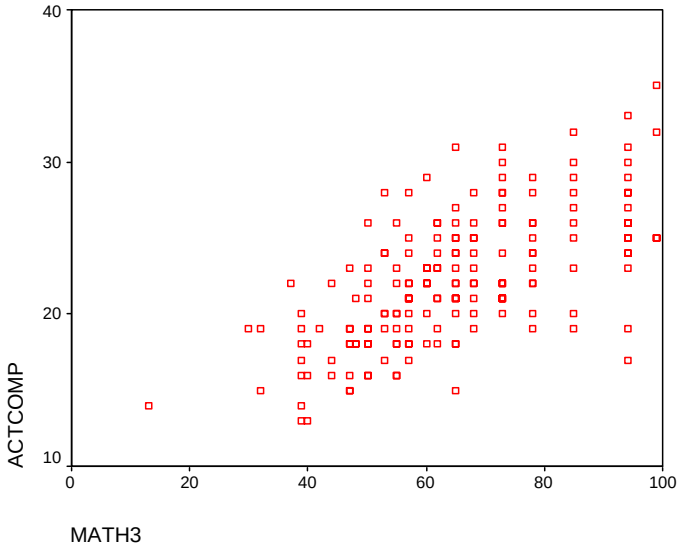

Figure 16. ACT Composite vs. Mathematics3 (School C)

A multiple regression for School C, shown in Table 6, combining all third grade TN-CTBS test factors, produced a correlation of .747, which, when squared, produced a coefficient of determination of .558. Thus, 55.8% of variability of ACT Composite scores could be attributed to the third grade TN-CTBS for School C.

Table 7

School C Regression: ACT Composite vs. 3rd Grade TN-CTBS

Model Summary[b]

Model	R	r^2	Adjusted r^2	Std. Error of the Estimate
1	.747[a]	.558	.550	2.91074

a Predictors: (Constant), MATH3, LANG3, READ3

b Dependent Variable: ACTCOMP

Rank Relationship of Third Grade Tests and ACT Composite

The SPSS program allows for each category of data to be ranked. Spreadsheets were created from the rank data for each school. The rank data were analyzed to determine if individual students (within a cohort) maintained their relative rank as measured by their scores on third grade standardized tests and high school ACT Composite. Correlation coefficients were computed for each of the third grade test rank factors and ACT Composite ranks.

Multiple regressions combining 3rd grade test rank factors were generated. Coefficients of determination were generated for each multiple regression.

Data for each cohort include a spreadsheet of test ranks, tables showing the correlation of rank variables, corresponding scatterplots, and multiple regressions given in table form, including r^2.

School A Rank Relationships

Definitions for the columns of the spreadsheet of rank data for School A (Appendix D) were given these titles: Rank of Reading3, Rank of Mathematics3, Rank of Language3, Rank of Using Information3, Rank of Thinking3, Rank of Complete Battery3, Rank of School Ability Index3, and Rank of ACT Composite.

Positive correlations for all third grade SAT/9 rank factors as related to the ACT Composite rank for School A are shown in Table 7. The correlations ranged from .646 to .847, indicating a strong relationship between the third grade SAT/9 rank and ACT Composite rank. Rank of Complete Battery3 (.847) and Rank of Thinking3 (.837) were extremely strong individual predictor variables of ACT rank. All correlations were significant at the 0.01 level (2-tailed).

Table 8

School A Rank Correlations: ACT Composite vs. 3rd Grade SAT/9

Rank Correlations - School A

		RANK of READ3	RANK of MATH3	RANK of LANG3	RANK of USEINFO3	RANK of THINK3	RANK of COMPBAT3	RANK of SAI3	RANK of ACTCOMP
RANK of READ3	Pearson Correlati	1	.591**	.706**	.742**	.892**	.852**	.610**	.762**
	Sig. (2-tailed)	.	.000	.000	.000	.000	.000	.000	.000
	N	45	45	45	45	45	45	45	45
RANK of MATH3	Pearson Correlati	.591**	1	.716**	.852**	.759**	.852**	.782**	.761**
	Sig. (2-tailed)	.000	.	.000	.000	.000	.000	.000	.000
	N	45	45	45	45	45	45	45	45
RANK of LANG3	Pearson Correlati	.706**	.716**	1	.707**	.846**	.812**	.674**	.646**
	Sig. (2-tailed)	.000	.000	.	.000	.000	.000	.000	.000
	N	45	45	45	45	45	45	45	45
RANK of USEINF	Pearson Correlati	.742**	.852**	.707**	1	.859**	.914**	.782**	.798**
	Sig. (2-tailed)	.000	.000	.000	.	.000	.000	.000	.000
	N	45	45	45	45	45	45	45	45
RANK of THINK3	Pearson Correlati	.892**	.759**	.846**	.859**	1	.951**	.753**	.837**
	Sig. (2-tailed)	.000	.000	.000	.000	.	.000	.000	.000
	N	45	45	45	45	45	45	45	45
RANK of COMPB	Pearson Correlati	.852**	.852**	.812**	.914**	.951**	1	.729**	.847**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.	.000	.000
	N	45	45	45	45	45	45	45	45
RANK of SAI3	Pearson Correlati	.610**	.782**	.674**	.782**	.753**	.729**	1	.753**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.	.000
	N	45	45	45	45	45	45	45	45
RANK of ACTCOM	Pearson Correlati	.762**	.761**	.646**	.798**	.837**	.847**	.753**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.
	N	45	45	45	45	45	45	45	45

**.Correlation is significant at the 0.01 level (2-tailed).

Scatterplots shown in Figures 17-23 confirmed the test rank correlations for School A.

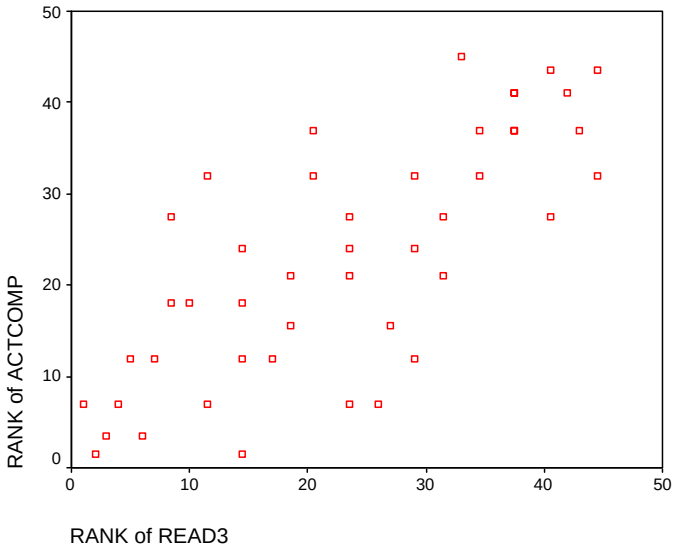

Figure 17. ACT Comp. Rank vs. Reading3 Rank (School A)

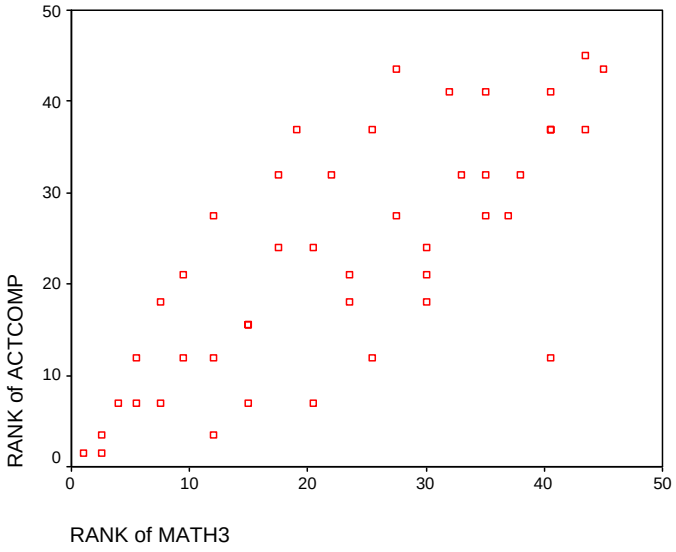

Figure 18. ACT Comp. Rank vs. Mathematics3 Rank (School A)

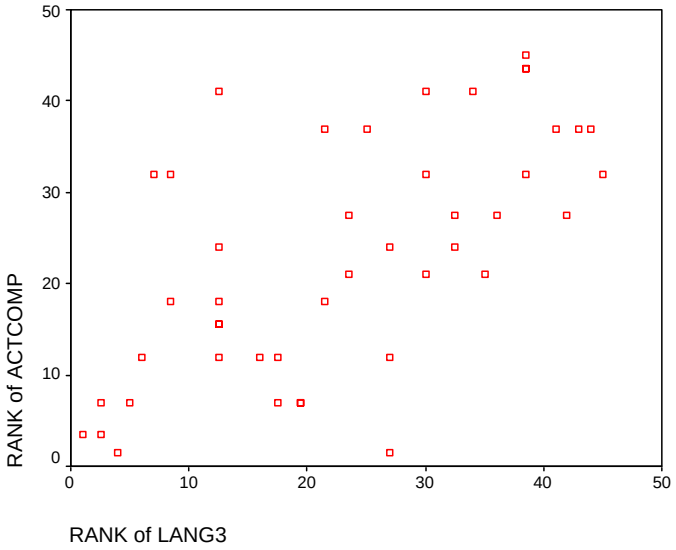

Figure 19. ACT Comp. Rank vs. Language3 Rank (School A)

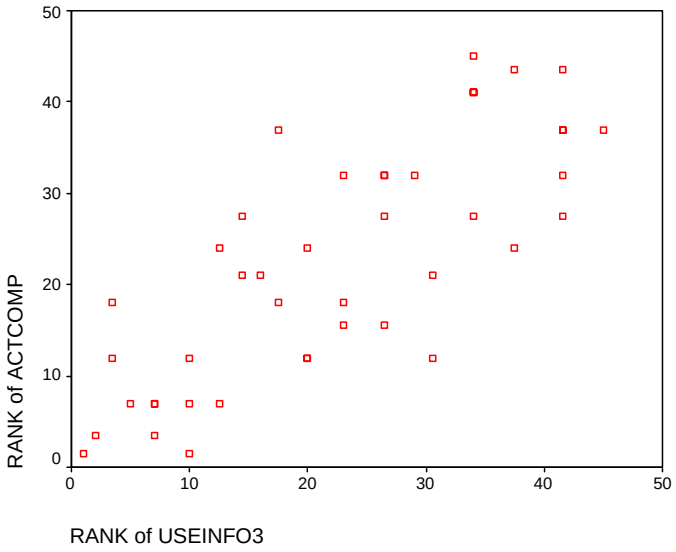

Figure 20. ACT Comp. Rank vs. Using Inform3 Rank (School A)

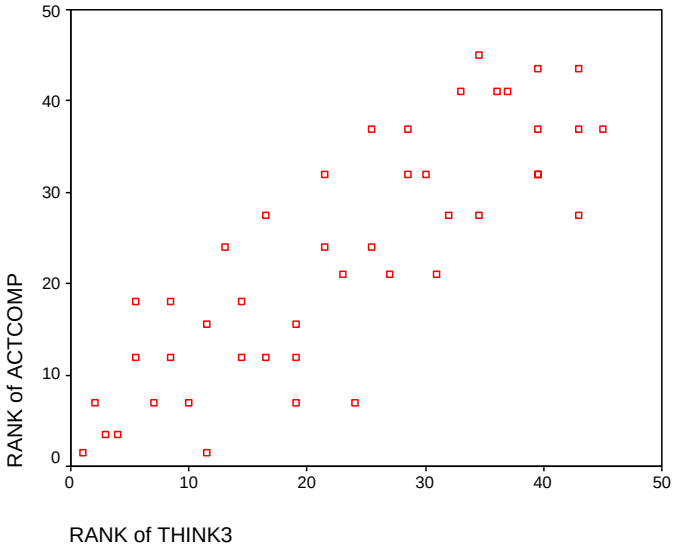

Figure 21. ACT Comp. Rank vs. Thinking3 Rank (School A)

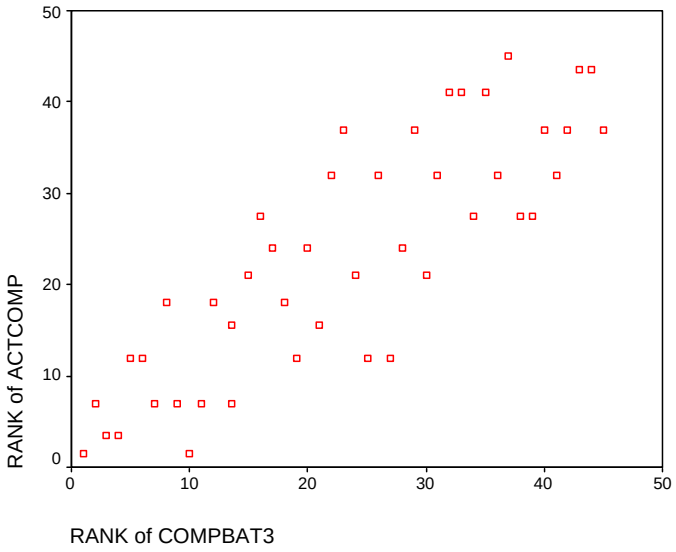

Figure 22. ACT Comp. Rank vs. Complete Battery3 Rank (School A)

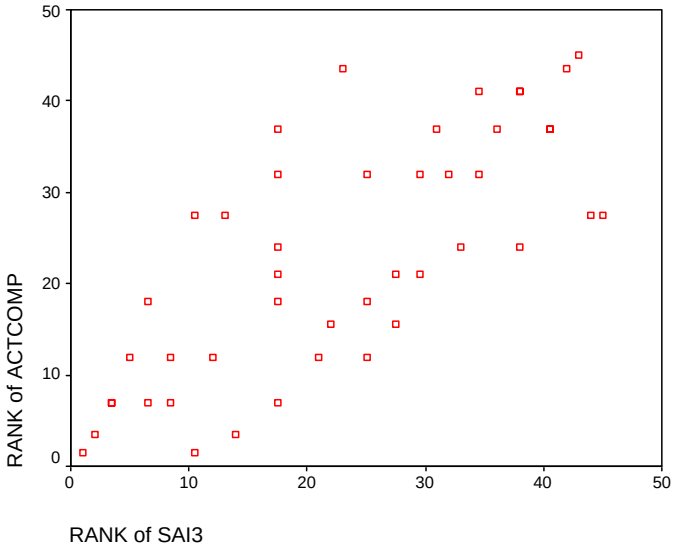

Figure 23. ACT Comp. Rank vs. School Ability In3 Rank (School A)

A multiple regression for School A rank, shown in Table 8, combining all third grade SAT/9 test rank factors, produced a correlation of .887, which, when squared, produced a coefficient of determination of .786. Thus, 78.6% of variability of ACT Composite rank could be attributed to the third grade SAT/9 rank for School A.

Table 9

School A Rank Regression: ACT Composite vs. 3rd Grade SAT/9

Model Summary[b]

Model	R	r²	Adjusted r²	Std. Error of the Estimate
1	.887	.786	.746	6.600173

a Predictors: (Constant), RANK of SAI3, RANK of READ3, RANK of LANG3, RANK of MATH3, RANK of USEINFO3, RANK of THINK3, RANK of COMPBAT3

b Dependent Variable: RANK of ACTCOMP

School B Test Rank Relationships

Definitions for the columns of the spreadsheet of rank data for School B (Appendix E) were given these titles: Rank of Reading3, Rank of Language3, Rank of Mathematics3, Rank of Total Battery3, Rank of Word Analysis3, Rank of Spelling3, and Rank of ACT Composite.

Positive correlations for all third grade CTBS/4 rank factors as related to the ACT Composite rank for School B are shown in Table 9. The correlations ranged from .361 to .917, with four of six factors indicating a strong relationship between the third grade CTBS/4 and ACT Composite. Reading3 (.917) and Total Battery3 (.804) were extremely strong

individual predictor variables of ACT Composite rank. All correlations were significant at the 0.01 level (2-tailed) except Word Analysis3.

Table 10

School B Rank Correlations: ACT Composite vs. 3rd Grade CTBS/4

Rank Correlations - School B

		RANK of READ3	RANK of LANG3	RANK of MATH3	RANK of TOTBATT3	RANK of WORDANA3	RANK of SPELL3	RANK of ACTCOMP
RANK of READ3	Pearson Correlation	1	.599**	.634**	.789**	.369	.644**	.917**
	Sig. (2-tailed)	.	.003	.001	.000	.083	.001	.000
	N	23	23	23	23	23	23	23
RANK of LANG3	Pearson Correlation	.599**	1	.713**	.889**	.432*	.634**	.694**
	Sig. (2-tailed)	.003	.	.000	.000	.040	.001	.000
	N	23	23	23	23	23	23	23
RANK of MATH3	Pearson Correlation	.634**	.713**	1	.887**	.310	.587**	.672**
	Sig. (2-tailed)	.001	.000	.	.000	.150	.003	.000
	N	23	23	23	23	23	23	23
RANK of TOTBATT3	Pearson Correlation	.789**	.889**	.887**	1	.493*	.760**	.804**
	Sig. (2-tailed)	.000	.000	.000	.	.017	.000	.000
	N	23	23	23	23	23	23	23
RANK of WORDANA	Pearson Correlation	.369	.432*	.310	.493*	1	.745**	.361
	Sig. (2-tailed)	.083	.040	.150	.017	.	.000	.091
	N	23	23	23	23	23	23	23
RANK of SPELL3	Pearson Correlation	.644**	.634**	.587**	.760**	.745**	1	.560**
	Sig. (2-tailed)	.001	.001	.003	.000	.000	.	.005
	N	23	23	23	23	23	23	23
RANK of ACTCOMP	Pearson Correlation	.917**	.694**	.672**	.804**	.361	.560**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.091	.005	.
	N	23	23	23	23	23	23	23

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Scatterplots shown in Figures 24-29 confirmed the test rank correlations for School B.

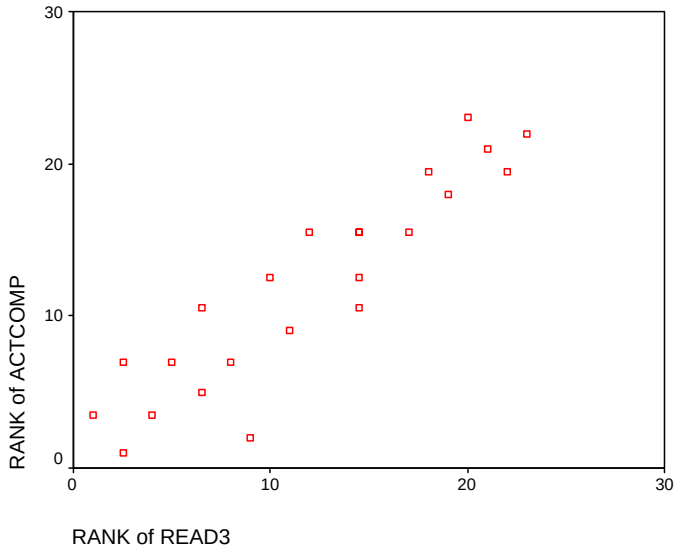

Figure 24. ACT Comp. Rank vs. Reading3 Rank (School B)

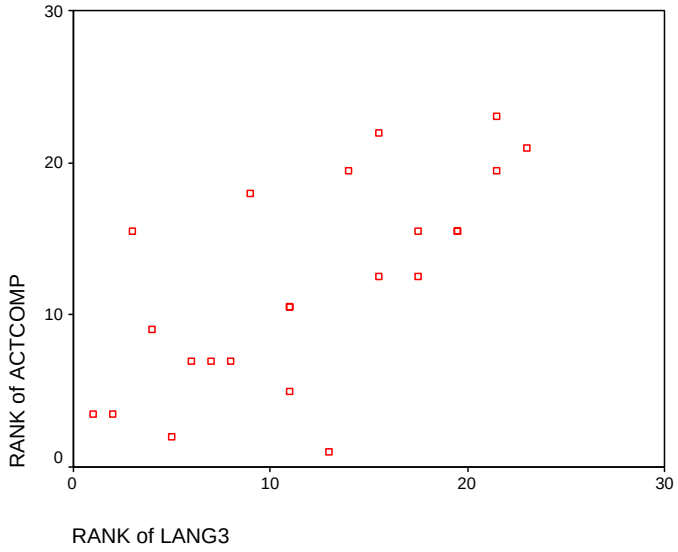

Figure 25. ACT Comp. Rank vs. Language3 Rank (School B)

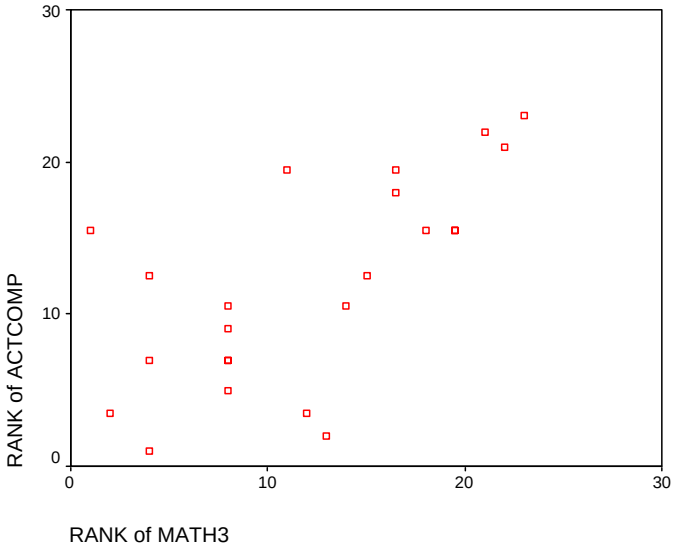

Figure 26. ACT Comp. Rank vs. Mathematics3 Rank (School B)

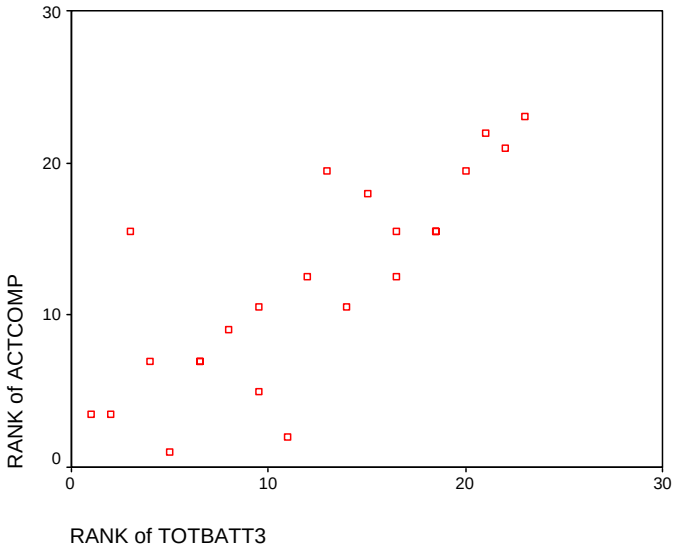

Figure 27. ACT Comp. Rank vs. Total Battery3 Rank (School B)

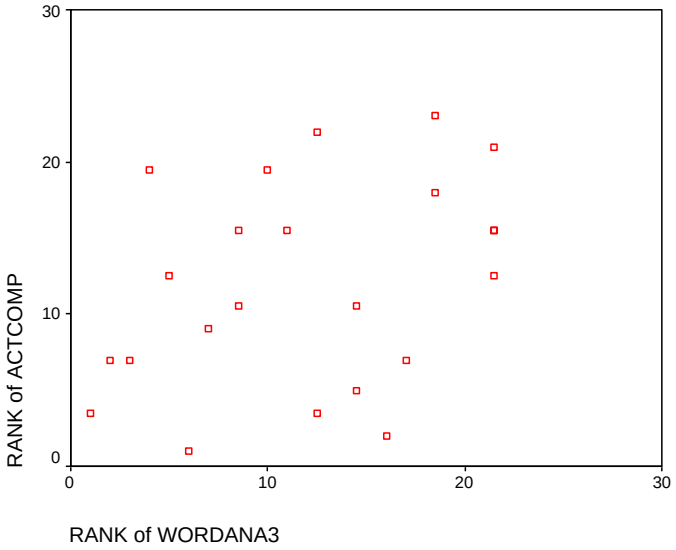

Figure 28. ACT Comp. Rank vs. Word Analysis3 Rank (School B)

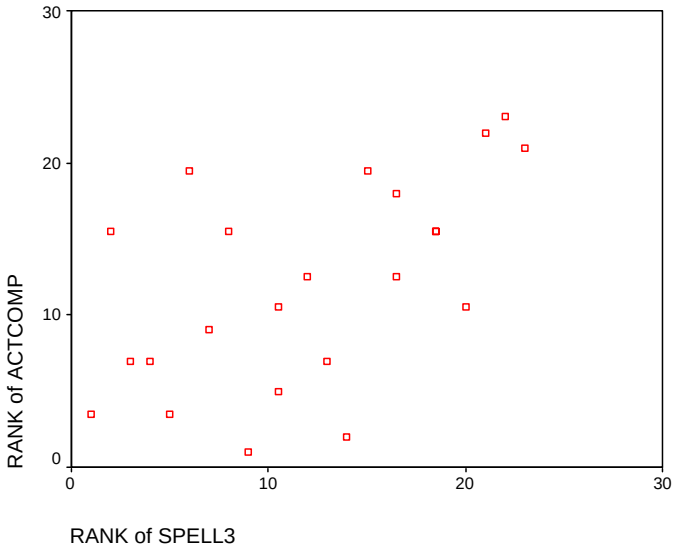

Figure 29. ACT Comp. Rank vs. Spelling3 Rank (School B)

A multiple regression for School B rank, shown in Table 10, combining all third grade CTBS/4 test rank factors, produced a correlation of .949, which, when squared, produced a coefficient of determination of .900. Thus, 90.0% of variability of ACT

Composite rank could be attributed to the third grade CTBS/4 rank for School B.

Table 11

School B Rank Regression: ACT Composite vs. 3rd Grade CTBS/4

Model Summary[b]

Model	r	r²	Adjusted r²	Std. Error of the Estimate
1	.949[a]	.900	.862	2.505478

a Predictors: Constant), RANK of SPELL3, RANK of MATH3, RANK of READ3, RANK of LANG3, RANK of WORDANA3, RANK of TOTBATT3

b Dependent Variable: RANK of ACTCOMP

School C Test Rank Relationships

Definitions for the columns of the spreadsheet of rank data for School C (Appendix F) were given these titles: Rank of Reading3, Rank of Language3, Rank of Mathematics3, and Rank of ACT Composite.

Positive correlations for all third grade TN-CTBS rank factors as related to the ACT Composite rank for School C are shown in Table 11. The correlations for each test rank factor, .723, .667, and .659, indicated a firm relationship between the third grade TN-CTBS rank and ACT Composite rank. All correlations were significant at the 0.01 level (2-tailed).

Table 12

School C Rank Correlations: ACT Composite vs. 3rd Grade TN-CTBS

Rank Correlations - School C

		RANK of READ3	RANK of LANG3	RANK of MATH3	RANK of ACTCOMP
RANK of READ3	Pearson Correlation	1	.711**	.733**	.723**
	Sig. (2-tailed)	.	.000	.000	.000
	N	182	182	182	182
RANK of LANG3	Pearson Correlation	.711**	1	.676**	.667**
	Sig. (2-tailed)	.000	.	.000	.000
	N	182	182	182	182
RANK of MATH3	Pearson Correlation	.733**	.676**	1	.659**
	Sig. (2-tailed)	.000	.000	.	.000
	N	182	182	182	182
RANK of ACTCOMP	Pearson Correlation	.723**	.667**	.659**	1
	Sig. (2-tailed)	.000	.000	.000	.
	N	182	182	182	182

**. Correlation is significant at the 0.01 level (2-tailed).

Scatterplots shown in Figures 30-32 confirmed the test rank correlations for School C.

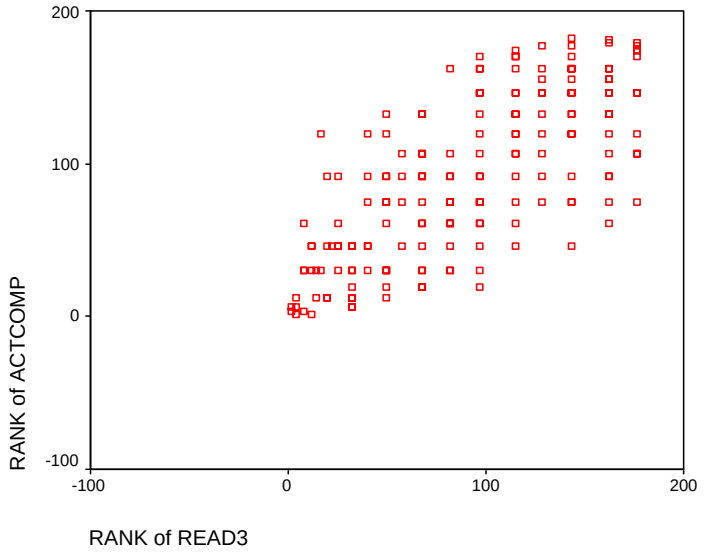

Figure 30. ACT Comp. Rank vs. Reading3 Rank (School C)

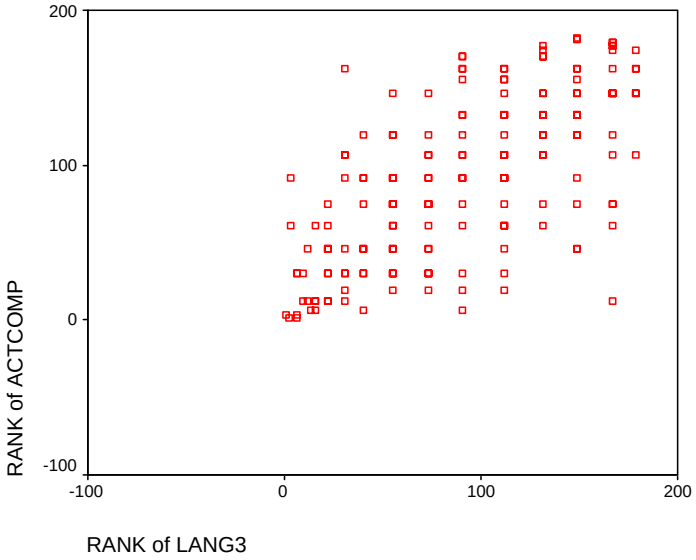

Figure 31. ACT Comp. Rank vs. Language3 Rank (School C)

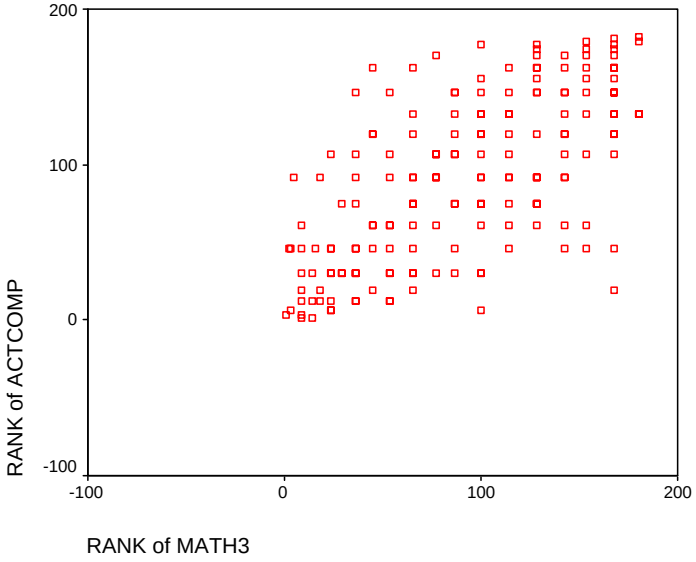

Figure 32. ACT Comp. Rank vs. Mathematics3 Rank (School C)

A multiple regression for School C rank, shown in Table 12, combining all third grade TN-CTBS test rank factors, produced a correlation of .766, which, when squared, produced a coefficient of determination of .586. Thus, 58.6% of variability of ACT

Composite rank could be attributed to the third grade TN-CTBS for School C.

Table 13

School C Rank Regression: ACT Composite vs. 3rd Grade TN-CTBS

Model Summary[b]

Model	r	r²	Adjusted r²	Std Error of the Estimate
1	.766[a]	.586	.579	34.079924

a Predictors: (Constant), RANK of MATH3, RANK of LANG3, RANK of READ3
b Dependent Variable: RANK of ACTCOMP

Correlations greater than .60 are rarely obtained in educational research (Fraenkel & Wallen, 1996). Most of the correlations and all of the regressions computed in this study were above .60 and were significant at the 0.01 level (two-tailed).

The statistical analysis of this study demonstrated a strong relationship between third grade standardized test scores and ACT Composite scores. A strong relationship was shown for cohort rank from third grade to high school, as measured by these tests. Chapter Five provides a summary of the results of this study and offers conclusions, inferences, implications, and recommendations for further research and study.

CHAPTER FIVE

Conclusions, Inferences, Implications, Recommendations for Further Research, and Summary

Introduction

The first purpose of this study was to determine the relationship between third grade standardized test scores and high school ACT Composite scores of individual students. The second purpose of the study was to determine the consistency of individual rank within a cohort as measured by standardized test scores in third grade and ACT Composite scores. Test score data from the Class of 2006 in three Michigan public schools was analyzed to determine these relationships. The schools varied in size and demographic make-up.

This researcher discovered high levels of correlation between the scores of three different, commonly used third grade standardized achievement tests and high school ACT Composite scores. The statistical tool of multiple regression using various sub-scores from the third grade tests was used to discover the strongest correlation to the ACT composite for each school district. Strong, positive correlations were found between third grade standardized test scores and ACT Composite scores. Strong, positive correlations were also found demonstrating consistent individual student rank within each tested cohort as measured by third grade test scores and high school ACT Composite scores.

Fraenkel & Wallen (1996) stated, "Only when a correlation of .65 or higher is obtained can individual predictions that are reasonably accurate for most purposes be made" p. 318). The correlations in this study between third grade standardized test scores and ACT composite scores easily exceeded their threshold for accurate prediction. The correlations (.872, .953, and .747) found in this study demonstrated a strong, consistent relationship between three commonly used third grade standardized tests and ACT composite scores.

Squaring r for each district's scores yielded .761, .908, and .558 respectively. This meant that third grade test scores accounted for more than 76 percent (School A), 90 percent (School B), and 55 percent (School C) of the variance of ACT composite scores in high school.

Correlations of individual rank within each cohort were computed to confirm the strong relationship between third grade test scores and ACT composite scores. The correlations computed for rank were similar (.887, .949, and .766) and verified the consistency of individual test scores when compared within a cohort. Squaring r for each district's cohort rank meant that third grade rank accounted for more than 78 percent (School A), 90 percent (School B), and 58 percent (School C) of the variance of rank for ACT composite scores.

Conclusions

The correlations found by this researcher demonstrated strong relationships between consistency of standardized test scores and relative individual rank within a cohort as measured by standardized test scores. Students in this study achieved ACT Composite scores in high school largely predicated on their third grade standardized test scores. The transition from third grade through high school did not considerably alter academic achievement as measured by these tests. Students found their place in a hierarchy of achievement as measured by standardized tests during third grade and maintained this relative rank at the end of high school.

A direct interpretation of this research is that the ability to perform on standardized tests is determined by third grade and relatively inflexible throughout a student's academic career. This interpretation parallels research that showed consistent achievement gaps as measured by standardized tests between various subgroups of the student population (Adams, 1994; Barton, 2005; Maylone, 2002; Institute of Educational Sciences National Center for Educational Statistics, 2005; Thorndike, 1951).

Inferences and Implications

Despite continued concern about standardized test score gaps, they have remained consistent over time. In chapter one, the difficulty of meeting the demands of No Child Left Behind (NCLB) and its requirement of Adequate Yearly Progress (AYP)

were expressed. Standardized tests are used as a measuring system to make sure that students are making AYP and to monitor the quality of schooling delivered (Michigan Department of Education, 2006A; Popham, 2001). Continually changing federal and state standards create an inconsistent measuring system, and the frequent shift of political winds has turned achievement standards into moving targets. The state superintendent of education in Michigan stated that it will be difficult to show definitive trends in achievement because the educational system will continue to be refined and adjusted at both the federal and state levels (Chambers, 2005).

Evidence was presented in Chapter Two that showed that standardized test scores are primarily influenced by factors outside of school control. Berliner & Biddle (1998) considered poverty to be the single greatest barrier to high achievement in the American public schools. Given the penchant for using standardized tests as a measure of accountability for education, this study provided another reason to question this practice. Viewing standardized tests as a gauge of quality results in an inappropriate perception of schools. Maylone (2002) demonstrated a direct correlation between socioeconomic status and student achievement on standardized tests and concluded that it is inappropriate to judge or rank schools based on standardized test performance. Using the ACT as a part of the Michigan Merit Exam (MME) will continue to expose the achievement gaps between different racial, ethnic, and socioeconomic groups.

This researcher demonstrated the predictability of ACT scores as early as third grade and perhaps the strong relationship held by all standardized tests. Individual standardized test scores in this study were largely intractable. Given the climate of accountability, the results of this study should encourage many people to reflect on the use of standardized tests. They do not measure the quality of schools; they measure the demographics of schools. All of the pressures and sanctions of NCLB have not resulted in a significant rise in achievement or closed racial achievement gaps (Lee, 2006). There have been no gains in NAEP reading scores and no convincing evidence that the pressure associated with high-stakes testing leads to any important benefits for students' achievement (Nichols, Glass, & Berliner, 2005).

The insistence that all students will meet state proficiency standards flies in the face of the wide achievement gaps that have been consistently measured since the early 1970s (National Center for Education Statistics, 2005). Cialdani (2001) said, "There are certain disturbing things we simply would rather not realize" (p. 55). The NCLB goal of having all students proficient in reading and mathematics by 2014 will not be attained (Mathis, 2006; Rothstein, Jacobsen, & Wilder, 2006). In the best case scenario, it is projected that half of the schools in Michigan will fail to meet AYP requirements by 2014 (Wiley, Mathis, & Garcia, 2005). The ACT will continue to show distinct score gaps for Michigan students as it has in the past (Institute of Educational National Center for Educational

Statistics, 2005). We can continue to castigate educators for their failure to close test score gaps or recognize that the ability to perform on standardized tests is determined early in a child's life and is relatively unresponsive to instruction.

Student demographics have consistently served as primary determinants of standardized test scores (Adams, 1994; Barton, 2005; Maylone, 2002; Institute of Educational Sciences National Center for Educational Statistics, 2005; Thorndike, 1951). Achievement as measured by standardized tests in public, private, parochial, and charter schools is determined primarily by the student population; i.e. low-achieving students populate low-achieving schools. It is unfair to measure groups of students by race/ethnicity or socioeconomic status because there are many exceptions and we run the risk of stereotyping individual students or groups. Children who score poorly on standardized tests are fated to remain at the bottom of an educational caste system without a different measurement of success.

English (2002) stated that the achievement deficit of African-American and Latino students will never be resolved because flawed tests have been used to assess pupil progress. He stated, "In particular, IQ and its derivative achievement test cousins have always shown that socioeconomic status (SES) is a crucial variable in explaining test score variance. SES is part of the concept of cultural capital, and this form of capital is a potent predictor of student success" (p. 298). English concluded that, "As long as the tests in use ignore the impact

of cultural capital and educators support them because they are neutral, the achievement gap will be a permanent feature in American education" (p. 309).

Raising the academic achievement levels of children from low-income families, according to Rothstein (2004), requires ameliorating the social and economic conditions of their lives, not just reforming schools on the basis of the association of social and economic disadvantage with student achievement gaps. However, educators are required to meet the mandates of NCLB and Education YES! in Michigan, despite having little control over the social and economic factors affecting their students.

The data from this study demonstrate the predictability of ACT scores in high school. Little has been done in the past twenty years to effectively close the achievement gaps as measured by standardized tests. The Michigan Merit Core Curriculum that mandates eighteen specific credits for all Michigan high school students in the fall of 2007 (McMillan, 2006) represents more of the same curriculum that has produced achievement gaps. If the gaps are to be effectively closed, the educational system must be reorganized and instructional practice changed.

Nelson (2006A) suggests that more effective educational practice may help low-income and minority students close achievement gaps. For example, an expansion of early childhood education with an emphasis on quality preschools for economically disadvantaged students and extended school days and years, including full-day prekindergarten and kindergarten, may

better meet the social and educational needs of low SES children. Beyond academic remediation, interventions for older children from low-income families, such as school-based health care and access to cultural enrichment activities may help close achievement gaps between these children and children from families with middle class incomes. (Nelson, 2006B). Better recruitment of, and training specific to teachers serving high-poverty schools may also improve achievement of low-income students at all levels (Nelson, 2006B).

Smaller class size in grades K-3 is one intervention that helps close the achievement gap, particularly for African-American and low-income students (Finn & Achilles, 1999; Smith, Molnar, & Zahorik, 2003; Wasley & Lear, 2001). Despite consistent evidence of the positive effect of small class size in early elementary as measured at the high school level, it has not been overwhelmingly embraced by the nation's schools.

Recommendations for Further Research and Study

The findings of this researcher may dismay the critics of public education. Politicians and those members of the public who are demanding accountability of our schools through the use of standardized test scores may be disappointed to be reminded again that this measurement is ineffective and misleading. The review of literature and findings of this study generated topics and recommendations for further investigations. Studies are needed to meet the following goals:

1. To ascertain more definitively when ability on standardized tests is apparent and to establish the predictive value of different standardized tests. Are there consistently reliable measures prior to third grade?
2. To specifically define factors that predict individual performance on standardized tests and develop dependable means of academic intervention based on these factors.
3. To improve instruction for individual students through more effective analysis of standardized test scores. Using scores to compare different groups of students is irrelevant to improving instruction.
4. To continue examination of non-instructional factors inside and outside of the school setting that affect student achievement. If these factors are dependable, address them directly in the earliest educational setting.
5. To improve instructional methods that mitigate the negative effects of a child's cultural and socioeconomic background.
6. To develop more meaningful measurements of academic achievement to meet the demand for accountability.
7. To define objective appraisals of teacher ability. Proponents of merit pay (Holland, 2005) based on the improvement of standardized test scores need to seek alternative assessments of teacher efficacy.

Summary

Our desire to be consistent with what we have already done (Cialdini, 2001) may explain our reluctance to eliminate the use of standardized tests as a measure of instructional effectiveness. The results of this study demonstrate that an individual's level of performance on standardized tests is apparent early in and consistent throughout the student's academic career. Because standardized test scores remain relatively constant, this should serve as an additional admonition that standardized tests are not effective measures of instruction or the quality of schools.

The disparity of disaggregated student achievement as measured by standardized tests continues, and there is no evidence to suggest the possibility that all students will attain academic proficiency. State and federal achievement standards and the goals and mandates of NCLB cannot be fulfilled when using standardized tests as a measurement.

References

ACT. (2006). *ACT Online Prep, The Real ACT Prep Guide*. Retrieved July 26, 2006, from ACT Web Site: http://www.actstudent.org/testprep/index.html

Adams, E. (1994). The effects of cost, income, and socio-economic variables on student scholastic aptitude scores (Doctoral dissertation, Ball State University, 1994). *Dissertation Abstracts International, 55. no. 08A*, 2276.

Allen, J., & Sconing, J. (2005, August). *Using ACT Assessment Scores to Set Benchmarks for College Readiness*. Retrieved October 10, 2006, from ACT, Inc. Web Site: http://www.act.org/research/reports/pdf/ACT_RR2005-3.pdf

Antonak, R. F. (1988). Relationships between group IQ and scholastic achievement at grades two, four, and six. *Educational Research Quarterly, 12*(2), 23-29.

Association for Supervision and Curriculum Development (ASCD). (2005). *A Lexicon of Learning*. Retrieved June 5, 2005, from ASCD Web Site: http://www.ascd.org/portal/site/ascd/menuitem.4247f922ca8c9ecc8c2a9410d3108a0c/

Austin, L. (2005, August 4). *Spellings tells lawmakers No Child Left Behind is good politics*. Retrieved August 5, 2005, from Dallas-Fort Worth Star-Telegram Web Site: http://dfw.com/startelegram/news/state/12304747.htm

Barton, P. E. (2004). Why Does the Gap Persist? *Educational Leadership, 62*(7), 8-13.

Barton, P. E. (2005). Achievement Gaps: Past and Present. *Principal, 84*(4), 12-16.

Berliner, D. C., & Biddle, B. I. (1998). *The Lamentable Alliance Between the Media and School Critics*. Retrieved April 28, 2004, from Arizona State University Web Site: http://courses.ed.asu.edu/berliner/readings/alliancew.htm

Biddle, B. J., & Berliner, D. C. (2002). *What Research Says About Unequal Funding for Schools in America*. Retrieved December 4, 2004, from Arizona State University Web Site: http://www.asu.edu/educ/epsl/EPRP/EPRP-0206-102-EPRP.htm

Bracey, G. (2000, December 5). *High Stakes Testing*. Retrieved March 18, 2004, from Arizona State University Web Site: http://www.asu.edu/educ/epsl/EPRU/documents/cerai-00-32.htm

Carlson, K. (2004). Test Scores by Race and Ethnicity. *Phi Delta Kappan, 85*(5), 379-380.

Chambers, K. (2005, August 20). *Most schools make the grade*. Retrieved August 20, 2005, from Holland Sentinel Web Site: http://www.hollandsentinel.com/stories/082005/local_20050820017.shtml

Chen, C., Lee, S., & Stevenson, H. W. (1996). Long-term prediction of academic achievement of American, Chinese, and Japanese adolescents. *Journal of Educational Psychology, 88,* 750-759.

Cialdini, R. B. (2001). *Influence: Science and Practice* (4th ed.). Boston: Allyn and Bacon.

Dively, J. (2004). Introduction to the Special Issue on Funding and Equity. *National Association of Secondary School Principals Bulletin, 88*(640), 1.

English, F. W. (2002). On the Intractability of the Achievement Gap in Urban Schools and the Discursive Practice of Continuing Racial Discrimination. *Education and Urban Society, 34*(3), 298-311.

Fair Test. (2005, April 21). *The ACT: Biased, Inaccurate, Coachable, and Misused*. Retrieved August 20, 2005, from http://fairtest.org/facts/act.html

Fair Test: The National Center for Fair & Open Testing. (2006). *Norm-Referenced Achievement Tests*. Retrieved April 9, 2006, from http://www.fairtest.org/facts/nratests.html

Finn, J. D., & Achilles, C. M. (1999). Tennessee's Class Size Study: Findings, Implications, Misconceptions. *Educational Evaluation and Policy Analysis, 21*(2), 97-109.

Fraenkel, J. R., & Wallen, N. E. (1996). *How to design and evaluate research in education* (3rd Ed.). San Francisco: McGraw-Hill, Inc.

Frey, A. (2002). Predictors of Placement Recommendations for Children with Behavioral or Emotional Disorders. *Behavioral Disorders, 27*(2), 126-136.

Fuchs, M., & Migdail, S. R. (1977). Correlation of McCarthy Scale Scores with ERB Achievement Tests Over a Three and Four Year Interval. , Retrieved July 1, 2000 from ERIC (ED162747).

Funk, S. G., Sturner, R. A., & Green, J. A. (1986). Preschool prediction of early school performance: relationship of McCarthy scales of children's abilities prior to school entry to achievement in kindergarten, first, and second grades. *Journal of School Psychology, 24,* 181-194.

Georgiou, S. N. (1999). Parental attributions as predictors of involvement and influences on child achievement. *The British Journal of Educational Psychology, 69,* 409-429.

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (1994). Role of parental motivational practices in children's academic intrinsic motivation and achievement. *Journal of Educational Psychology, 86,* 104-113.

Grossman, F. M., & Johnson, K. M. (1983). Validity of the Slosson and Otis-Lennon in Predicting Achievement of Gifted Students. *Educational and Psychological Measurement, 43*(2), 617-622.

Harcourt Assessment, Inc. (2006). *Harcourt Assessment*. Retrieved February 19, 2006, from http://harcourtassessment.com/HAIWEB/Cultures/en-us/default

Illlinois State Board of Education. (2006, August 16). *Illinois students' ACT test scores continue upward trend*. Retrieved September 14, 2006, from http://www.isbe.state.il.us/news/2006/aug16.htm

Institute of Educational Sciences National Center for Educational Statistics. (2005). *ACT score averages and standard deviations, by selected characteristics, sex, and race/ethnicity: Selected years, 1995 through 2004*. Retrieved December 29, 2006, from U.S. Department of Education Web Site: http://nces.ed.gov/programs/digest/d05/tables/dt05_130.asp

Jencks, C., & Phillips, M. (1998). America's Next Achievement Test: Closing the Black-White Test Score Gap. *The American Prospect, 40,* 44.

Kober, N. (2001). *It takes more than testing: Closing the Achievement Gap*. Retrieved August 13, 2006, from Center on Educational Policy Web Site: http://www.ctredpol.org/improvingpublicschools/closingachievementgap.pdf

Lee, J. (2006). *Tracking achievement gaps and assessing the impact of NCLB on the*. Retrieved October 8, 2006, from Harvard University Web Site: http://www.civilrightsproject.harvard.edu/research/esea/nclb_naep_lee.pdf

Lee, K., & Reimann, C. B. (2003, July). *Who's Attending Michigan's Priority Schools?* Retrieved August 14, 2005, from The Educational Policy Center at Michigan State University Web Site: http://www.epc.msu.edu/

Lee, V. E., & Burkam, D. T. (2002, September). *Inequality at the Starting Gate: Social Background Differences in Achievement as Children Begin School*. Retrieved January 20, 2007, from Economic Policy Institute Web Site: http://www.epi.org/content.cfm/books_starting_gate

Mathis, W. J. (2006, September). *The Accuracy and Effectiveness of Adequate Yearly Progress, NCLB's School Evaluation System*. Retrieved December 28, 2006, from Arizona State University Web Site: http://epsl.asu.edu/epru/documents/EPSL-0609-212-EPRU.pdf

Maylone, N. (2002). The Relationship of Socioeconomic Factors and District Scores on the Michigan Educational Assessment Program Tests: An Analysis (Doctoral dissertation, Eastern Michigan University, 2002). *Dissertation Abstracts International*.

McGraw-Hill Education. (2006). *Our Products*. Retrieved February 19, 2006, from McGraw-Hill, Inc. Web Site: http://www.ctb.com/products/product_summary.jsp?FOLDER%3C%3Efolder_id=1408474395222551

McLeod, S., D'Amico, J. J., & Protheroe, N. (2003). *K-12 Principals Guide to No Child Left Behind.* Arlington, VA: Educational Research Service.

McMillan, D. (2006). Michigan Merit Core Curriculum Guides Students Toward the 21st Century. *Leading Change, Spring*(2006), 3.

Michigan Department of Education (MDE). (2006A). *AYP Glossary of Terms*. Retrieved February 19, 2006, from MDE Web Site:

http://www.michigan.gov/mde/0,1607,7-140-22709_22875-85950--,00.html

Michigan Department of Education. (2006B). *Preparing for the ACT*. Retrieved July 26, 2006, from Michigan Department of Education Web Site: http://www.michigan.gov/mde/0,1607,7-140-22709_35150---,00.html

National Center for Education Statistics (U.S. Department of Education). (2004). *The Nation's Report Card: Reading Highlights 2003* (NCES 2004-452 ed.) [Brochure]. Jessup, MD: Author.

National Center for Educational Statistics (U.S. Department of Education). (2004). *The Nation's Report Card: Mathematics Highlights 2003* (NCES 2004-451 ed.) [Brochure]. Jessup, MD: Author.

Navarette, R. (2003, August 17). Lack of Teacher Preparation Hurts Poor Kids. *Detroit News,* p. A15.

Nelson, A. (2006A, April). *Closing the Gap: Early Childhood Education*. Retrieved February 2, 2007, from Association for Supervision and Curriculum Development (ASCD) Web Site: http://www.ascd.org/portal/site/ascd/template.MAXIMIZE/menuitem.c30040c1b9063eeeb85516f762108a0c/?javax.portlet.tpst=d5b9c0fa1a493266805516f762108a0c_ws_MX&javax.portlet.prp_d5b9c0fa1a493266805516f762108a0c_viewID=issue_view&javax.portlet.prp_d5b9c0fa1a493266805516f762108a0c_journalmoid=7902e5a329cca010VgnVCM1000003d01a8c0RCRD&javax.portlet.begCacheTok=token&javax.portlet.endCacheTok=token

Nelson, A. (2006B, Fall). *Overcoming the Income Gap*. Retrieved February 2, 2007, from Association for Supervision and Curriculum Development (ASCD) Web Site: http://www.ascd.org/portal/site/ascd/template.MAXIMIZE/menuitem.c30040c1b9063eeeb85516f762108a0c/?javax.portlet.tpst=d5b9c0fa1a493266805516f762108a0c_ws_MX&javax.portlet.prp_d5b9c0fa1a493266805516f762108a0c_viewID=issue_view&javax.portlet.prp_d5b9c0fa1a493266805516f762108a0c_journalmoid=8d9666745d99e010VgnVCM1000003d01a8c0RCRD&javax.portlet.begCacheTok=token&javax.portlet.endCacheTok=token

Nichols, S. L., Glass, G. V., & Berliner, D. C. (2005, September). *High-Stakes Testing and Student Achievement: Problems for the No Child Left Behind Act*. Retrieved November 13, 2006, from Arizona State University Web Site: http://epsl.asu.edu/epru/documents/EPSL-0509-105-EPRU.pdf

Noble, J. P., Roberts, W. L., & Sawyer, R. L. (2006, October). *Student Achievement, Behavior, Perceptions, and Other Factors Affecting ACT Scores*. Retrieved January 20, 2007, from ACT, Inc. Web Site: http://www.act.org/research/reports/pdf/ACT_RR2006-1.pdf

Noble, J., Davenport, M., Schiel, J., & Pommerich, M. (1999, October). *High School Academic and Noncognitive Variables Related to the ACT Scores of Racial/Ethnic and Gender Groups*. Retrieved November 21, 2004, from ACT Web Site Web Site: http://www.act.org/research/reports/pdf/ACT_RR99-06.pdf

Olszewski-Kubilius, P., Kulieke, M. J., & Shaw, B. (1990). Predictors of achievement in mathematics for gifted males and females. *Gifted Child Quarterly, 34,* 64-71.

Popham, W. J. (2001). *The Truth About Testing: An Educator's Call to Action.* Alexandria, VA: Association for Supervision and Curriculum Development.

Popham, W. J. (2005, October 10-11). *Classroom Assessment: Staying Instructionally Afloat in an Ocean of Accountability*. Paper presented at the meeting of the Educational Testing Service Annual Invitational Conference. New York.

Powers, S., Thompson, D., & Azevedo, B. (1983). The predictive validity of the Stanford mathematics test across race and sex. *Educational and Psychological Measurement, 43,* 645-649.

Prewett, P. N., & Fowler, D. B. (1992). Predictive ability of the Slosson intelligence test with the WISC-R and WRAT-R level 1. *Psychology in the Schools, 29,* 17-21.

Putnam, J. (2006, March 14). Race disparity in MEAP scores has lawmakers seeking answers. *Muskegon Chronicle,* p. 1A.

Ravitch, D. (2002). *A Brief History of Testing and Accountability*. Retrieved September 3, 2006, from Hoover Institution Web Site: http://www.hoover.org/publications/digest/4495866.html

Roberts, W. L., & Noble, J. P. (2004). *Academic and Noncognitive Variables Related to PLAN Scores*. Retrieved October 10, 2006, from ACT, Inc. Web Site: http://www.act.org/research/reports/pdf/ACT_RR2004-1.pdf

Rock, D. A., & Stenner, A. J. (2005). *Assessment Issues in the Testing of Children at School Entry*. Retrieved April 11, 2006, from The Future of Children Web Site: http://www.futureofchildren.org/usr_doc/pg_15_rock-stenner.pdf

Roeber, Ed. (2006). *Planning for the Michigan Merit Examination*. Retrieved February 22, 2006, from Michigan Department of Education Web Site: http://www.michigan.gov/mde/0,1607,7-140-22709_35150---,00.html

Roscigno, V. J. (1998). What Matters When Trying to Understand the Educational Achievement Gap Between Black and White Students? *Social Forces, 76,* 1033-1066.

Rosenbach, J. H., & Rusch, R. R. (1991). IQ and achievement: 1930's to 1980's. *Psychology in the Schools, 28,* 304-309.

Rothstein, R. (2004). The Achievement Gap: A Broader Picture. *Educational Leadership, 62*(3), 40-43.

Rothstein, R., Jacobsen, R., & Wilder, T. (2006, November 14). *'Proficiency for All' - An Oxymoron*. Retrieved January 20, 2007, from Economic Policy Institute Web Site: http://www.epi.org/webfeatures/viewpoints/rothstein_20061114.pdf

Smith, P., Molnar, A., & Zahorik, J. (2003). Class Size Reduction: A Fresh Look at the Data. *Educational Leadership, 61*(1), 72-74.

SRA/McGraw-Hill (2003). *Scoring High TerraNova CTBS A Test Prep Program.* Columbus, OH: SRA/McGraw-Hill.

Standard & Poor's. (2006). *SchoolMatters*. Retrieved July 26, 2006, from http://www.schoolmatters.com/

Stone, B. J. (1992). Prediction of achievement by Asian-American and white children. *Journal of School Psychology, 30,* 91-99.

Thorndike, R. (1951). Community Variables as Predictors of Intelligence and Academic Achievement. *Journal of Educational Psychology, 42,* 321-338.

U.S. Department of Education. (2005). *No Child Left Behind*. Retrieved August 13, 2005, from ED.gov Web Site: http://www.ed.gov/nclb/landing.jhtml?src=fb#

Wasley, P. A., & Lear, R. J. (2001). Small Schools, Real Gains. *Educational Leadership, 58*(6), 22-27.

Weller, L. D., Schnittjer, C. J., & Tuten, B. A. (1992). Predicting achievement in grades three through ten using the Metropolitan readiness test. *Journal of Research in Childhood Education, 6,* 121-130.

West, J., Denton, K., & Germino-Hausken, E. (2000). *America's Kindergartners, NCES 2000-070*. Retrieved March 22, 2004, from U.S. Department of Education. National Center for Educational Statistics. Web Site: http://nces.ed.gov/

Wiley, E. W., Mathis, W. J., & Garcia, D. R. (2005, September). *The Impact of the Adequate Yearly Progress Requirement of the Federal "No Child Left Behind" Act on Schools in the Great Lakes Region*. Retrieved November 13, 2006, from Arizona State University Web Site: http://epsl.asu.edu/epru/documents/EPSL-0509-109-EPRU.pdf

Witt, E. A., Dunbar, S. B., & Hoover, H. D. (1994). A multivariate perspective on sex differences in achievement and later performance among adolescents. *Applied Measurement in Education,* 7(3), 241-254.

Woodruff, D. J. (2003). *Relationships Between EPAS Scores and College Preparatory Course Work in High School*. Retrieved October 10, 2006, from ACT, Inc. Web Site: http://www.act.org/research/reports/pdf/ACT_RR2003-5.pdf

APPENDICES

Appendix A School A Data: 3rd Grade SAT/9 and ACT Composite

Read3	Math3	Lang3	UseInfo3	Think3	ComBat3	SAI3	ACTComp
65.6	61.0	64.2	51.6	61.0	60.2	70.1	23
81.1	72.8	47.9	75.8	74.7	70.7	77.0	27
57.0	63.5	62.3	53.7	62.3	56.2	62.9	22
54.8	68.5	60.4	79.6	64.9	65.6	77.0	23
47.4	52.6	57.0	53.7	58.7	56.6	48.9	24
65.6	57.5	46.3	63.5	67.7	64.1	64.2	25
61.0	57.5	47.9	61.0	57.5	57.8	53.7	23
59.8	64.9	55.3	84.6	64.9	61.0	75.8	26
41.9	37.7	32.3	32.3	41.3	41.4	23.0	17
68.5	50.0	57.0	55.9	67.0	62.4	53.7	22
62.3	48.9	53.2	47.4	59.8	48.8	53.7	18
78.2	67.0	84.6	84.6	89.6	75.1	86.9	24
61.0	42.5	53.2	48.9	62.9	54.9	31.5	18
38.3	61.0	51.6	47.4	45.2	45.3	46.3	18
54.8	64.9	47.9	61.0	58.1	65.3	38.3	19
47.4	68.5	46.3	63.5	58.1	58.0	60.4	21
72.8	86.9	93.3	84.6	93.3	79.1	79.6	26
78.2	93.3	78.2	84.6	86.9	83.0	81.1	29
53.7	78.2	62.3	68.5	69.3	68.6	68.5	25
75.8	86.9	69.3	75.8	79.6	74.2	74.7	27
59.8	62.3	41.9	66.3	61.0	60.7	53.7	25
68.5	78.2	74.7	75.8	78.2	76.5	51.1	24
70.1	89.6	78.2	75.8	78.2	74.7	84.6	30
61.0	79.6	64.2	66.3	73.7	71.5	93.3	24
75.8	86.9	89.6	84.6	86.9	82.6	79.6	26
26.3	45.2	32.3	45.2	34.4	37.8	31.5	18
61.0	68.5	71.8	71.8	71.8	68.5	64.2	22
57.0	54.3	47.9	63.5	54.8	54.9	62.9	20
55.9	86.9	60.4	71.8	58.7	62.8	60.4	19
41.3	50.0	50.0	48.9	47.9	45.1	55.3	19
72.8	74.7	99.0	66.3	86.9	74.4	74.7	25
46.8	52.6	40.2	39.0	44.1	44.0	50.0	19
75.8	78.2	62.3	75.8	84.6	70.5	77.0	27
53.7	54.3	37.1	51.6	52.1	52.1	44.7	18

75.8	59.3	58.7	58.1	67.7	65.7	53.7	26
54.8	37.7	60.4	48.9	54.8	51.3	48.9	15
93.3	67.0	78.2	79.6	89.6	83.5	59.3	29
32.3	31.5	35.8	18.9	25.3	33.3	18.9	15
93.3	82.7	78.2	84.6	86.9	81.6	60.4	25
54.8	63.5	55.3	58.1	47.9	53.3	53.7	21
86.9	89.6	81.1	89.6	89.6	84.7	67.0	26
52.6	48.9	47.9	39.0	44.1	47.7	44.7	21
65.6	45.2	51.6	61.0	59.8	59.9	46.3	19
63.5	54.3	47.9	66.3	59.8	60.3	58.1	20
36.5	52.6	15.4	47.4	39.0	41.8	52.6	17

Appendix B School B Data: 3rd Grade CTBS/4 and ACT Composite

Read3	Lang3	Math3	TotBatt3	WordAna3	Spelling3	ACTComp
57	55	61	58	71	59	18
89	69	77	81	62	74	30
53	60	53	55	38	28	21
66	74	75	73	99	67	25
77	91	79	85	99	95	28
50	45	45	47	62	29	19
48	57	53	52	36	27	21
71	63	72	70	84	63	26
66	74	75	73	99	67	25
63	52	53	56	50	35	22
46	46	56	49	33	18	19
55	61	52	55	73	55	21
70	77	72	75	57	61	27
78	68	54	67	39	31	27
65	72	73	71	54	23	25
59	69	52	60	99	63	24
48	65	52	54	46	42	17
66	64	70	68	54	72	23
68	47	37	50	59	37	25
76	77	99	91	84	88	33
54	64	53	57	67	47	23
54	64	53	57	67	47	20
66	72	71	71	45	48	24

Appendix C School C Data: 3rd Grade TN-CTBS and ACT Composite

Read3	Lang3	Math3	ACTComp
22	1	13	14
50	45	68	21
58	58	65	21
48	54	55	19
56	58	60	23
56	54	60	22
48	34	44	16
65	54	94	17
53	71	65	20
53	86	55	16
58	66	44	17
65	58	60	20
81	71	65	24
41	43	39	16
93	71	99	25
75	66	68	28
81	48	73	28
70	51	47	19
53	45	60	18
53	54	48	21
48	45	40	16
70	66	62	25
62	54	55	20
81	77	94	24
70	71	73	30
93	66	73	27
58	43	53	20
99	66	68	23
22	40	47	15

32	29	39	20
99	71	78	23
58	71	55	23
93	77	94	28
93	77	73	21
62	66	85	20
93	61	85	27
62	61	73	28
32	48	40	18
81	71	85	26
65	54	50	21
93	86	73	20
81	77	94	25
46	51	62	19
75	58	57	21
75	86	73	31
70	54	78	20
46	54	60	22
65	61	78	22
81	51	60	22
53	61	62	24
93	54	62	26
93	61	65	22
65	77	68	26
99	86	94	26
65	71	53	24
81	99	94	28
36	31	39	13
62	58	57	21
42	45	39	19
62	51	62	21
81	58	57	21
46	58	50	19
93	77	78	28

81	86	65	31
70	54	73	24
62	48	50	23
58	66	57	20
81	99	94	26
62	66	73	22
81	66	68	24
65	77	68	25
48	48	39	17
75	66	65	26
58	61	65	25
99	71	94	31
65	61	85	29
65	66	78	19
53	54	55	22
65	71	62	21
58	58	53	17
58	54	48	18
81	86	78	26
65	51	65	18
70	48	47	23
65	61	62	23
53	58	48	18
99	86	73	21
81	66	94	27
93	71	94	25
53	61	57	25
50	58	57	24
48	37	50	16
93	77	65	24
48	77	57	19
48	58	62	18
62	54	50	19
37	48	50	16

75	77	65	27
99	86	85	30
99	86	78	26
93	86	99	32
62	54	65	18
36	45	47	19
70	99	73	26
37	31	47	18
70	71	73	29
75	66	73	22
50	54	68	19
62	61	68	20
46	45	55	20
58	48	44	22
41	29	37	22
41	48	42	19
29	24	40	13
93	86	73	26
99	58	94	23
50	51	55	18
99	71	94	29
48	61	65	15
93	66	68	25
65	99	85	28
81	86	73	21
36	31	47	18
53	66	50	22
70	71	85	25
53	61	73	21
29	45	55	16
58	86	60	23
70	77	78	24
41	43	47	16
62	66	65	22

58	61	57	22
75	66	60	23
70	66	85	23
75	99	94	26
53	61	57	17
75	77	99	25
48	58	50	18
99	86	94	24
50	58	78	22
58	61	65	23
65	54	65	21
81	77	94	19
39	61	50	18
65	61	57	28
70	77	99	25
81	66	53	28
65	66	53	20
58	66	50	18
58	66	78	25
99	86	85	32
48	51	47	15
53	48	57	18
58	58	53	19
50	37	32	19
70	66	65	21
70	71	60	29
81	77	65	25
70	86	94	28
65	77	55	26
39	54	53	24
46	54	39	18
29	43	32	15
99	99	94	30
70	71	50	26

32	31	39	14
75	51	78	24
81	86	94	26
56	51	85	19
81	61	78	29
56	54	57	21
62	45	55	18
93	58	68	22
93	99	62	23
93	77	94	33
65	58	62	26
58	51	73	22
65	66	68	22
70	77	57	22
36	45	30	19
32	34	57	18
81	77	99	35

Appendix D School A Ranked Data: 3rd Grade SAT/9 and ACT Composite

RRead3	RMath3	RLang3	RUseInfo3	RThink3	RComBat3	RSAI3	RACT
29	20.5	32.5	12.5	21.5	20	33	24
42	32	12.5	34	33	33	38	41
18.5	23.5	30	14.5	23	15	27.5	21
14.5	30	27	37.5	25.5	28	38	24
8.5	12	23.5	14.5	16.5	16	10.5	27.5
29	17.5	8.5	23	28.5	26	29.5	32
23.5	17.5	12.5	20	13	17	17.5	24
20.5	25.5	21.5	41.5	25.5	23	36	37
6	2.5	2.5	2	4	3	2	3.5
31.5	9.5	23.5	16	27	24	17.5	21
26	7.5	19.5	7	19	9	17.5	7
40.5	27.5	42	41.5	43	38	44	27.5
23.5	4	19.5	10	24	13.5	3.5	7
4	20.5	17.5	7	7	7	8.5	7
14.5	25.5	12.5	20	14.5	27	5	12
8.5	30	8.5	23	14.5	18	25	18
34.5	40.5	44	41.5	45	40	40.5	37
40.5	45	38.5	41.5	39.5	43	42	43.5
11.5	35	30	29	30	31	32	32
37.5	40.5	34	34	36	35	34.5	41
20.5	22	7	26.5	21.5	22	17.5	32
31.5	35	36	34	34.5	39	13	27.5
33	43.5	38.5	34	34.5	37	43	45
23.5	37	32.5	26.5	32	34	45	27.5
37.5	40.5	43	41.5	39.5	42	40.5	37
1	5.5	2.5	5	2	2	3.5	7
23.5	30	35	30.5	31	30	29.5	21
18.5	15	12.5	23	11.5	13.5	27.5	15.5
17	40.5	27	30.5	16.5	25	25	12

5	9.5	16	10	8.5	6	21	12
34.5	33	45	26.5	39.5	36	34.5	32
7	12	6	3.5	5.5	5	12	12
37.5	35	30	34	37	32	38	41
11.5	15	5	12.5	10	11	6.5	7
37.5	19	25	17.5	28.5	29	17.5	37
14.5	2	27	10	11.5	10	10.5	1.5
44.5	27.5	38.5	37.5	43	44	23	43.5
2	1	4	1	1	1	1	1.5
44.5	38	38.5	41.5	39.5	41	25	32
14.5	23.5	21.5	17.5	8.5	12	17.5	18
43	43.5	41	45	43	45	31	37
10	7.5	12.5	3.5	5.5	8	6.5	18
29	5.5	17.5	20	19	19	8.5	12
27	15	12.5	26.5	19	21	22	15.5
3	12	1	7	3	4	14	3.5

Appendix E School B Rank Data: 3rd Grade CTBS/4

and ACT Composite

RRead3	RLang3	RMath3	RTotBatt3	RWorAn3	RSpell3	RACT
9	5	13	11	16	14	2
23	15.5	21	21	12.5	21	22
5	7	8	6.5	3	4	7
14.5	19.5	19.5	18.5	21.5	18.5	15.5
21	23	22	22	21.5	23	21
4	1	2	1	12.5	5	3.5
2.5	6	8	4	2	3	7
19	9	16.5	15	18.5	16.5	18
14.5	19.5	19.5	18.5	21.5	18.5	15.5
11	4	8	8	7	7	9
1	2	12	2	1	1	3.5
8	8	4	6.5	17	13	7
18	21.5	16.5	20	10	15	19.5
22	14	11	13	4	6	19.5
12	17.5	18	16.5	8.5	2	15.5
10	15.5	4	12	21.5	16.5	12.5
2.5	13	4	5	6	9	1
14.5	11	14	14	8.5	20	10.5
17	3	1	3	11	8	15.5
20	21.5	23	23	18.5	22	23
6.5	11	8	9.5	14.5	10.5	10.5
6.5	11	8	9.5	14.5	10.5	5
14.5	17.5	15	16.5	5	12	12.5

Appendix F School C Rank Data: 3rd Grade TN-CTBS and ACT Composite

RRead3	RLang3	RMath3	RACT
1.5	1	1	3.5
40.5	22	114	74.5
67.5	73	100	74.5
32.5	55	53.5	46.5
57.5	73	77	107
57.5	55	77	91.5
32.5	9.5	18	12.5
97	55	168	19
49.5	131.5	100	60.5
49.5	167	53.5	12.5
67.5	111.5	18	19
97	73	77	60.5
143.5	131.5	100	119.5
19.5	15.5	9	12.5
162	131.5	180	132
128.5	111.5	114	162.5
143.5	31	128	162.5
115	40.5	23.5	46.5
49.5	22	77	30
49.5	55	29	74.5
32.5	22	14	12.5
115	111.5	86.5	132
81.5	55	53.5	60.5
143.5	149	168	119.5
115	131.5	128	174
162	111.5	128	155.5
67.5	15.5	45	60.5
176.5	111.5	114	107
1.5	13	23.5	6.5

7.5	3.5	9	60.5
176.5	131.5	142.5	107
67.5	131.5	53.5	107
162	149	168	162.5
162	149	128	74.5
81.5	111.5	153.5	60.5
162	90.5	153.5	155.5
81.5	90.5	128	162.5
7.5	31	14	30
143.5	131.5	153.5	146
97	55	36	74.5
162	167	128	60.5
143.5	149	168	132
25	40.5	86.5	46.5
128.5	73	65.5	74.5
128.5	167	128	177
115	55	142.5	60.5
25	55	77	91.5
97	90.5	142.5	91.5
143.5	40.5	77	91.5
49.5	90.5	86.5	119.5
162	55	86.5	146
162	90.5	100	91.5
97	149	114	146
176.5	167	168	146
97	131.5	45	119.5
143.5	179	168	162.5
11.5	6.5	9	1.5
81.5	73	65.5	74.5
22	22	9	46.5
81.5	40.5	86.5	74.5
143.5	73	65.5	74.5
25	73	36	46.5
162	149	142.5	162.5

143.5	167	100	177
115	55	128	119.5
81.5	31	36	107
67.5	111.5	65.5	60.5
143.5	179	168	146
81.5	111.5	128	91.5
143.5	111.5	114	119.5
97	149	114	132
32.5	31	9	19
128.5	111.5	100	146
67.5	90.5	100	132
176.5	131.5	168	177
97	90.5	153.5	170
97	111.5	142.5	46.5
49.5	55	53.5	91.5
97	131.5	86.5	74.5
67.5	73	45	19
67.5	55	29	30
143.5	167	142.5	146
97	40.5	100	30
115	31	23.5	107
97	90.5	86.5	107
49.5	73	29	30
176.5	167	128	74.5
143.5	111.5	168	155.5
162	131.5	168	132
49.5	90.5	65.5	132
40.5	73	65.5	119.5
32.5	11.5	36	12.5
162	149	100	119.5
32.5	149	65.5	46.5
32.5	73	86.5	30
81.5	55	36	46.5
14.5	31	36	12.5

128.5	149	100	155.5
176.5	167	153.5	174
176.5	167	142.5	146
162	167	180	179.5
81.5	55	100	30
11.5	22	23.5	46.5
115	179	128	146
14.5	6.5	23.5	30
115	131.5	128	170
128.5	111.5	128	91.5
40.5	55	114	46.5
81.5	90.5	114	60.5
25	22	53.5	60.5
67.5	31	18	91.5
19.5	3.5	5	91.5
19.5	31	16	46.5
4	2	14	1.5
162	167	128	146
176.5	73	168	107
40.5	40.5	53.5	30
176.5	131.5	168	170
32.5	90.5	100	6.5
162	111.5	114	132
97	179	153.5	162.5
143.5	167	128	74.5
11.5	6.5	23.5	30
49.5	111.5	36	91.5
115	131.5	153.5	132
49.5	90.5	128	74.5
4	22	53.5	12.5
67.5	167	77	107
115	149	142.5	119.5
19.5	15.5	23.5	12.5
81.5	111.5	100	91.5

67.5	90.5	65.5	91.5
128.5	111.5	77	107
115	111.5	153.5	107
128.5	179	168	146
49.5	90.5	65.5	19
128.5	149	180	132
32.5	73	36	30
176.5	167	168	119.5
40.5	73	142.5	91.5
67.5	90.5	100	107
97	55	100	74.5
143.5	149	168	46.5
16.5	90.5	36	30
97	90.5	65.5	162.5
115	149	180	132
143.5	111.5	45	162.5
97	111.5	45	60.5
67.5	111.5	36	30
67.5	111.5	142.5	132
176.5	167	153.5	179.5
32.5	40.5	23.5	6.5
49.5	31	65.5	30
67.5	73	45	46.5
40.5	11.5	3.5	46.5
115	111.5	100	74.5
115	131.5	77	170
143.5	149	100	132
115	167	168	162.5
97	149	53.5	146
16.5	55	45	119.5
25	55	9	30
4	15.5	3.5	6.5
176.5	179	168	174
115	131.5	36	146

7.5	6.5	9	3.5
128.5	40.5	142.5	119.5
143.5	167	168	146
57.5	40.5	153.5	46.5
143.5	90.5	142.5	170
57.5	55	65.5	74.5
81.5	22	53.5	30
162	73	114	91.5
162	179	86.5	107
162	149	168	181
97	73	86.5	146
67.5	40.5	128	91.5
97	111.5	114	91.5
115	149	65.5	91.5
11.5	22	2	46.5
7.5	9.5	65.5	30
143.5	149	180	182