

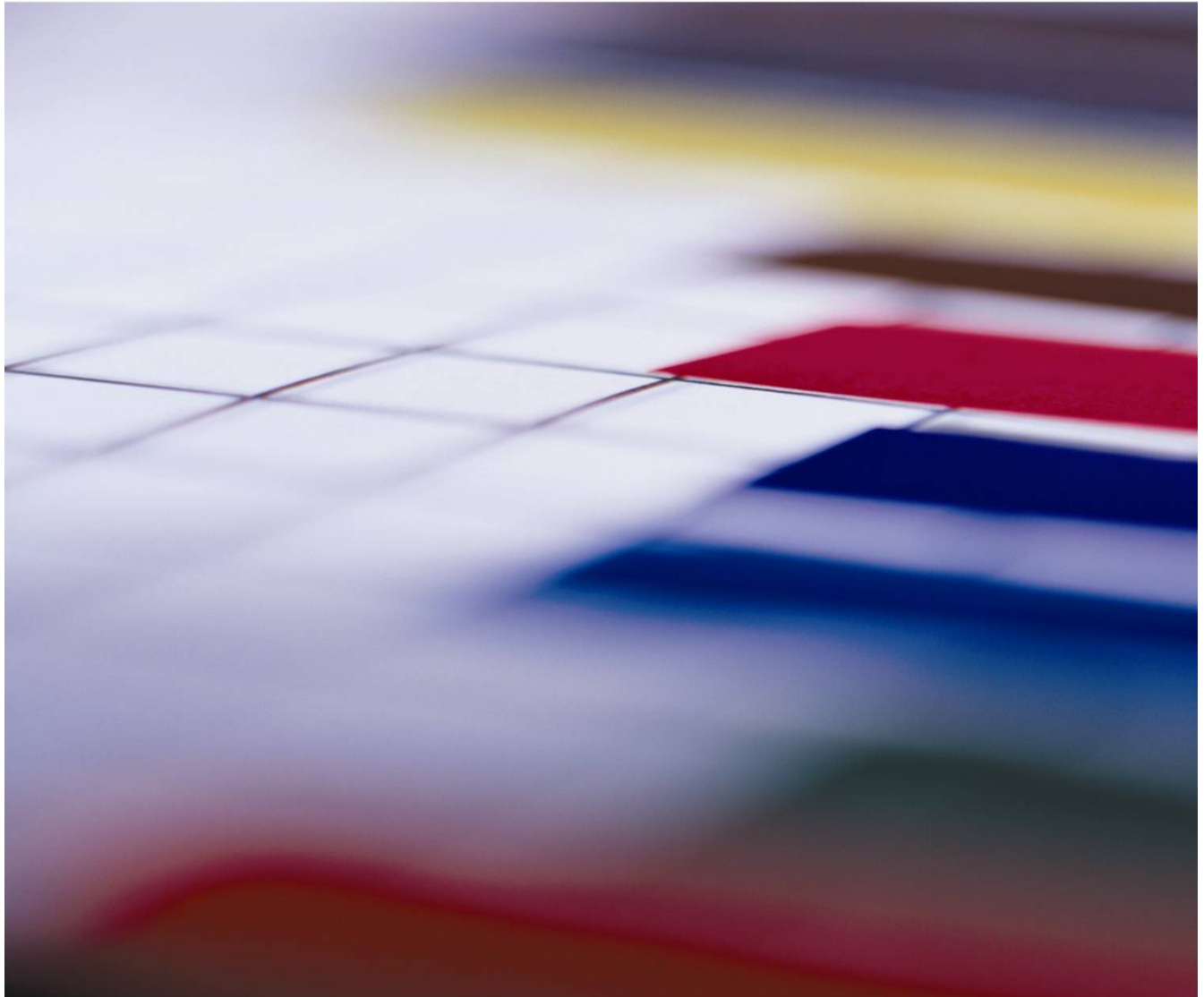


Center for Research in Educational Policy

The University of Memphis
325 Browning Hall
Memphis, Tennessee 38152
Toll Free: 1-866-670-6147

Analysis of MOST Student Achievement

2006-2007





Center for Research in Educational Policy

The University of Memphis
325 Browning Hall
Memphis, Tennessee 38152
Toll Free: 1-866-670-6147

Analysis of MOST Student Achievement

2006-2007

June 2008

Todd Zoblotsky
Aaron McDonald
Erika Strahl Layton
Center for Research in Educational Policy

2006-2007 MOST STUDENT ACHIEVEMENT ANALYSIS

Introduction

The Memphis Opportunity Scholarship Trust (MOST) was established to provide educational choices to low-income families in the Memphis area. This is primarily accomplished by providing economically disadvantaged students with tuition assistance scholarships. Students receiving scholarships have elected to attend a variety of private schools.

The overarching purpose of this study was to examine the academic achievement of students receiving MOST scholarships. A challenge to this examination was that the different schools attended by MOST students did not all use a “common” assessment of student achievement. As a result of the variety of tests and different types of available data, two broad approaches were used to examine achievement. The first consisted of a descriptive study of kindergarten and first grade students not attending Catholic Diocese Jubilee Schools. The second broad approach involved the Jubilee Schools. A more rigorous analytical examination was used with this group of MOST students because of access to individual student level data for both MOST and non-MOST students attending the Jubilee Schools. Both approaches are discussed in the following sections.

Achievement Analysis for MOST Students attending Non-Jubilee Schools

This analysis examined the academic achievement of kindergarten and first grade students attending private schools other than Jubilee Schools via tuition assistance provided through MOST. The purposes of this study were to (a) examine the progress of this cohort of students in relation to grade level peers and national or independent school test norms and (b) set a baseline for a longitudinal study which could potentially follow this cohort of students over time. To accomplish these purposes, student achievement was analyzed from two perspectives. First, the students receiving MOST scholarships were compared to class or grade-level averages for their school. Second, students were compared to test norms. In most cases, national norms were utilized. For the two schools using the Comprehensive Testing Program (CTP-4), norms for “independent schools” were provided in lieu of national norms. The research questions that guided this analysis were as follows:

1. How did MOST students perform on the 2007 standardized test utilized by their school relative to the norms or averages for students in their class or grade?
2. How did MOST students perform on the 2007 standardized test utilized by their school relative to the national or independent school norms (averages) for their grade?

Methodology

MOST student participants

Approximately 192 Kindergarten and first grade MOST students enrolled in 45 non-Jubilee schools in 2006-07 academic year. Of these students, 114 were excluded from the analysis for the following reasons: 68 attended schools where their grade was not tested; 9 had either left the school or discontinued receiving the scholarship; 12 were excluded because their

parents could not be contacted to provide consent to participate in the study; 9 parents did not give permission for their child’s scores to be used in the study; and the test results for 16 students could not be obtained from their schools.

These exclusions left 78 students whose scores were analyzed. In some cases, either class/grade norms or national/system averages could not be obtained, so these students’ scores were analyzed using the data available. Five different standardized tests were used in the participating schools: (1) Stanford Achievement Test (versions 9 and 10), (2) Iowa Tests of Basic Skills (ITBS), (3) Comprehensive Testing Program 4 (CTP 4), (4) TerraNova, and (5) Gates-MacGinitie Reading Test (Gates). Table 1 lists the number of students whose scores were obtained by school, grade and test.

Table 1: Students by school and test

School	Test	Kindergarten	First Grade
Breath of Life Christian Academy	Stanford 10	2	5
Brinkley Heights Urban Academy	Stanford 10	13	9
Central Baptist School	Stanford 10		1
Christ the King Lutheran School	Stanford 9		1
Christ the Rock Christian Academy	Stanford 10	3	3
Evangelical Christian School	CTP 4		3
First Assembly Christian School	Stanford 10		1
Heritage Baptist Academy	Stanford 10	3	
Holy Rosary School	ITBS		3
Immaculate Conception Cathedral School	ITBS		3
The LaGrange School	Stanford 10	1	
Lord’s Tabernacle Christian Academy	TerraNova	3	
Macon Road Baptist School	Stanford 10	3	2
Margolin Hebrew Academy	CTP 4		3
Rossville Christian Academy	Stanford 10	1	
St. Ann-Bartlett School	ITBS		1
St. Anne School	ITBS		1
St. Francis of Assisi School	ITBS		1
St. Mary’s Episcopal School	Gates	2	
St. Michael School	ITBS		5
St Paul School	ITBS		4
World Overcomers Christian Academy	Stanford 10	1	

Procedure

Participating schools were contacted by mail during the fall of 2006 and asked to return a form listing the type of standardized test they administered, the grades to which the test was administered, and the dates of testing. A postage-paid envelope was provided for the schools to return the forms to the Center for Research in Educational Policy (CREP) at The University of Memphis. Follow-up telephone calls were also made to collect this information.

An informed consent form was sent to parents or guardians of applicable MOST scholarship recipients in late fall, 2006. Addresses were provided by MOST administrators. The consent forms were passive, meaning parents were only required to return them if they did *not* want their child's information to be included in the analysis. A postage-paid envelope was provided for parents/guardians to return the forms. Two schools (St. Paul and Central Baptist School) requested that active consent forms (meaning parents had to return the form for their child to be included) be sent out before releasing scores. This request was accommodated and copies of the signed forms were provided to these schools.

The majority of schools administered the standardized tests in spring of 2007, and thus, collection of test scores began in May. Schools were contacted by telephone and arrangements were made on a case-by-case basis to collect student scores. In some cases, the schools mailed or faxed the scores to CREP. In others, CREP personnel visited the schools to obtain the scores. MOST administrators assisted with communication to the schools in order to collect additional scores. Paper-based scores were obtained in all cases.

Measures and Analysis

For most tests that were used in the schools, a National Percentile Rank (NPR) was available for the student's classroom or grade level as well as the individual student. A

percentile rank shows the percentage of students in the normative or comparison group that a particular score surpassed, and is useful to look at relative position within a normative (norm) or peer group. For example, a student scoring at the 75th percentile outscored 75% of the students in the comparison group. Similarly, for the classroom or grade level NPRs, if the first grade students at a particular school had a NPR of 63, then that grade scored higher than 63% of first grade students in the nationally representative norm group. A NPR of 50 is average, with scores below 50 being below average, and scores above 50 being above average.

There were two tests that did not provide NPRs. For the CTP-4, norms were comprised of student scores from “independent schools” taking the test. Thus, these norms are not supposed to be nationally representative (like a NPR), but instead, are a relative or ranked score for independent schools that administered the CTP-4. Another exception was the Gates-McGinitie Test, for which stanine scores were provided. A stanine is a standard score between 1 and 9, with a mean of 5 and a standard deviation of 2. Therefore, a stanine score of 5 is average, with scores below and above 5 considered below and above average, respectively.

Based on the available data, students’ individual NPR was compared to their class or grade level NPR for their school. Any individual student score above the class/grade NPR was considered “Above Average.” Scores below the class/grade NPR were considered “Below Average,” and scores identical to the class/grade NPR were considered “Average.” For example, if a school’s grade level NPR was 75 and an individual student’s score was 74 or lower, he or she was listed as “Below Average.” If his or her score was at the 76th percentile or higher, he or she was listed as “Above Average.” If his or her score was at the 75th percentile, he or she was considered “Average.”

For the second analysis, students were grouped into four categories (quartiles) based on their NPR: 0% to 25%, 26% to 50%, 51% to 75% and 76% to 100%. For the Gates-McGinitie Test, students were grouped into 6th, 7th, 8th or 9th stanine as no students fell into the lower stanines. This shows how MOST students tended to score in comparison with national normative groups. It is important to note that these analyses do not infer or test “statistical significance.” That is, if a student is above or below average, this does not mean that they are statistically significantly above or below average. Instead, these absolute determinations (i.e., above, below) are only descriptive.

Results

Summaries of results are presented by type of test and grade level. The results for class or grade level show the number of MOST scholarship students who fell below, at, or above the average scores of their peers in the same classroom or grade at their school (this information varied by school). The results for National Percentile Rank indicate the number of MOST students who fell into quartiles. Results for stanines indicate the number of MOST students for each stanine from 6 to 9.

Stanford Achievement Test

Kindergarten

Table 2 and Table 3 present results for Kindergarten MOST students in all schools who took the Stanford Achievement Test (versions 9 and 10).

- Compared to their peers, a slight majority (58%) of MOST students scored below average on the Total Reading and Mathematics subtests.
- The majority (65%) of MOST students scored below average when compared to their peers on the Environment and Listening subtests.

- Sixty-two percent of MOST students scored below average when compared to their peers on both the Basic and Complete Batteries of the Stanford test.
- Thirty-seven percent of MOST students' NPR scores fell in the highest quartile for Total Reading, while 30% fell in the lowest quartile.
- Almost half (48%) of MOST students' NPR scores fell in the lowest quartile for Mathematics, while 41% scored above the national average of 50.
- Slightly over half (56%) of MOST students' Complete Battery NPR scores fell below the national average of 50.

Table 2: Stanford Kindergarten Class/Grade Results N = 26

	Below Average		Average		Above Average	
	Number	Percent	Number	Percent	Number	Percent
Total Reading	15	57.7	1	3.8	10	38.5
Sounds & Letters	14	53.8	1	3.8	11	42.3
Word Reading	15	57.7	1	3.8	10	38.5
Sentence Reading	13	50.0	2	7.7	11	42.3
Mathematics	15	57.7	2	7.7	9	34.6
Environment	17	65.4	4	15.4	5	19.2
Listening	17	65.4	1	3.8	8	30.8
Basic Battery	16	61.5	1	3.8	9	34.6
Complete Battery	16	61.5	1	3.8	9	34.6

Table 3: Stanford Kindergarten National Percentile Rank Results N = 27

	0-25%		26-50%		51-75%		76-100%	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total Reading	8	29.6	3	11.1	6	22.2	10	37.0
Sounds & Letters	7	25.9	10	37.0	0	0.0	10	37.0
Word Reading	8	29.6	2	7.4	5	18.5	12	44.4
Sentence Reading	3	11.1	8	29.6	4	14.8	12	44.4
Mathematics	13	48.1	3	11.1	6	22.2	5	18.5
Environment	16	59.3	3	11.1	5	18.5	3	11.1
Listening	9	33.3	7	25.9	8	29.6	3	11.1
Basic Battery	8	29.6	5	18.5	6	22.2	8	29.6
Complete Battery	8	29.6	7	25.9	5	18.5	7	25.9

1st Grade

Table 4 and Table 5 present results for 1st grade MOST students in all schools that utilized the Stanford Achievement Test (9 and 10).

- Compared to their peers, a slight majority (57%) of MOST students scored below average on the Total Reading subtest.
- Half (50%) of MOST students scored above average when compared to their peers on the Total Mathematics subtest.
- Half (50%) of MOST students scored below average when compared to their peers on the Language subtest.
- Slightly over half (52%) of MOST students scored below average on the Complete Battery when compared to their peers.
- The majority (67%) of MOST students' NPR scores fell within the middle two quartiles for Total Reading.
- In the Total Mathematics subtest, a slight majority (52%) of MOST students' NPR scores fell above the national average of 50.
- MOST students' NPR scores for the Language subtest were evenly distributed between the bottom and top halves of national scores (50% for each).
- In the Environment subtest, the majority (70%) of MOST students' NPR scores fell in the lowest quartile.
- Most (76%) of the MOST students' NPR scores for the Complete Battery fell within the middle two quartiles of scores.

Table 4: Stanford 1st Grade Class/Grade Results N = 22

	Below Average		Average		Above Average		Category N
	Number	Percent	Number	Percent	Number	Percent	
Total Reading*	12	57.1	0	0.0	9	42.9	21
Word Study Skills	11	50.0	3	13.6	8	36.4	22
Word Reading	10	45.5	2	9.1	10	45.5	22
Sentence Reading*	12	60.0	2	10.0	6	30.0	20
Reading Comprehension	13	59.1	1	4.5	8	36.4	22
Total Mathematics	10	45.5	1	4.5	11	50.0	22
Mathematics Problem Solving	10	45.5	1	4.5	11	50.0	22
Mathematics Procedures	10	45.5	2	9.1	10	45.5	22
Language	11	50.0	1	4.5	10	45.5	22
Spelling	9	40.9	4	48.2	9	40.9	22
Environment	10	45.5	4	18.2	8	36.4	22
Listening	10	45.5	3	13.6	9	40.9	22
Basic Battery*	10	47.6	1	4.8	10	47.6	21
Complete Battery*	11	52.4	1	4.8	9	42.8	21

*Not all students had a score for every category.

Table 5: Stanford 1st Grade National Percentile Rank Results N = 22

	0-25%		26-50%		51-75%		76-100%		Category N
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	
Total Reading*	3	14.3	8	38.1	6	28.6	4	19.0	21
Word Study Skills	3	13.7	6	27.3	9	40.1	4	18.2	22
Word Reading	4	18.2	8	36.4	2	9.1	8	36.4	22
Sentence Reading*	3	15.0	13	65.0	0	0.0	4	20.0	20
Reading Comprehension	4	18.2	7	31.8	6	27.3	5	22.7	22
Total Mathematics	1	4.5	9	40.1	5	22.7	7	31.8	22
Mathematics Problem Solving	5	22.7	5	22.7	7	31.8	5	22.7	22
Mathematics Procedures	1	4.5	5	22.7	8	36.4	8	36.4	22
Language	3	13.7	8	36.4	3	13.7	8	36.4	22
Spelling	1	4.5	6	27.3	9	40.1	6	27.3	22
Environment	9	70.1	4	18.2	6	27.3	3	13.7	22
Listening	7	31.8	4	18.2	7	31.8	4	18.2	22
Basic Battery*	1	4.8	9	42.9	7	33.3	4	19.0	21
Complete Battery*	1	4.8	9	42.9	7	33.3	4	19.0	21

*Not all students had a score for every category.

Iowa Test of Basic Skills for 1st Grade

Table 6 and Table 7 present results for 1st grade MOST students in all schools who took the Iowa Test of Basic Skills (ITBS).

- The majority of students were Above Average on all subtests except for Total Reading, while students were evenly divided (50% each) between Below and Above Average.

- A slight majority (56%) of MOST students' scores were above average in the Language subtest when compared to their peers.
- In the Mathematics subtest, the majority (67%) of MOST students' scores were above average when compared to their peers.
- Core Total scores for a slight majority (56%) of MOST students were above average when compared to the scores of their peers.
- In Total Reading, the majority (67%) of MOST students' NPR scores fell above the national average of 50.
- The majority (61%) of MOST students' NPR scores for the Language subtest fell in the top quartile and none were in the bottom quartile.
- In the Mathematics subtest, the majority (72%) of MOST students' NPR scores fell within the middle two quartiles of scores.
- The majority (78%) of MOST students' NPR Core Total scores were above the national average of 50.

Table 6: ITBS 1st Grade Class/Grade Results N = 18

	Below Average		Average		Above Average		Category N
	Number	Percent	Number	Percent	Number	Percent	
Vocabulary	8	47.1	0	0.0	9	52.9	17
Reading: Words	5	31.3	0	0.0	11	68.8	16
Reading: Comprehension	8	47.1	0	0.0	9	52.9	17
Reading: Total	9	50.0	0	0.0	9	50.0	18
Word Analysis	7	41.2	0	0.0	10	58.8	17
Listening	7	41.2	1	5.9	9	52.9	17
Language	8	44.4	0	0.0	10	55.6	18
Mathematics	6	33.3	0	0.0	12	66.7	18
Core Total	8	44.4	0	0.0	10	55.6	18

*Not all students had a score for every category.

Table 7: ITBS 1st Grade National Percentile Rank Results N = 18

	0-25%		26-50%		51-75%		76-100%	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Vocabulary	4	22.2	0	0.0	5	27.8	9	50.0
Reading: Words	1	5.6	4	22.2	3	16.7	10	55.6
Reading: Comprehension	2	11.1	4	22.2	4	22.2	8	44.4
Reading: Total	2	11.1	3	16.7	6	33.3	7	38.9
Word Analysis	2	11.1	4	22.2	5	27.8	7	38.9
Listening Language	1	5.6	3	16.7	2	11.1	12	66.7
Mathematics	0	0.0	3	16.7	4	22.2	11	61.1
Core Total	1	5.6	7	38.9	6	33.3	4	22.2
	3	16.7	1	5.6	7	38.9	7	38.9

Comprehensive Testing Program 4 (CTP4) for 1st Grade

Table 8 and Table 9 present results for 1st grade MOST students in all schools who took the CTP4 test (please note that these results are based on a small number (N=3) of MOST students).

- Two-thirds (67%) of MOST students' scores in Auditory Comprehension, Reading Comprehension, Word Analysis, and Mathematics were below average when compared to their peers.
- Half (50%) of MOST students' NPR scores were in the top quartile in Auditory Comprehension.
- In Reading Comprehension, half (50%) of MOST students' NPR scores were in the bottom quartile.
- In Word Analysis and Mathematics, half (50%) of MOST students' NPR scores were in the 26-50% range.

Table 8: CTP4 1st Grade Class/Grade Results N = 3

	Below Average		Average		Above Average	
	Number	Percent	Number	Percent	Number	Percent
Auditory Comprehension	2	66.7	0	0.0	1	33.3
AC: Vocabulary in Context	3	100.0	0	0.0	0	0.0
AC: Explicit Information	1	33.3	0	0.0	2	66.7
AC: Inference	2	66.7	0	0.0	1	33.3
AC: Analysis	1	33.3	0	0.0	2	66.7
Reading Comprehension	2	66.7	0	0.0	1	33.3
RC: Vocabulary in Context	2	66.7	1	33.3	0	0.0
RC: Explicit Information	3	100.0	0	0.0	0	0.0
RC: Inference	1	33.3	0	0.0	1	33.3
RC: Analysis	2	66.7	0	0.0	1	33.3
Word Analysis	2	66.7	0	0.0	1	33.3
WA: Sight Words	2	66.7	0	0.0	1	33.3
WA: Phonic Analysis	2	66.7	0	0.0	1	33.3
WA: Structure Analysis	2	66.7	0	0.0	1	33.3
Mathematics	2	66.7	0	0.0	1	33.3
M: Num. Sense/Oper. W Whole Num.	2	66.7	0	0.0	1	33.3
M: Geometry & Spatial Sense	1	33.3	0	0.0	2	66.7
M: Measurement	3	100.0	0	0.0	0	0.0
M: Conceptual Understanding	2	66.7	0	0.0	1	33.3
M: Procedural Knowledge	2	66.7	0	0.0	1	33.3
M: Problem Solving	2	66.7	0	0.0	1	33.3

Table 9: CTP4 1st Grade National Percentile Rank Results N = 6

	0-25%		26-50%		51-75%		76-100%	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Auditory Comprehension	2	33.3	0	0.0	1	16.7	3	50.0
AC: Vocabulary in Context*	0	0.0	0	0.0	3	100.0	0	0.0
AC: Explicit Information*	0	0.0	0	0.0	1	33.3	2	66.7
AC: Inference*	0	0.0	0	0.0	2	66.7	1	33.3
AC: Analysis*	0	0.0	0	0.0	0	0.0	3	100.0
Reading Comprehension	3	50.0	1	16.7	1	16.7	1	16.7
RC: Vocabulary in Context*	0	0.0	1	33.3	2	66.7	0	0.0
RC: Explicit Information*	0	0.0	1	33.3	2	66.7	0	0.0
RC: Inference*	0	0.0	1	33.3	1	33.3	1	33.3
RC: Analysis*	0	0.0	2	66.7	1	33.3	0	0.0
Word Analysis	0	0.0	3	50.0	1	16.7	2	33.3
WA: Sight Words*	0	0.0	0	0.0	0	0.0	3	100.0
WA: Phonic Analysis*	0	0.0	0	0.0	0	0.0	3	100.0
WA: Structure Analysis*	0	0.0	1	33.3	1	33.3	1	33.3
Mathematics	0	0.0	3	50.0	2	33.3	1	16.7
M: Num. Sense/Oper. W Whole Num.*	0	0.0	1	33.3	1	33.3	1	33.3
M: Geometry & Spatial Sense*	0	0.0	0	0.0	1	33.3	2	66.7
M: Measurement*	0	0.0	2	66.7	1	33.3	0	0.0
M: Conceptual Understanding*	0	0.0	1	33.3	1	33.3	1	33.3
M: Procedural Knowledge*	0	0.0	0	0.0	2	66.7	1	33.3
M: Problem Solving*	0	0.0	1	33.3	1	33.3	1	33.3

*Not all students had a score for every category.

Gates-McGinitie Reading Test for Kindergarten

Table 10 and Table 11 present results for 1st grade MOST students in all schools who took the Gates-McGinitie Reading test (please note that these results are based on a small number (N=2) of MOST students).

- MOST students' Total scores were evenly distributed between the below average and above average categories when compared to their peers (50% in each).
- MOST students Total scores were evenly distributed between the seventh and ninth stanines (50% in each).

Table 10: Gates Kindergarten Class/Grade Results N = 2

	Below Average		Average		Above Average	
	Number	Percent	Number	Percent	Number	Percent
Total	1	50.0	0	0.0	1	50.0
Initial Consonants	0	0.0	2	100.0	0	0.0
Final Consonants	1	50.0	1	50.0	0	0.0
Vowels	1	50.0	1	50.0	0	0.0
Basic Story	1	50.0	1	50.0	0	0.0

Table 11: Gates Kindergarten National Stanine Results N = 2

	6		7		8		9	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total	0	0.0	1	50.0	0	0.0	1	50.0
Initial Consonants	0	0.0	0	0.0	0	0.0	2	100.0
Final Consonants	0	0.0	1	50.0	0	0.0	1	50.0
Vowels	1	50.0	0	0.0	0	0.0	1	50.0
Basic Story	0	0.0	0	0.0	1	50.0	1	50.0

TerraNova for Kindergarten

Table 12 presents results for kindergarten MOST students in all schools who took the TerraNova test. Class/grade averages were not available for this test (please also note that these results are based on a small number (N=3) of MOST students).

- All MOST students' NPR scores in the Reading and Language subtests, as well as Total scores, were in the top quartile.

- All MOST students' NPR scores were above the national average of 50 in the Mathematics subtest.

Table 12: TerraNova Kindergarten National Percentile Rank Results N = 3

	0-25%		26-50%		51-75%		76-100%	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Reading	0	0.0	0	0.0	0	0.0	3	100.0
Language	0	0.0	0	0.0	0	0.0	3	100.0
Mathematics	0	0.0	0	0.0	2	66.7	1	33.3
Total	0	0.0	0	0.0	0	0.0	3	100.0

Summary

Overall, MOST students appeared to perform comparably with national norm groups, with several instances of percentile rankings above the national average. There was much more variation in outcomes when compared to their grade or classroom peers. It should be noted that several of these grade and classroom comparisons were based on very small sample sizes. An important point to reiterate is that no statistically reliable or causal conclusions can be drawn from this descriptive examination.

Achievement Analysis for Students Attending Jubilee Schools

This examination of student achievement focused on MOST scholarship students attending Jubilee Schools. Unlike the descriptive approach at non-Jubilee Schools, a more rigorous quasi-experimental design was used for these students. In this design, MOST students were inferentially compared with other students attending the same Jubilee School. The MOST and non-MOST students were “matched” on variables such as ethnicity, gender, and prior achievement. This is also referred to as a “matched-pair” design. In addition to the matched pair design, the longitudinal performance of MOST students was also compared with norms from the Iowa Tests of Basic Skills (ITBS) using a “repeated measures” design that examined data from Fall 2004 to Spring 2007. The questions that guided this examination were as follows:

1. What are the longitudinal trends of MOST student performance on the ITBS relative to national norms in Total Reading?
2. What are the longitudinal trends of MOST student performance on the ITBS relative to national norms in Total Language?
3. What are the longitudinal trends of MOST student performance on the ITBS relative to national norms in Total Mathematics without Computation?
4. How did MOST students attending Jubilee schools perform on 2007 Iowa Tests of Basic Skills (ITBS) Total Reading relative to non-MOST comparison students attending Jubilee schools who were matched on grade level, ethnicity, gender, and 2006 test scores?
5. How did MOST students attending Jubilee schools perform on 2007 Iowa Tests of Basic Skills (ITBS) Total Language relative to non-MOST comparison students attending Jubilee schools who were matched on grade level, ethnicity, gender, and 2006 test

scores?

6. How did MOST students attending Jubilee schools perform on 2007 ITBS Total Mathematics without Computation relative to non-MOST comparison students attending Jubilee schools who were matched on grade level, ethnicity, gender, and 2006 test scores?
7. How did MOST students attending Jubilee schools perform on the ITBS in Spring 2007 by grade level relative to national norms in Reading, Language Arts, Math, Social Studies, and Science?

Methods

Participants

MOST student participants. There were a total of 164 MOST students attending one of seven Jubilee Schools during the 2006-2007 academic year, which included grades Kindergarten through eight. Of these, parental consent to participate in the study was obtained from 147 (89.6%). The numbers of initially participating MOST students by school were as follows: 34 at De La Salle, 17 at Holy Names, 4 at Little Flower, 22 at St. Augustine, 14 at St. Joseph, 41 at St. John, and 15 at St. Patrick. From the group of MOST students for which parent consent was obtained, six students only had the student name and school name data available (i.e., had no test data available), and were therefore not included in the analyses. Of the remaining 141 MOST students, the majority were African American (77.3%; $n = 109$), whereas 12.8% ($n = 18$) were Hispanic, about 4.3% ($n = 6$) were White, 3.5% ($n = 5$) were of other unspecified ethnicities, and 2.1% ($n = 3$) were missing the ethnicity data. A slight majority of students were female (51.8%; $n = 73$), while 45.4% ($n = 64$) were male and 2.8% ($n = 4$) were missing the gender data.

Participants for repeated measures analyses. Out of 141 MOST students available, the following students were dropped: 59 students who did not attend the same school across the four time points used in the analysis: Fall 2004, Spring 2005, Spring 2006 and Spring 2007. Of the 82 MOST students remaining, 14 were excluded from the Reading analysis, 1 from the Language analysis, and 2 from the Math analysis due to missing scores for any of the four time points. Thus, 68, 81 and 80 MOST students (41.5%, 49.4%, and 48.8% of the total sample, respectively) were available for separate repeated measures analyses in Reading, Language, and Math. Table 13 summarizes the number of students selected for repeated measures analyses by school and subject area.

Table 13: Number of MOST Students Selected for Repeated Measures Analyses by School and Subject Area

School	Reading	Language	Math
De La Salle	16	16	16
Holy Names	12	12	12
St. Augustine	10	12	11
St. Joseph	8	8	8
St. John	19	24	24
St. Patrick	3	9	9
Total	68	81	80

Participants for matched sample analyses. Of the 141 MOST students, 42, 43 and 34 were dropped from the treatment (MOST) groups for the Reading, Language and Math analyses respectively due to missing Iowa Tests of Basic Skills (ITBS) Normal Curve Equivalent (NCE) scores for Spring 2006 and/or Spring 2007. The control groups (by subject area) were selected from 1,162 observations in a longitudinal ITBS Jubilee student test data file provided by the Memphis Catholic Diocese for Jubilee Schools. After excluding those students with missing

scores for Spring 2006 and/or Spring 2007 and those who were identified as MOST students, 270, 311 and 301 non-MOST Jubilee students were initially selected for the Reading, Language and Math analyses, respectively. In the end, 60, 63 and 63 MOST Jubilee students (36.6%, 38.4%, and 38.4% of the total sample respectively) were matched to non-MOST Jubilee students for the analyses on Reading, Language and Math ITBS scores based on: Grade level in spring 2006 and spring 2007, ethnicity, gender, and a difference in spring 2006 Normal Curve Equivalent (NCE) scores between MOST and non-MOST students of +/- 3 NCE points.

Measures

NCE scores from the Total Reading, Total Language, and Total Mathematics without Computation subtests of the ITBS were employed as outcomes in the study. The ITBS administrations used in the study were from Fall 2004, Spring 2005, Spring 2006, and Spring 2007. NCE scores were used because of their similarity to Percentile Ranks in terms of interpretability and their ability, unlike Percentile Ranks, to be used mathematically in calculations.

Analyses

Repeated measures analyses. Repeated measures analyses of variance (ANOVA) with test occasion as a within-subjects factor, were used to examine the trend in MOST student performance on the ITBS tests (i.e., if student performance changed significantly) across four test occasions. Total Reading, Total Language, and Total Math without Computation NCE scores were used as dependent variables for the analyses by subject area. The analyses examined the trend in student performance, with data from Fall 2004 serving as the baseline or starting point indicating student's initial level of achievement, and then at each subsequent spring. It should be noted that Fall 2004 was used as the baseline as this was the first year data were available.

Students had varying levels of years attending each school by this time. These analyses examined trends for students who had NCE scores available for all four administrations.

For each analysis, the result of the sphericity test was initially examined. The sphericity assumption must be met in order for the repeated measures analysis to be statistically valid. If the sphericity test was not significant (i.e., the assumption of sphericity was met), then a repeated measures analysis was performed. If the assumption of sphericity was violated, then the Huynh-Feldt correction was applied before conducting the repeated measures analysis to adjust the univariate test degrees of freedom and make the analysis statistically sound. The univariate approach was used to test the significance of the effects of test occasion. For a significant test occasion effect, paired t-tests using the Bonferroni adjustment were performed to make pairwise comparisons of the mean NCE scores across the four test occasions. As test occasion had $k = 4$ levels, the adjusted significance level was set as $0.05/(k*(k-1)/2) = 0.0083$.

Matched samples analyses. Group equivalence on NCE scores for Spring 2006 for the matched MOST and non-MOST Jubilee students was confirmed with one-way ANOVAs for each subject area to determine if the two groups examined were similar in achievement prior to spring 2007. Cohen's d effect size was computed as the mean difference of NCE scores (MOST – non-MOST) divided by the pooled standard deviation to also confirm the similarities between MOST and non-MOST Jubilee student groups in the matched samples. The effect size (or d) indicates the number of standard deviations by which, for example, the MOST and non-MOST group means differ. Thus, an effect size of +0.50 would indicate a half of a standard deviation advantage—a highly substantial educational impact. Generally, in education, effect sizes exceeding +0.20 would be considered meaningful and fairly strong when obtained for a whole-

school intervention. In addition, correlations were performed to test the relationship between Spring 2006 and Spring 2007 ITBS NCE scores.

Analysis of Covariance (ANCOVA) was used to assess the impact of MOST program participation on Spring 2007 ITBS Total Reading, Total Language, and Total Math without Computation NCE scores respectively, with students' Spring 2006 ITBS NCE score used as the covariate. ANCOVA statistically equates (adjusts the means of) the groups in 2006-07 on the covariate, meaning that any differences in achievement in spring 2007 can be evaluated as if the groups had similar achievement in spring 2006. Effect sizes were calculated for both unadjusted and adjusted mean differences within each subject area. Given the small sample size, it was not possible to disaggregate comparisons by grade level and school.

Descriptive Analyses

In addition to the inferential analyses, mean (average) MOST student performance on all ITBS subtests in Spring 2007 was computed by grade level.

Results

Repeated Measures Analyses Results

Reading. As the test of sphericity was rejected ($\chi^2(5)=17.116$, $p=0.0043$), the data did not meet the assumption of repeated measures analysis, and the Huynh-Feldt correction was used to adjust the univariate test degrees of freedom. The results indicated that MOST students' Total Reading NCE scores did not change significantly over four test occasions ($F(3,201) = 1.52$, $p = 0.2137$). As shown in Figure 1, MOST student performance in Reading was relatively stable across time.

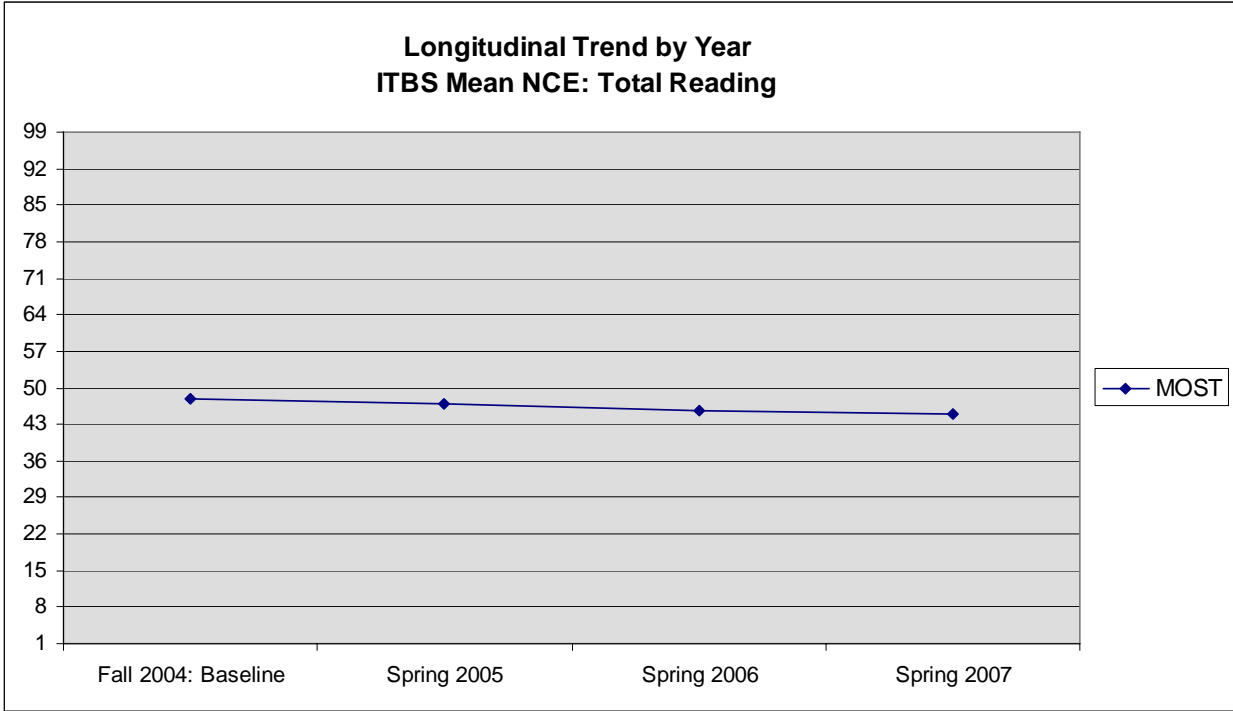


Figure 1: Longitudinal Trend in Total Reading Mean NCE for MOST Students

Language. Mauchly's test of sphericity ($\chi^2(5) = 10.376, p = 0.0652$) indicated that the assumption of sphericity was met. There was no significant effect on Total Language NCE scores for test occasion ($F(3,240) = 0.78, p = 0.5063$). As reflected in Figure 2, MOST student performance in Language did not change significantly over time.

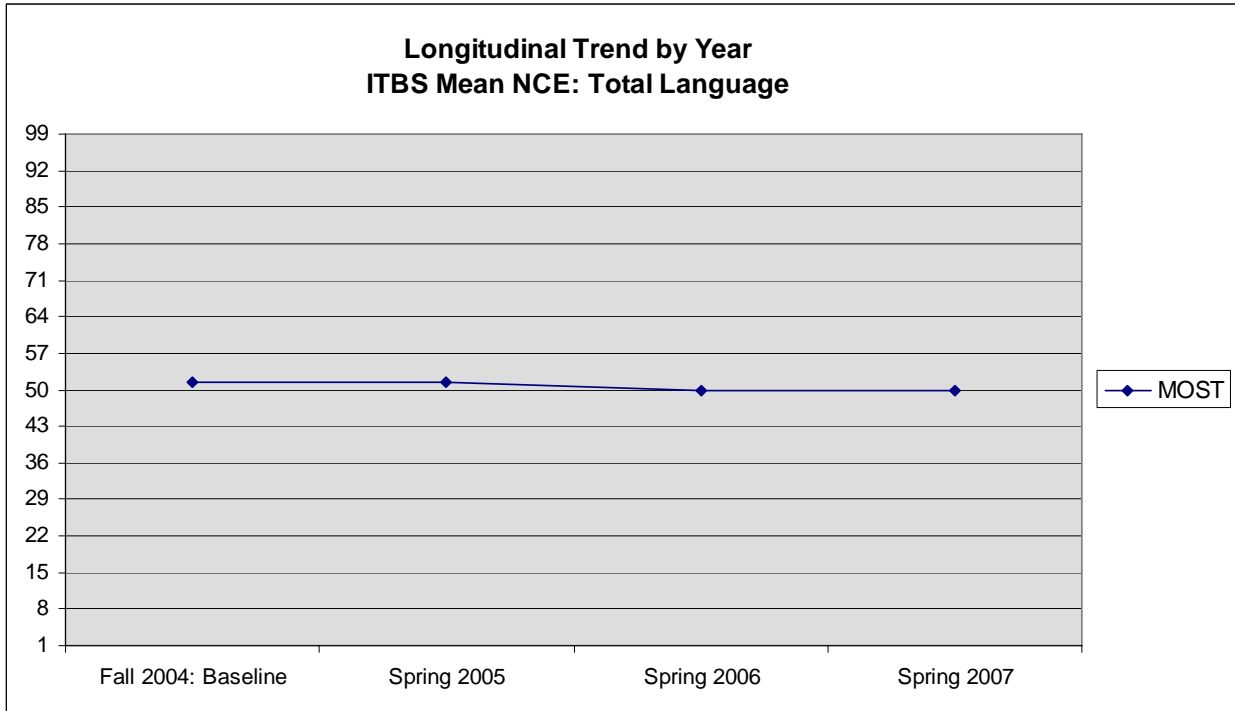


Figure 2: Longitudinal Trend in Total Language Mean NCE for MOST Students

Math. Mauchly's test of sphericity ($\chi^2(5) = 7.285, p = 0.2003$) indicated that the data met the assumption of sphericity. Test occasion had a significant effect on Total Mathematics without Computation NCE scores ($F(3,237) = 2.97, p = 0.0326$). Follow-up paired t-tests using the Bonferroni adjustment revealed that the mean score of 42.8 for Spring 2007 was significantly lower than the mean score of 46.9 for Spring 2005 ($t = -2.565, p = 0.0122$). Figure 3 shows that MOST student performance across the four test occasions did not vary dramatically.

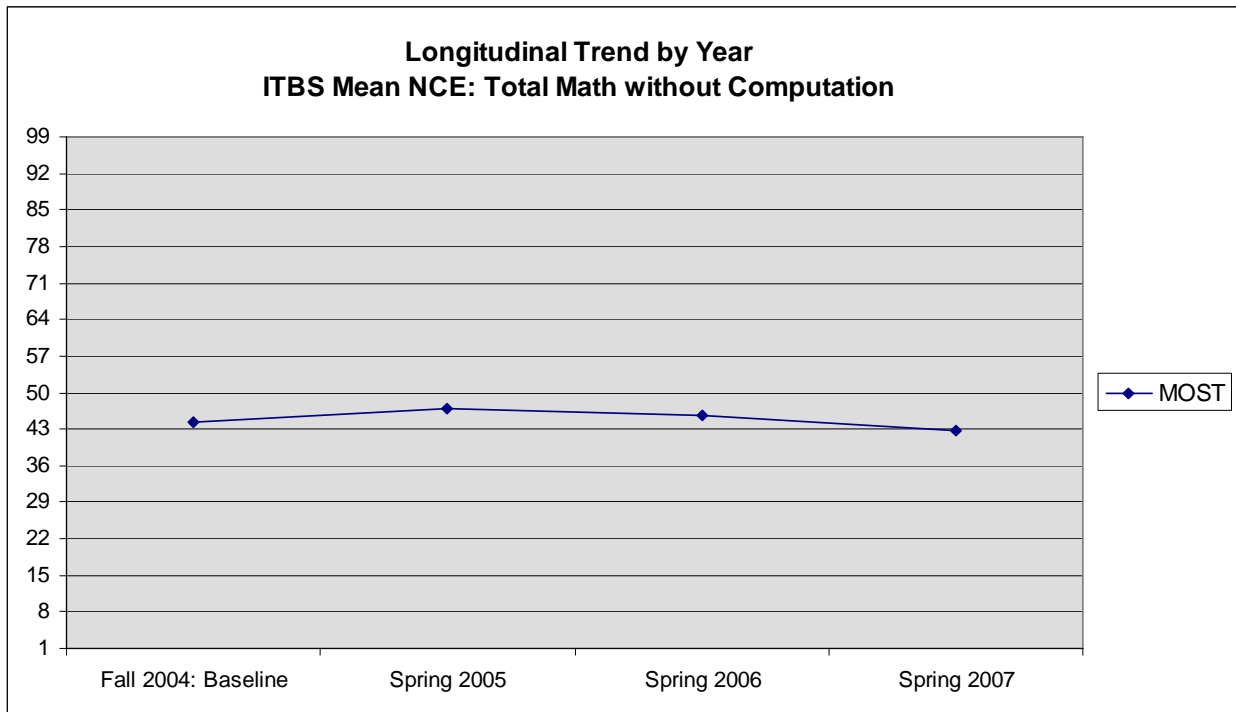


Figure 3: Longitudinal Trend in Total Math without Computation Mean NCE for MOST Students

Matched Sample Analysis Results

Reading. Utilizing the ITBS Total Reading NCE in Spring 2006, one-way ANOVA confirmed the similarity of the matching process ($F(1,118) = 0.01, p = 0.9100$), with a near zero effect size (see Table 14). Correlation between NCE scores for Spring 2006 and Spring 2007 was moderately strong ($r = 0.785$). The ANCOVA revealed no significant differences in the Spring 2007 Total Reading NCE scores between MOST and Non-MOST Jubilee students after controlling for prior year achievement ($F(1,117) = 2.25, p = 0.1359$). However, the adjusted effect size was -0.276 (see Table 15), reflecting a notable, but non-significant, program effect in Reading favoring Non-MOST students.

Language. Using the ITBS Total Language NCE in Spring 2006, one-way ANOVA supported the adequacy of the matching process ($F(1,124) = 0.00, p = 0.9532$), with an effect size near zero (see Table 14). Correlation between NCE scores for Spring 2006 and Spring 2007 was moderately strong ($r = 0.749$). As seen in Table 15, the ANCOVA result indicated no significant differences in the Spring 2007 Total Language NCE scores between MOST and Non-Most Jubilee students after controlling for prior year achievement ($F(1,123) = 0.01, p = 0.9199$), with a very small adjusted effect size that favored controls (-0.019).

Math. Based on the ITBS Total Math without Computation NCE in Spring 2006, one-way ANOVA supported the equality of the groups ($F(1,124) = 0.00, p = 0.9858$), with an effect size approaching zero (see Table 14). Correlation between NCE scores for Spring 2006 and Spring 2007 was moderate ($r = 0.651$). As shown in Table 15 the ANCOVA revealed no significant differences in the Spring 2007 Total Math without Computation NCE scores between MOST and Non-Most Jubilee students after controlling for prior year achievement ($F(1,123) = 0.05, p = 0.8208$), with a very small adjusted effect size that favored MOST students (0.041).

Table 14: Baseline Means, Standard Deviations, and Effect Sizes for MOST and non-MOST Students

Group	Reading			Language			Math		
	N	M	SD	N	M	SD	N	M	SD
MOST Students	60	47.05	15.90	63	49.95	15.11	63	45.37	14.99
Non-MOST students	60	47.38	16.31	63	49.79	15.19	63	45.41	15.05
Effect Size	$d=-0.021$			$d=0.011$			$d=-0.003$		

Table 15: ANCOVA Means, Standard Deviations, and Effect Sizes for MOST and non-MOST Students

Group	Reading				Language				Math			
	N	M	SD	Adj. M	N	M	SD	Adj. M	N	M	SD	Adj. M
MOST Students	60	42.92	15.78	43.04	63	50.30	15.35	50.24	63	42.86	15.65	42.87
Non-MOST students	60	45.68	13.90	45.56	63	50.37	15.76	50.43	63	42.38	17.42	42.36
Effect Size	$d=-0.187$				$d=-0.005$				$d=0.029$			
Adj. Effect Size	$d=-0.276$				$d=-0.019$				$d=0.041$			

Mean 2007 Performance by Grade Level and Subject

Table 16 provides the Spring 2007 ITBS mean NCE scores for MOST students by grade level and subject area.

- The mean NCE scores for MOST students in Kindergarten and grade 4 exceeded the national average of 50 on every test they participated in. The mean scores in Total Reading and Total Language for Kindergarten students and in Social Studies and Science for 4th graders substantially exceeded the national average.
- The mean NCE scores were substantially below the national average in Total Math without Computation for 3rd, 6th, 7th, and 8th graders, and in Reading and Science for 7th graders. The means for 7th and 8th grades should be interpreted with caution, however,

due to the small sample sizes for these grades ($n = 4$ and $n = 8$, respectively).

- The Total Language mean NCE scores for grades 5, 6, and 7 were only slightly below the national average of 50 (49.31, 49.68, and 49.50 respectively).

Table 16: Spring 2007 Mean NCE scores by Grade and Subtest for MOST Students

Grade	Statistics	Total Reading	Total Language	Total Math without Computation	Social Studies	Science
Kindergarten	Mean	70.71	65.64	57.36		
	N	7	11	11		
Grade 1	Mean	52.88	59.19	52.56	48.57	39.57
	N	16	16	16	16	14
Grade 2	Mean	46.05	46.50	43.35	42.56	41.25
	N	20	20	20	16	16
Grade 3	Mean	40.61	46.04	39.07	45.22	44.00
	N	28	28	28	27	27
Grade 4	Mean	54.67	53.42	51.25	67.00	63.75
	N	12	12	12	8	8
Grade 5	Mean	42.58	49.31	42.23	41.29	41.67
	N	26	26	26	21	21
Grade 6	Mean	46.32	49.68	39.84	48.93	46.87
	N	19	19	19	15	15
Grade 7	Mean	33.75	49.50	34.00	44.25	35.50
	N	4	4	4	4	4
Grade 8	Mean	47.25	47.75	36.75	45.25	45.00
	N	8	8	8	8	8
Total	Mean	46.82	50.94	43.98	46.53	44.18
	N	140	144	144	113	113

Summary of achievement findings for MOST students attending Jubilee Schools

- Total Math without Computation was the only subject area to show significant differences in test occasion, with the mean score for Spring 2007 significantly lower than Spring 2005.
- After controlling for prior achievement, there were no significant differences in NCE scores between MOST and matched non-MOST students in any subject area. While not

significant, the adjusted effect size in Reading (-0.276) was noteworthy and favored non-MOST students.

- Given the low sample sizes for MOST students included for the repeated measures and matched samples analyses (below 50% and below 40% respectively), the results should be interpreted with caution as the MOST student samples included in the analyses may not be representative of all MOST students attending Jubilee schools.
- The mean NCE scores for MOST students in Kindergarten and grade 4 exceeded the national average of 50 on every test they participated in.
- The mean NCE scores were substantially below the national average in Total Math without Computation for 3rd, 6th, 7th, and 8th graders, and in Reading, Math and Science for 7th graders. The means for 7th and 8th grades should be interpreted with caution, however, due to the small sample sizes for these grades ($n = 4$ and $n = 8$, respectively).