.

The Effects of Core Knowledge School Factors on State Test Achievement in North Carolina

Herbert J. Walberg[1]

Joseph Meyer[2]

March 2004

[1] Emeritus Research Professor of Education and Psychology, University of Illinois and Distinguished Visiting Fellow, Stanford University Hoover Institution.

[2] Vice President for Research, Metalytics, Lincoln, Illinois

I. Introduction

Statistical analyses were made of annual achievement progress on North Carolina's Department of Public Instruction tests. The analyses showed that Core Knowledge schools excelled the other schools in achievement progress in 8 of 10 comparisons of reading and mathematics in the five grade levels available for analysis.

The present report, however, which precedes the Core Knowledge vs. other schools comparison, concerns the achievement performance of Core Knowledge schools compared to one another and asks the following question: What factors lead to the achievement success of Core Knowledge schools? The analysis makes use of responses to the standard "Official Core Knowledge School Renewal Form 2003-2004," which concerns curricula, the use of TASA tests, grading periods, and other matters that may have bearing on the degree of implementation and achievement success of Core Knowledge schools.

II. Method of Research

A. Sample

The analyses make use of achievement test and demographic information about students in the Core Knowledge schools for the last two school years, 2001-2002, and 2002-2003, available from the North Carolina Department of Public Instruction.[3] Of the eight Core Knowledge schools in the state, data are unavailable for two—one a private school that did not participate in the testing program and the other a school that at the time

[3] The web site of the data source is http://www10.ncschoolcats.com.

of testing served only primary grade students that did not participate in the testing for the two years. This left six schools, although two lacked sixth- and seventh-grade data since they served earlier grades at the time the Department of Public Instruction collected the state data.

B. North Carolina State Achievement Testing Program

Like the other states, North Carolina has an elementary- and secondary-school testing program concentrated on mathematics and reading skill. The Department of Public Instruction describes the program as follows:

> "The competency goals and objectives adopted in 1998 included the Reading Comprehension and Mathematics In response to legislation passed by the 1989 North Carolina General Assembly, the State Board of Education developed and initially implemented End-of-Grade Tests for grades 3 through 8 in the areas of reading and mathematics effective with the 1992–93 school year. These curriculum-based multiple-choice achievement tests are specifically aligned to the North Carolina Standard Course of Study and include a variety of strategies to measure the academic performance of North Carolina students.
>
> The North Carolina State Board of Education tests for each grade are organized into four strands: (1) number sense, numeration, and numerical operations; (2) spatial sense, measurement, and geometry; (3) patterns, relationships, and functions; and (4) data, probability, and statistics. The mathematics EOG tests are administered in two parts: Calculator Inactive and Calculator Active. Students are not allowed to use calculators during the Calculator Inactive part of the test. Students are allowed to use calculators during the Calculator Active part of the test. Both parts of the test require students to interpret information from problems in context in order to generate the appropriate responses to the test questions. The North Carolina End-of-Grade (EOG) Test–Reading Comprehension assesses reading by having students read both literary and informational selections and then answer questions related to the selections. Knowledge of vocabulary is assessed indirectly through application and understanding of terms within the context of the selections and questions.
>
> The selections chosen for the reading tests reflect reading for various purposes such as literary experience, gaining information, and performing a task. Literary texts include fiction, poetry, drama, and literary nonfiction

> such as biographies, letters, journals, and essays. Informational texts include content areas (art, science, mathematics, social studies, etc.) and consumer or practical selections (pamphlets, reviews, recipes, how-to, etc.).
>
> Understanding Scores for the EOG Tests: Students take the state-required multiple-choice North Carolina EOG Tests in Reading and Mathematics during the final weeks of the school year. Reports of student scores are printed soon after scoring and sent to schools for distribution to parents."[4]

The North Carolina state test seems a good choice for evaluating school policies, practices, and curricula for several reasons. Tests, particularly national commercial tests, may vary greatly in the degree that they reflect the goals of a given school's curriculum and instructional emphases. For example, because schools may adapt their curricula to the commercial tests they use, such as the Metropolitan Achievement Tests, they are likely to do better than other schools on the tests they have chosen. Tests required by states, however, put schools on an even footing, and reflect what the representatives of citizens in the state think is important.

Moreover, because of the federal No Child Left Behind legislation, state requirements, wider availability of school "report cards, and the pressures of accountability and choice, most schools are under increasingly under pressure to perform well on the state states. Along with the National Assessment of Educational Progress, state tests are becoming "the currency of the realm.'

Finally, all regular schools in each state are required to participate in state testing programs. Hence, the complete universe of schools can be analyzed rather than subjectively choosing typical or, in the case of comparative studies, "matched schools," neither of which is considered scientific.

---

[4] Public Schools of North Carolina, State Board of Education, "Assessment Brief: Understanding North Carolina End-of-Grade Testing," March 1, 2004 • Vol. 5, No. 3. http://www.ncpublicschools.org/accountability/testing/briefs/ABriefEOG04.pdf.

C. Renewal Information on Core Knowledge Schools

The staffs of schools that wished to remain an official Core Knowledge school filled out the "Official Core Knowledge School Renewal Form 2003-2004." The questions that can be usefully asked about the items on the Renewal Form are as follows:

Item	Question
Percent Implementation	Is higher achievement attributable to degree of implementation?
Reading program	Are some programs more effective than others are?
Mathematics program	Are some programs more effective than others are?
Baltimore Project	Does the Baltimore Project lead to higher achievement?
TASA tests employed	Do TASA tests lead to higher achievement?
Pearson textbooks	Do Pearson textbooks lead to higher achievement particularly in History and Geography?
Total student population	Do small schools achieve more than larger ones?
Demographic profile	Do schools with concentrations of racial/ethnic groups achieve more than others do?
Percentage of students on free/reduced price lunch	Do schools with fewer students in poverty achieve more?
Number of teachers	Do schools with more or fewer teachers achieve more than others do?
Multi-age classes	Do schools with multi-age classes achieve more than others do?
State, district,	Do schools subject to various standards achieve more

Item	Question
or local standards	than others do?
Willing to share standards with others	Do schools willing to share standards achieve more than other schools do?
Planning time for teachers per month	Do schools where teachers have more planning time achieve more than other schools do?
Common planning time per month	Do schools where teachers have larger amounts of common planning time achieve more than other schools do?
Six or nine week grading periods	Do schools with shorter grading periods achive more than other schools do?
Issues or barriers to full implementation of Core Knowledge at school	Do schools with issues or barriers achieve less well than other schools do?
Core Knowledge professional development	Do schools completing more workshops or particular types of workshops do better than other schools do?
Representatives sent to National Conference	Do schools that participate in the conference to greater degrees achieve more than other schools do?
Other staff development outside Core Knowledge	Do schools that participate in certain types of additional staff development achieve more than other schools do?
Goals for Core Knowledge	Do schools with some goals achieve more than other schools do?
Upcoming Core Knowledge events	Do schools with plans for certain kinds of events do better than other schools do?

D. Statistical Procedures

Initial analysis of the 1592 eligible Core Knowledge students' test results showed that the much of the variation in their scores, about 80 percent, is attributable to differences among students rather than differences among the Core Knowledge schools. As many studies have shown, achievement is a continuously accumulative process, and variations among schools in any given year make for relatively small differences in students' achievement compared with their previous experiences at home and, in the later grade levels, in school. During the first 18 years of life, for example, only about 8 percent of the time is spent in school. For this reason, variations in the quality of schooling and particular school features and practices are often dwarfed and difficult to detect compared to the family influences on intellectual development and achievement.[5]

Because the majority of the variation was attributable to differences among students, the analysis was designed to take into account the variations among them. Specifically, during the analysis, "value added" gains from the 2001-2002 to the 2002-2003 school year were calculated. As explained further below, the analysis also took into account the poverty and minority status of each student. Only students with complete information for both school years were included in the analyses.

In addition, the initial analyses confirmed that the data were statistically clustered within schools, which could be expected since students are influenced by features and conditions within their schools and communities that tend may tend to make them similar to

[5] Herbert J. Walberg, "Improving Educational Productivity: An Assessment of Extant Research," a paper prepared for the conference and book The Scientific Basis of Educational Productivity sponsored by the American Psychological Association and the Mid-Atlantic Laboratory for Student Success; to be published by Information Age Publishing, Greenwich, CT., 2004.

one another and different from students in other schools. For example, a highly effective principal or school board may confer higher test scores on students within their purview, which sets them apart from students in other schools.

The consequence of such "clustered" or "correlated effects" is that the student scores within a school are not independent of one another as required for statistical inference. Thus, the basis of estimating school effects is a combination of the number of schools, the number of students in a school, and the underlying correlation structure (i.e. how the test scores of students in the same school are correlated with each other). Even though the sample of students, 1592, for the present evaluation is seemingly very large, the valid sample size is smaller, which avoids coming to misleading positive or negative conclusions that have often characterized previous studies of school effects. To account precisely and simultaneously for such individual student variations and clustered school effects, generalized linear models were employed.[6]

To be discussed below, descriptive statistics about the sample in terms of frequencies, percentages, means, and standard deviations are shown in Tables 1 and 2. Table 3 displays each school's response to the Core Knowledge survey. The pretest-, ethnicity- and poverty-adjusted regression residuals from the GEE models were computed separately for each combination of subject and grade. The median values are shown in Table 4.[7] In the models, for ease of interpretation across grades, the residuals were

[6] Also called hierarchical linear models. The method employed Generalized Estimating Equations, a statistically efficient way of fitting such data. See K. Y. Liang and S. L Zeger, Longitudinal data analysis using generalized linear models, Biometrika 73: 13-22, 1986 and S.L Zeger and K.Y. Liang. Longitudinal data analysis for discrete and continuous outcomes. Biometrics, 42(1): 121-30, 1986). The statistical package SAS 8.02 GENMOD procedure was employed. See SAS/STAT User's Guide, Version 8. (Cary, NC: SAS Institute Inc., 1999).

[7] The corresponding mean and standard deviation are shown in Appendix A.

converted to z-scores, where, for example, 0.2 (or -0.2) indicated that the score was 0.2 standard deviation above (or below) the Core Knowledge school average, and 0 indicates that it a school was average. Similarly, in Table 5, we summarized the regression residuals by school characteristics. Characteristics which all six schools shared (e.g. all attended Getting Started workshop) were excluded here. The level of statistical significance was set at 0.05 (two-sided). Each of the tables deserves comment.

III. Results

A. Student Demographic Characteristics of Core Knowledge Schools

Table 1 shows the characteristics of students in each Core Knowledge school separately, all Core Knowledge schools, and the 533,919 students in non-Core Knowledge schools in North Carolina. The samples contain only students with complete achievement information for the two recent academic years and with complete demographic information on minority status. The number of Grade 3 through Grade 7 students is 1592, and the number in each Core Knowledge school varies from 91 to 386.

The numbers of students vary even more widely at certain grades. School E, for example, has more than four times as many third graders as does School A. There are substantially fewer students in the later grades, and three schools had no qualifying students in Grades 6 and 7.[8]

As Table 1 shows, few American Indian, Asian, Hispanic, and Multi-Racial students are represented in either Core Knowledge or other schools. Greater percentages of Blacks

[8] There were insufficient numbers of schools and students available analyzing Grade 8 scores, and the achievement test for Grade 8 is less comparable to the prior grade than for the other grades represented.

attend Core Knowledge schools than other schools as do slightly smaller percentages of Whites.

Core Knowledge schools have greater percentages of students whose families fully pay for their lunches. Smaller percentages of Core Knowledge students qualify for free lunch—an index of higher poverty than reduced-price lunch status.

B. Race/Ethnicity and Poverty Effects on Achievement

Table 2A and 2B show the reading and mathematics achievement levels of Black, Hispanic, White, and Other students. Because their numbers were so small, American Indian, Asian, and Multi-Racial students were included in the Other category. As the National Assessment of Educational Progress and other large-scale national and state surveys have shown, Whites score higher on average than Blacks and Hispanics on both the reading and mathematics achievement tests. As also found in previous surveys, poverty pervasively lowers achievement. Because of these differences, race/ethnicity and poverty status were taken into consideration in the analysis, as was the effect of each student's pretest or score for the previous academic year 2001-2002 in calculating indexes of annual progress.

C. Characteristics of Core Knowledge Schools

Table 3 shows the results of the "Official Core Knowledge School Renewal Form 2003-2004." Several trends are noteworthy.

First, with respect to several Core Knowledge principles implicitly and plausibly featured as positive in the Renewal Form, all schools reported affirmatively. These include attendance at a Getting Started Workshop, participation in a Lessons and Assessments

Workshop, used Pearson textbooks, employing state standards, and having staff development other than national conference participation.

These findings may deserve celebration since they suggest strong and uniform adherence to Core Knowledge implementation ideals and since such adherence is thought to increase achievement. From a research point of view, however, it means that schools with and without these ideals are not represented such that their possible can be detected in the present sample.

Nearly the same is true of items in which only one or two schools faltered. These include having no barriers to implementation, attendance at a Core Knowledge Institute, participation in an Overview Workshop, and using the Open Court reading program and the Saxon mathematics program.[9] The same problem may be seen in the degree or percentage of Core Knowledge implementation, which varied from 94 to 100 percent, aside from one school with a still high 86 percent implementation percentage.

Again, if the schools vary only slightly the possible causes of achievement differences, it is unlikely that the effects of the causes may be found. To find causes confidently would require larger differences in the possible causes and larger samples of schools, say 20 or 30. In the present study, relative implementation uniformity and the small sample is compounded by missing and ambiguous responses to several of the questions.

[9] None of the schools made use of the Baltimore Project or employed the TASA tests, which again left no comparison group. The goals for the school year were highly diverse, which made them impossible to categorize or rate.

D. Achievement Progress in the Sample Schools

Table 4 shows the annual progress of each of the schools adjusted for previous achievement, race/ethnicity, and poverty. The results in Table 4 are medians[10] of achievement progress[11] of each school for each combination of grade and subject. A positive number indicates the school did better than the average of all Core Knowledge schools with available data, and negative score means the school did less well than the others did.

Given the uniformity of the schools' compliance with Core Knowledge noted in Table 3, correspondingly little variation in scores might be expected. This is the exact result. The largest progress indicator in the table is .26 for mathematics in School D, which corresponds to a percentile of 60. The worst result, -.39, for School C in mathematics has a corresponding percentile of 35. The other results reported in Table 4 were even closer to the 50th percentile.

Thus, the North Carolina Core Knowledge schools uniformly and largely abided by Core Knowledge ideals represented on the Renewal From and attained relatively the same achievement results relative to one another. Again, these are desirable results from the point of Core Knowledge students, but less than ideal for detecting the causes of their differences. Perhaps a physical analogy may be useful: How can training effects be uncovered from an extremely close and fast 50-yard dash of runners uniformly well trained.

[10] The corresponding means and standard deviations are shown in the Appendix.

[11] Technically, these are medians of Z-score standardized regression residuals.

IV. School Characteristic Effects on Academic Progress

Tables 5A and 5B for reading and mathematics progress show what can be inferred from the previous analyses and comments. The last column shows that in no case did a Core Knowledge school characteristic have a statistically significant effect on academic progress adjusted for the previous year's scores, race-ethnicity, and poverty. The customary level of statistical significance is .05, and the lowest level in the tables is .22, which is not even close.

Thus, the analysis suggests an answer to the chief question: What factors lead to the achievement success of Core Knowledge schools? The answer is that, among the factors represented on the Renewal Form, none can be detected. This does not mean that a different sample would lead to different results. Larger samples with greater variations in implementation of Core Knowledge principles might show large and significant effects on achievement progress. It seems likely, for example, that if true implementation effects exist, that a sample with many schools with percent implementation rates that vary between 10 and 100 percent would differ significantly in achievement progress.

V. Conclusion

To be documented in the final report in this series, statistical analyses showed that North Carolina Core Knowledge schools excelled the other schools in the state in achievement progress in 8 of 10 comparisons of reading and mathematics in the five grade levels available for analysis. As the present report documents, the Core Knowledge schools in North Carolina also very largely adhered to the Core Knowledge Foundation's implementation requirements that are monitored and promoted within the Foundation's Renewal Form.

Because none of the schools in the sample did a poor job of implementing the requirements, no association between the school requirements and annual achievement progress was found. In future research, a larger sample of schools in another state or in multiple states that have wider variability in the implementation requirements might show that some factors significantly influence outcomes.

On the other hand, if Core Knowledge schools elsewhere are similar to those in North Carolina in largely meeting the present renewal requirements, additional research may not show such implementation effects. If so, the Foundation may wish to consider two courses. The first is to raise the renewal bar even higher so that schools can strive for even more advanced implementation levels.

The second course would be to encourage wider and deeper use of teaching methods and school policies that previous research has shown to promote achievement in a variety of curricula and that may enable Core Knowledge to achieve greater heights of learning.[12] The second course would then emphasize policies and practices that can be improved rather than the Core Knowledge curriculum and implementation requirements, which may be difficult to improve further.

[12] Provided with the submission of this report, see Herbert J. Walberg, "Improving Educational Productivity: An Assessment of Extant Research" in Rena Subotnik and Herbert J. Walberg, editors, *The Scientific Basis of Educational Productivity* (Greenwich, CT.: Information Age Publishing, in process) prepared for a conference sponsored by the American Psychological Association and the Mid-Atlantic Laboratory for Student Success, May 2004.

Table 1. Student characteristics: North Carolina state database, 3rd – 7th grade

	Core Knowledge Schools*						All CK Schools	Other Schools
	<A>	<B>	<C>	<D>	<E>	<F>		
# of Students:								
Grade 3	31	79	71	80	130	77	468	106,002
Grade 4	30	61	63	80	98	61	393	104,091
Grade 5	30	65	47	78	114	45	379	107,337
Grade 6	0	59	49	76	0	0	184	108,170
Grade 7	0	58	38	72	0	0	168	108,319
Total	91	322	268	386	342	183	1592	533,919
% American Indian	0	<1%	0	<1%	<1%	0	<1%	2%
% Asian	0	<1%	<1%	1%	<1%	0	<1%	2%
% Black	9%	56%	63%	8%	24%	95%	40%	30%
% Hispanic	0	5%	1%	<1%	16%	1%	5%	6%
% Multi-racial	0	<1%	1%	<1%	3%	<1%	1%	2%
% White	91%	38%	34%	89%	56%	4%	53%	58%
% Free Lunch	4%	20%	22%	2%	22%	27%	16%	37%
% Reduced Pay Lunch	0	10%	9%	0	7%	13%	7%	9%
% Full Pay Lunch	96%	70%	69%	98%	71%	60%	77%	54%

*Grades available in each school: <A> K-5; <B> K-7; <C> K-7; <D> 1-7; <E> K-5; <F> K-5.
**Limited English Proficiency
***Physically, mentally, or cognitively impaired.

Table 2A. Student characteristics and post-test reading performances, by grade, ethnicity, and poverty: Unadjusted mean ± SD

	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7	
	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD
Black	200	242.2 ± 8.5	163	249.1 ± 8.4	141	252.6 ± 7.0	68	254.4 ± 7.0	57	258.9 ± 8.4
Hispanic	22	241.8 ± 10.2	15	249.6 ± 9.4	10	250.0 ± 9.6	5	256.4 ± 4.9	2	271.5 ± 6.4
White	219	250.7 ± 8.1	193	255.6 ± 8.6	205	258.5 ± 7.4	106	260.5 ± 7.6	104	264.8 ± 7.4
Other	14	246.9 ± 10.4	7	256.1 ± 9.8	8	257.6 ± 3.6	4	260.8 ± 5.4	3	252.0 ± 9.6
Free lunch	63	241.8 ± 8.5	70	248.7 ± 8.2	61	251.7 ± 7.2	25	254.0 ± 6.7	12	260.8 ± 7.1
Reduced pay	33	242.7 ± 9.4	20	248.1 ± 7.2	23	253.3 ± 5.2	6	256.3 ± 3.3	15	257.1 ± 8.3
Full pay	359	247.6 ± 9.3	288	253.8 ± 9.1	280	257.1 ± 7.8	152	258.9 ± 8.0	139	263.4 ± 8.3

Note:
(1) Because their numbers were small, American Indian, Asian, and Multi-Racial students were included in the Other category.
(2) Three of the six schools do not have 6^{th} and 7^{th} grade.

Table 2B. Student characteristics and post-test math performances, by grade, ethnicity, and poverty: Unadjusted mean ± SD

	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7	
	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD
Black	201	248.1 ± 5.7	163	254.3 ± 7.1	141	256.8 ± 7.0	68	259.5 ± 6.3	57	262.9 ± 9.0
Hispanic	22	247.4 ± 5.3	15	256.7 ± 9.4	11	255.0 ± 8.7	5	265.4 ± 4.2	2	274.0 ± 4.2
White	219	253.9 ± 5.4	193	260.7 ± 6.1	207	263.4 ± 7.9	106	267.4 ± 7.6	105	271.8 ± 9.8
Other	14	251.9 ± 6.5	7	260.4 ± 6.6	8	260.0 ± 5.3	4	265.8 ± 6.1	3	259.7 ± 14.0
Free lunch	63	247.4 ± 5.5	70	254.7 ± 7.9	61	256.8 ± 8.3	25	260.0 ± 7.6	13	265.5 ± 8.6
Reduced pay	33	248.2 ± 7.0	20	254.3 ± 6.5	24	258.4 ± 6.9	6	257.0 ± 3.2	15	263.6 ± 6.5
Full pay	359	251.8 ± 6.0	288	258.8 ± 7.0	282	261.5 ± 8.0	152	265.4 ± 7.8	139	269.4 ± 10.8

Note:
(1) Because their numbers were small, American Indian, Asian, and Multi-Racial students were included in the Other category.
(2) Three of the six schools do not have 6^{th} and 7^{th} grade.

Table 3. Core Knowledge school characteristics

School	<A>	<B>	<C>	<D>	<E>	<F>
# of Teachers	13[a]	38	13	37	38	14
Planning Time (hours/month)*	4	13/26[b]	15	15	13	15/20[c]
Grading Period (weeks)	12	(NA)	(NA)	(NA)	9	(NA)
% Implementation	100%	94%	97%	95%	100%	86%
Barriers to Implementation	No	No	No	No	No	Yes
Attended Core Knowledge Institute	No	Yes	Yes	Yes	No	Yes
Attended Overview Workshop	No	Yes	Yes	Yes	Yes	Yes
Attended Getting Started Workshop	Yes	Yes	Yes	Yes	Yes	Yes
Attended Lessons and Assessments Workshop	Yes	Yes	Yes	Yes	Yes	Yes
Use Baltimore Project	No	No	No	No	No	No
Take TASA Tests	No	No	No	No	No	No
Use "Open Court" reading and "Saxon" math program	No	Yes	Yes	Yes	No	Yes
Use Pearson Textbooks	Yes	Yes	Yes	Yes	Yes	Yes
Use State Standards	Yes	Yes	Yes	Yes	Yes	Yes
Use District Standards	(NA)	Yes	Yes	(NA)	(NA)	Yes
Use Local Standards	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
Willing to Share Standards	Yes	No	No	No	Yes	No
Has Multi-age Classes	Yes	(NA)	(NA)	(NA)	Yes	(NA)
# of Representatives Sent to National Conference	3	4	3	2	6 ?	2
# of Staff Who Have Presented at National Conference	0	0	0	0	12	0
Had Other Staff Development	Yes	Yes	Yes	Yes	Yes	Yes

*Surveys which gave this in periods/day were converted assuming 40 minutes/period and 20 days/month.
[a]Reported 13 teachers and 8 paraprofessionals; here we used 13.
[b]13 hours/month for K – 5th grade, 26 hours/month for 6th and 7th grade.
[c]15 hours/month for K - 2nd grade, 20 hours/month for 3rd – 5th grade.
Note: Open Court reading and Saxon math were combined because they were commonly offered together.

Table 4. Comparison of the six Core Knowledge schools: Median regression residuals of z-scores, adjusting for ethnicity and poverty, by subject, grade, and school

Reading

School	<A>	<B>	<C>	<D>	<E>	<F>
Grade 3	0.12	-0.03	0.09	-0.11	0.14	-0.10
Grade 4	0.11	0.03	0.21	0.22	-0.10	-0.34
Grade 5	0.08	0.21	0.24	-0.03	0.03	-0.23
Grade 6	(NA)	-0.00	-0.06	0.12	(NA)	(NA)
Grade 7	(NA)	0.01	-0.13	0.15	(NA)	(NA)

Math

School	<A>	<B>	<C>	<D>	<E>	<F>
Grade 3	0.14	0.11	0.11	-0.09	-0.03	-0.15
Grade 4	0.09	0.23	0.34	-0.18	-0.06	-0.22
Grade 5	0.21	-0.02	0.26	-0.24	0.07	-0.29
Grade 6	(NA)	0.19	-0.39	0.26	(NA)	(NA)
Grade 7	(NA)	-0.11	0.11	0.03	(NA)	(NA)

Note:
(1) Schools <A>, <E>, and <F> do not have 6^{th} and 7^{th} grade.
(2) For the corresponding mean ± SD, see Appendix A.

Table 5A. Core Knowledge school characteristics and student reading performances: Regression residuals of z-scores (median presented for each grade, mean ± SD presented for all grades combined), adjusting for ethnicity and poverty

		Grade					All grades	P-value*
		3	4	5	6	7		
# of Teachers	<20	0.06	-0.07	-0.04	-0.06	-0.13	-0.03 ± 0.74	0.41
	20+	0.03	0.04	0.05	0.05	0.09	0.01 ± 0.62	
Planning Time (hours/month)	4	0.12	0.11	0.08	(NA)	(NA)	0.03 ± 0.70	0.44
	13-15	0.05	0.06	0.05	0.03	0.10	0.04 ± 0.65	
	20-26	-0.10	-0.34	-0.23	-0.00	0.01	-0.15 ± 0.66	
% Implementation	80%	-0.10	-0.34	-0.23	(NA)	(NA)	-0.25 ± 0.72	0.34
	90%	-0.04	0.12	0.10	(NA)	(NA)	0.03 ± 0.67	
	100%	0.14	-0.07	0.03	(NA)	(NA)	0.03 ± 0.59	
Barrier to Implementation	Yes	-0.10	-0.34	-0.23	(NA)	(NA)	-0.25 ± 0.72	0.22
	No	0.05	0.08	0.05	(NA)	(NA)	0.03 ± 0.64	
Attended Core Knl. Institute	Yes	-0.05	0.04	0.02	(NA)	(NA)	-0.01 ± 0.68	0.62
	No	0.14	-0.07	0.03	(NA)	(NA)	0.03 ± 0.59	
Attended Overview Workshop	Yes	0.03	-0.01	0.03	(NA)	(NA)	-0.00 ± 0.66	0.53
	No	0.12	0.11	0.08	(NA)	(NA)	0.03 ± 0.70	
Use "Open Court" reading & "Saxon" math program	Yes	-0.05	0.04	0.02	(NA)	(NA)	-0.01 ± 0.68	0.62
	No	0.14	-0.07	0.03	(NA)	(NA)	0.03 ± 0.59	
Use District Standards	Yes	0.01	-0.07	0.10	-0.04	-0.04	-0.02 ± 0.68	0.47
	Missing	0.05	0.05	0.02	0.12	0.15	0.01 ± 0.63	
Willing to Share Standards	Yes	0.14	-0.07	0.03	(NA)	(NA)	0.03 ± 0.59	0.62
	No	-0.05	0.04	0.02	(NA)	(NA)	-0.01 ± 0.68	
Has Multi-age Classes	Yes	0.14	-0.07	0.03	(NA)	(NA)	0.03 ± 0.59	0.62
	Missing	-0.05	0.04	0.02	(NA)	(NA)	-0.01 ± 0.68	
# of Representatives Sent to National Conference	2	-0.10	-0.02	-0.11	0.12	0.15	-0.08 ± 0.70	0.59
	3	0.10	0.18	0.14	-0.06	-0.13	0.08 ± 0.72	
	4	-0.03	0.03	0.21	-0.00	0.01	0.02 ± 0.60	
	6	0.14	-0.09	0.03	(N/A)	(N/A)	0.03 ± 0.55	
Staff Presented at National Conference	Yes	0.14	-0.09	0.03	(NA)	(NA)	0.03 ± 0.55	0.76
	No	-0.03	0.05	0.03	(NA)	(NA)	-0.01 ± 0.68	

*Chi-square test from GEE model, testing for the effect of each characteristic.
Note: Only three schools have 6th and 7th grade. No summary statistic was provided when all three schools share the same characteristics (e.g. none had barriers to implementation).

Table 5B. Core Knowledge school characteristics and student math performances: Regression residuals of z-scores (median presented for each grade, mean ± SD presented for all grades combined), adjusting for ethnicity and poverty

		Grade					All grades	P-value*
		3	4	5	6	7		
# of Teachers	<20	0.06	-0.10	-0.04	-0.39	0.11	0.00 ± 0.68	0.79
	20+	-0.03	-0.01	-0.07	0.20	-0.04	-0.01 ± 0.52	
Planning Time (hours/month)	4	0.14	0.09	0.21	(NA)	(NA)	0.15 ± 0.59	0.35
	13-15	-0.02	0.04	-0.05	-0.02	0.04	0.01 ± 0.58	
	20-26	-0.15	-0.22	-0.29	0.19	-0.11	-0.12 ± 0.57	
% Implementation	80%	-0.15	-0.22	-0.29	(NA)	(NA)	-0.23 ± 0.61	0.46
	90%	0.01	0.06	-0.09	(NA)	(NA)	0.02 ± 0.60	
	100%	0.03	-0.01	0.10	(NA)	(NA)	0.04 ± 0.51	
Barrier to Implementation	Yes	-0.15	-0.22	0.01	(NA)	(NA)	-0.23 ± 0.61	0.22
	No	0.02	0.04	-0.29	(NA)	(NA)	0.02 ± 0.57	
Attended Core Knl. Institute	Yes	-0.03	-0.06	-0.13	(NA)	(NA)	-0.02 ± 0.60	0.32
	No	0.03	-0.01	0.10	(NA)	(NA)	0.04 ± 0.51	
Attended Overview Workshop	Yes	-0.03	-0.06	-0.09	(NA)	(NA)	-0.01 ± 0.58	0.25
	No	0.14	0.09	0.21	(NA)	(NA)	0.15 ± 0.59	
Use "Open Court" reading & "Saxon" math program	Yes	-0.03	-0.06	-0.13	(NA)	(NA)	-0.02 ± 0.60	0.32
	No	0.03	-0.01	0.10	(NA)	(NA)	0.04 ± 0.51	
Use District Standards	Yes	0.05	0.04	-0.08	0.04	-0.06	0.00 ± 0.64	0.52
	Missing	-0.04	-0.07	-0.02	0.26	0.03	-0.01 ± 0.52	
Willing to Share Standards	Yes	0.03	-0.01	0.10	(NA)	(NA)	0.04 ± 0.51	0.32
	No	-0.03	-0.06	-0.13	(NA)	(NA)	-0.02 ± 0.60	
Has Multi-age Classes	Yes	0.03	-0.01	0.10	(NA)	(NA)	0.04 ± 0.51	0.32
	Missing	-0.03	-0.06	-0.13	(NA)	(NA)	-0.02 ± 0.60	
# of Representatives Sent to National Conference	2	-0.11	-0.21	-0.24	0.26	0.03	-0.11 ± 0.56	0.22
	3	0.13	0.22	0.24	-0.39	0.11	0.12 ± 0.69	
	4	0.11	0.23	-0.02	0.19	-0.11	0.05 ± 0.55	
	6	-0.03	-0.06	0.07	(N/A)	(N/A)	0.01 ± 0.48	
Staff Presented at National Conference	Yes	-0.03	-0.06	0.07	(NA)	(NA)	0.01 ± 0.48	0.84
	No	-0.01	-0.01	-0.09	(NA)	(NA)	-0.01 ± 0.60	

*Chi-square test from GEE model, testing for the effect of each characteristic.
Note: Only three schools have 6th and 7th grade. No summary statistic was provided when all three schools share the same characteristics (e.g. none had barriers to implementation).

Appendix A. Comparison of the six Core Knowledge schools: Mean (± SD) regression residuals of z-scores, adjusting for ethnicity and poverty, by subject, grade, and school

Reading

School	<A>	<B>	<C>	<D>	<E>	<F>
Grade 3	0.04 ± 0.70	-0.06 ± 0.71	0.15 ± 0.66	-0.08 ± 0.65	0.10 ± 0.60	-0.14 ± 0.73
Grade 4	0.06 ± 0.67	-0.01 ± 0.64	0.19 ± 0.80	0.18 ± 0.74	-0.04 ± 0.53	-0.37 ± 0.73
Grade 5	-0.02 ± 0.76	0.19 ± 0.52	0.20 ± 0.85	-0.11 ± 0.71	-0.00 ± 0.51	-0.28 ± 0.69
Grade 6	(NA)	-0.01 ± 0.44	-0.06 ± 0.64	0.01 ± 0.66	(NA)	(NA)
Grade 7	(NA)	-0.00 ± 0.57	-0.03 ± 0.62	0.00 ± 0.61	(NA)	(NA)

Math

School	<A>	<B>	<C>	<D>	<E>	<F>
Grade 3	0.19 ± 0.60	0.07 ± 0.57	0.08 ± 0.60	-0.17 ± 0.58	0.00 ± 0.47	-0.10 ± 0.68
Grade 4	0.08 ± 0.57	0.23 ± 0.66	0.23 ± 0.81	-0.13 ± 0.52	-0.04 ± 0.52	-0.33 ± 0.55
Grade 5	0.17 ± 0.62	-0.11 ± 0.50	0.45 ± 0.90	-0.20 ± 0.47	0.04 ± 0.47	-0.31 ± 0.55
Grade 6	(NA)	0.21 ± 0.37	-0.44 ± 0.52	0.22 ± 0.45	(NA)	(NA)
Grade 7	(NA)	-0.12 ± 0.51	0.15 ± 0.40	0.01 ± 0.52	(NA)	(NA)

Note: Schools <A>, <E>, and <F> do not have 6^{th} and 7^{th} grade.