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

Herbert J. Walberg[1]

Joseph Meyer[2]

April 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

Extensive statistical analyses were made of annual achievement progress on North Carolina's Department of Public Instruction tests. The data included 1,592 students in six Core Knowledge schools and 533,919 students in more than 1,300 other schools. The analyses showed that Core Knowledge schools excelled the other schools in achievement progress in eight of ten comparisons of reading and mathematics in the five grade levels available for analysis.

The present report adds the data for the other schools in North Carolina but employs the same sample of Core Knowledge schools and the same tests and demographic indicators as analyzed in the previous report. So that the present report can be read independently in the absence of the previous report, the information about the sample and tests are included in this report.

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 of testing served only primary grade students that did not participate in the testing for both

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

years. This left six schools, although two lacked sixth- and seventh-grade data since they served only earlier grades at the time the Department of Public Instruction compiled 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]

C. Choice of North Carolina State Tests

The North Carolina state tests seem 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,

---

[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.

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 increasingly under pressure to perform well on the state tests. 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, subjectively and usually incompletely "matched schools," neither of which is considered scientific.[5]

D. Statistical Procedures

Initial analysis of the test scores of 1,592 eligible Core Knowledge students' and 533,919 students in other schools showed that much of the variation in their scores, about 80 percent, is attributable to differences among students rather than differences among the Core Knowledge and other schools. As many studies have shown, achievement is a continuously accumulative process, and variations among schools in any given year account 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

[5] 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 (provided by the authors with the submission of the previous report).

life, for example, only about 8 percent of the total 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 family influences on intellectual development and achievement.[6]

Since 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, which means that only students exposed to the full year of the Core Knowledge or of other curricula were compared.

E. Clustered Data

The initial analysis 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 may tend to make them similar to one another and different from students in other schools. For example, a highly effective principal, means of instruction, or school board may confer higher test scores on students within a school, which sets them apart from students in other schools.

---

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

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, for the present evaluation is seemingly very large, the valid sample size is must be considered smaller to avoid 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[7] were employed for each combination of two subjects and five grades.

To be discussed further below, descriptive statistics about the sample in terms of frequencies and percentages are shown in Tables 1 and 2. Table 3 displays the contrast between Core Knowledge and other schools in achievement progress as well as the regression coefficients for pretest achievement, ethnicity, and poverty. Each of the tables deserves comment.

---

[7] 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).

III. Sample Sizes and Student Characteristics

A. Sample Sizes

Table 1 shows, for each grade level, the sample sizes of schools and students for Core Knowledge and other schools. The samples contain only students with complete achievement information for the two most recent academic years and with complete demographic information on minority status. The number of eligible Core Knowledge students in Grades 3 through 7 students is 1,592, and the number of students in other schools is 533,919.

Table 2 shows that the samples of non-Core Knowledge students for the five grades range from about 104,000 to 108,000. The samples of Core Knowledge students ranges from 379 to 468 in Grades 3 through 5. The samples range from 184 and 168 for Grades 6 and 7 because three Core Knowledge schools (A, E, and F) offered only earlier grades. Reflecting national trends for middle schools to be fewer and larger, there are fewer middle schools, somewhat less than 700, than primary schools, slightly more than 1,300, in North Carolina as indicated in Table 1.

B. Student Characteristics

As Table 2 shows, few American Indian, Asian, Hispanic, and Multi-Racial students are represented in either Core Knowledge or other schools. Greater percentages of Blacks attend Core Knowledge schools than other schools. Slightly smaller percentages of Whites attend Core Knowledge schools.

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.

Race/Ethnicity and Poverty Effects on Achievement

Many studies have shown racial/ethnic differences among students' achievement and the adverse effects of poverty. Consonant with previous research, Table 3 shows that poverty as indexed by free and reduced lunch status has effects on reading and mathematics achievement. Similarly, there are differences among students in the several racial/ethnic groups, which were taken into account in the analysis along with the previous year's achievement.[8]

IV. Achievement Progress in Core Knowledge and Other Schools

The last rows for reading and mathematics in Table 3 show the effect of Core Knowledge scores in contrast with other schools adjusted for the previous achievement, poverty, and racial/ethnic status. The results show that students in Core Knowledge schools outperformed those in other schools in eight of the ten instances (two subjects and five grades).[9] Although Core Knowledge schools outperformed in both subjects in Grades 4

---

[8] No doubt attributable to "colinearity," or correlation among the poverty and minority status indicators, the poverty and minority effects are inconsistent across subjects and grade levels. By definition, the designations within both the minority and poverty categories are inversely related; if a student, for example, is designated Hispanic they cannot also be Asian, and, similarly, if a student is eligible for reduced lunch, she or he cannot simultaneously be eligible for a free lunch. In addition, the two broad categories of poverty and minority status are collinear with one another since some racial/ethnic groups are more often eligible for free and reduced lunches. Such findings are neither new nor important for the analysis since these indicators and previous achievement scores are used merely to equate the groups for a fair comparison.

[9] Statistical probabilities of significance are reported here for the convenience of education researchers that attach importance to them, even though there is much debate

through 7, they underperformed in Grade 3 for some reason. Perhaps the third-grade Core Knowledge curriculum may be less well matched to the North Carolina tests than in the later grades in which the Core Knowledge students uniformly excelled. In any case, an 80 percent success rate is impressive.

V. Conclusion

This evaluation suggests an affirmative answer to the chief question: Do Core Knowledge Schools generally excel the academic progress of other schools adjusted for individual students' previous achievement scores and poverty and minority status. Unlike most previous curriculum evaluations, the comparison involves a huge sample, indeed, the entire population of eligible students in a state rather than a few hand picked Core Knowledge and comparison schools. The comparison also involves the North Carolina state

---

about them, and they are criticized by professional statisticians (see R. Allan Reese, "Does Significance Matter?" Significance: A Publication of the Royal Statistical Society, March 2004, pp. 39-40). Among other reasons, their assumption of a random sample from a well-defined population is rarely met.

Focusing on the general pattern of Core Knowledge schools exceeding other schools is particularly appropriate rather than the statistical probability of chance results. Since the sample actually comprises the population of more than a half-million eligible students in more than 1,300 schools, no ill-founded, mistaken inference of significance from sample to population is required. Moreover, though perhaps the largest Core Knowledge evaluation yet undertaken and one that enlarges the research to state tests and an entire eligible state population, the present research confirms previous evaluations showing Core Knowledge outperformed comparison schools. Thus, the results are internally consistent within the present study and consistent with the general conclusion of positive results in previous studies. See references at http://www.coreknowledge.org/CKproto2/about/eval.htm .

test, which can be assumed a priori to be neither fair nor unfair to Core Knowledge or comparison schools. In eight of ten comparisons, Core Knowledge schools excelled.

A previous report by the present authors[10] showed that the North Carolina Core Knowledge schools had employed uniformly and well the suggested Core Knowledge features and conditions such as participation in Core Knowledge Overview Workshop, Institutes, and the National Conference. Although the present report suggests that these same schools generally excelled other schools in achievement progress, the conclusion applies only to these North Carolina and similar Core Knowledge schools.

This report does not imply that schools that inadequately employ the Core Knowledge features and conditions can similarly excel. Nor does it imply that the present sample and other Core Knowledge schools cannot do better by more fully employing superior teaching techniques and school leadership that have generally raised achievement.[11]

Now that downloadable state databases are becoming more widely available, additional large-scale studies are feasible, which may reveal what makes for the biggest differences for Core Knowledge school success. Although the previous and the present research confirms that Core Knowledge schools excel, they may be capable of even better

---

[10] The Effects of Core Knowledge School Factors on State Test Achievement in North Carolina (Charlottesville, VA.: Core Knowledge, March 2004).

[11] 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 (provided with the submission of the previous report).

and more uniform results both through implementation of even more rigorous Core Knowledge features and conditions and the wider application of the best practices of teaching and school leadership.

Table 1. Distribution of students and schools

	Core Knowledge Schools Total: 1,592 students		Non-Core Knowledge Schools Total: 533,919 students	
Grade	# of Schools	# of Students/School: Median (Range)	# of Schools	# of Students/School: Median (Range)
3	6	78 (31-130)	1305	79 (1-266)
4	6	62 (30-98)	1302	77 (1-319)
5	6	56 (30-114)	1300	79 (1-381)
6	3	59 (49-76)	692	142 (1-710)
7	3	58 (38-72)	655	161 (1-609)

Table 2. Student characteristics: North Carolina state database, 3^{rd} – 7^{th} grade

	Core Knowledge Schools*						All CK	Other
	<A>	<B>	<C>	<D>	<E>	<F>	Schools	Schools
# 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.

Table 3. Core Knowledge participation as predictor for achievement gain (improvement in score from pre-test to post-test), adjusting for ethnicity and poverty status, by subject and grade

Effect	Reading Coefficents									
	3rd Grade		4th Grade		5th Grade		6th Grade		7th Grade	
	β	s.e.	β	s.e.	β	s.e.	β	s.e.	β	s.e.
Intercept	9.11	0.05	104.3	0.39	105.9	0.04	103.2	0.04	103.8	0.04
Am. Indian.	0.74**	0.23	0.23	0.24	0.10	0.18	0.35	0.19	0.55*	0.23
Asian	0.22	0.15	0.55**	0.14	-0.42**	0.12	0.02	0.11	0.62**	0.13
Black	-0.29**	0.06	0.16**	0.05	0.54**	0.05	-0.19**	0.05	0.37**	0.05
Hispanic	0.70**	0.10	0.66**	0.10	0.41**	0.08	0.15	0.09	0.90**	0.09
Multi-race	-0.23	0.14	0.08	0.12	0.23	0.13	-0.21	0.12	0.08	0.14
White	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)
Free Lunch	0.06	0.06	0.06	0.05	0.57**	0.05	-0.15**	0.05	0.28**	0.05
Reduced Pay	0.08	0.08	0.04	0.07	0.45**	0.06	-0.11	0.06	0.30**	0.07
Full Pay	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)
CK	-2.19*	0.63	1.59	0.72	0.05	0.85	0.70	0.11	0.13	0.22

Effect	Mathematics Coefficients									
	3rd Grade		4th Grade		5th Grade		6th Grade		7th Grade	
	β	s.e.	β	s.e.	β	s.e.	β	s.e.	β	s.e.
Intercept	15.6	0.05	7.36	0.05	6.11	0.05	5.29	0.08	4.26	0.07
Am. Indian.	0.81**	0.16	-0.12	0.17	0.05	0.17	0.22	0.16	-0.25	0.20
Asian	0.45**	0.11	1.37**	0.11	0.73**	0.11	0.86**	0.14	1.01**	0.14
Black	0.42**	0.04	0.49**	0.04	-0.09*	0.04	0.62**	0.05	-0.14**	0.05
Hispanic	0.84**	0.08	1.09**	0.07	0.43**	0.08	0.92**	0.09	0.40**	0.10
Multi-race	0.25*	0.10	0.11	0.10	0.02	0.11	0.40**	0.12	0.10	0.13
White	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)
Free Lunch	0.20**	0.04	-0.01	0.04	-0.32**	0.04	0.04	0.04	-0.41**	0.05
Reduced Pay	0.29**	0.06	0.04	0.05	-0.12*	0.05	0.23**	0.06	-0.24**	0.06
Full Pay	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)
CK	-2.43*	0.60	1.46	0.86	1.28	1.53	0.64	1.45	0.73	0.80

* $0.01 \le$ p-value < 0.05
** p-value < 0.01