

FOCUS ON RESULTS

AN ACADEMIC IMPACT ANALYSIS OF THE KNOWLEGE IS POWER PROGRAM (KIPP)



www.educationalpolicy.org

A research paper prepared by the Educational Policy Institute
for the KIPP Foundation

August 2005



Educational Policy Institute

Educational Policy Institute, Inc. (EPI) is a non-profit, non-partisan, and non-governmental organization dedicated to policy-based research on educational opportunity for all students. With offices in Virginia Beach (US), Toronto (CAN), and Melbourne (AUS), EPI is a collective association of researchers and policy analysts from around the world dedicated to the mission of enhancing our knowledge of critical barriers facing students and families throughout the educational pipeline.

The mission of EPI is to expand educational opportunity for low-income and other historically-underrepresented students through high-level research and analysis. By providing educational leaders and policymakers with the information required to make prudent programmatic and policy decisions, we believe that the doors of opportunity can be further opened for all students, resulting in an increase in the number of students prepared for, enrolled in, and completing postsecondary education.

For more information about the Educational Policy Institute, please visit our website:

www.educationalpolicy.org

FOCUS ON RESULTS:

An Academic Impact Analysis of the Knowledge Is Power Program (KIPP)

Educational Policy Institute
August 2005

Executive Summary

The Education Policy Institute (EPI) conducted an analysis of recent academic data collected from 24 KIPP schools on behalf of the KIPP Foundation. The purpose of this study was to evaluate these data to determine whether KIPP schools have had a positive impact on the student learning of 5th-grade students as demonstrated by standardized test results in comparison to national norms.

In this report, we examine the results on the Stanford Achievement Test for students who were 5th graders at KIPP schools in 2003-2004. These students comprise 27 cohorts across two different testing periods: (a) fall 2003 to spring 2004; and (b) fall 2003 to fall 2004.

Findings from this study show that 5th-grade cohorts at KIPP schools post substantially greater academic gains on the Stanford Achievement Test (SAT) than what is considered normal, a finding consistent with prior research on KIPP schools (Doran and Drury, 2002). Given that a growth score of zero on the normal curve equivalent (NCE) is considered “normal growth,” schools administering the tests first in the fall with a follow-up test in the spring enjoyed a mean gain of 10.1 in reading, 10.9 in language, and 17.4 in mathematics. Schools that first administered the test in the fall and then again the following fall recognized score gains of 7.5 in reading, 9.1 in language, and 11.6 in mathematics.

As the data indicate, KIPP 5th-grade cohorts experienced average gains of 9 to 17 points across all tests (see tables in Appendix B for school by school results).

Although not all KIPP schools can boast gains that move students one standard deviation or even one-half of a standard deviation, only one cohort in each testing period [fall-to-spring and fall-to-fall] lost ground. The fact that most schools did show increases well above normal growth rates in reading, language, and mathematics is laudable and worthy of continued investigation and practice.

These findings, while impressive, do not allow us to suggest that KIPP schools have necessarily found “the answer” to the educational woes of urban schools. However, the data suggest that these

Focus on Results: An Analysis of the KIPP Schools

schools are doing something right. Only after continued research will we be able to find out whether the practices utilized by KIPP schools are replicable and able to aid urban areas in the education of youth.

The report concludes with a recommendation for continued research in the outcomes and pedagogies utilized in KIPP schools as compared to other urban districts. Specifically, EPI recommends that future research incorporate both matched-student cohorts and empirically-driven comparison groups.

FOCUS ON RESULTS:

An Academic Impact Analysis of the Knowledge Is Power Program (KIPP)

Educational Policy Institute
August 2005

Middle school is a critical point in the maturation of youth, where decisions that impact future education and career opportunities begin to take shape. The college choice process introduced by Hossler and Gallagher (1989) suggests that the predisposition for postsecondary study has to be in place before any of the necessary steps can be taken. Cabrera and LaNasa (2000) suggest that this phase can begin and definitely ends in middle school. Strangely, little research has been done on middle schools where this issue is concerned.

KIPP began in 1994 when Dave Levin and Mike Feinberg launched a program for 5th-graders in a public school in inner-city Houston, Texas. In 1995, Feinberg's KIPP Academy Houston became a charter school, and Levin returned home to New York to establish KIPP Academy in the South Bronx.

The original KIPP Academies have a sustained record of high student achievement. The Texas Education Agency recognized KIPP Academy Houston as a Texas "Exemplary School" for its first eight years of existence, and the New York State Senate has recognized KIPP Academy New York as the highest-performing public middle school in the Bronx for eight consecutive years.

In 2000, Doris and Donald Fisher, co-founders of Gap Inc., formed a partnership with Feinberg and Levin to replicate KIPP's success. The non-profit KIPP Foundation recruits, trains, and supports outstanding teachers to open college preparatory public schools in high-need communities nationwide. Principals trained through the KIPP School Leadership Program have replicated the initial success of the original KIPP Academies. According to a 2002 report, the first three new KIPP schools each recorded statistically-significant gains in student achievement (Doran and Drury, 2002).

The mission of KIPP schools is to provide underserved students (predominantly in urban settings) with the knowledge, skills, and character needed to succeed in college and the competitive world beyond. To achieve this goal they are guided by five pillars: high expectations, choice and commitment, more time (e.g., longer school days), power to lead, and focus on results. Of these five principles, the focus on results has become a pressing concern for KIPP. The focus on results concerns all of the outcomes attributable to KIPP, like discipline, scholarship, leadership, academic performance

and high aspirations. Although performance on standardized tests is not the sole measure of KIPP success, it is an objective measure that can be used as part of a larger set of indicators to evaluate and possibly validate the efforts of KIPP school leaders, teachers, and students.

The KIPP Foundation contracted with the Educational Policy Institute to conduct an analysis of recent academic data collected from their schools. The purpose of this study was to evaluate these data to determine whether KIPP schools have had a positive impact on student learning and proficiency as demonstrated by standardized test results in comparison with national norms. This is considered a preliminary study to be followed by a more rigorous, student-level data analysis.

KIPP in Context

Discussions of student performance are often couched within the socioeconomic and racial/ethnic identities of student groups. Understanding who KIPP students are and how they compare to other schools nationally is a critical component of an academic performance study.

Table 1. Comparative Demographics: Percentage of Students by Race/Ethnicity

Group	White (non-Hispanic)	Black/African American	Hispanic/ Latino	Asian/ Pacific Islanders	American Indian/ Alaskan Native
KIPP	1.7	62.5	32.7	2.0	0.1
Urban ¹	23.1	37.1	32.8	6.4	0.6
Nation (2001-02) ²	58.9	16.9	18.5	4.4	1.3

Whereas White, non-Hispanic students make up 58.9 percent of schools nationwide, and 23.1 percent of urban schools, only 1.7 percent of KIPP students are non-Hispanic students. Conversely, KIPP serves a population that is 95.2 percent African American and Hispanic, while urban schools and the public schools nationally serve African American and of 69.9 percent and 25.4 respectively.

The widely-discussed achievement gap between White, non-Hispanic students and students of color, particularly Black students, has persisted. There is evidence that the gap is closing, but national data show that there is still some distance to go. Results from the 2004 NAEP long-term trend assessment, also known as the Nation's Report Card, show that the gap in performance between 4th- and 8th-grade White, non-Hispanic students and 4th- and 8th-grade African American and Hispanic students, is not significantly smaller now than it was in 1986 (in mathematics) and 1988 (in reading).

¹ Urban school statistics are taken from the Council of the Great City Schools, an association of 65 urban schools districts representing 7.3 million students. (<http://www.cgcs.org/about/about.html>)

² NCES (2002).

African American and Hispanic students continue to score well below average on this national assessment (NCES, 2003a).

While the term “urban schools” has become synonymous with “non-White schools,” such schools also serve a higher percentage of low-income students. Like other urban schools, KIPP schools serve a disproportionately low-income student population. For instance, 78 percent of KIPP students receive free- and reduced-price lunch, compared to 63 percent of students at urban schools and 40 percent at all schools nationwide.

Table 2. Comparative Demographics: Percentage of Students with Special Needs

Group	English Language Learners	Special Education	Free and Reduced-Priced Lunch
KIPP	10.3	8.0	78.0
Urban	17.0	12.9	62.6
Nation (2001-02)	5.0 ³	6.0 ⁴	39.7 ⁵

In addition, KIPP schools serve twice the percentage of English language learners as schools nationwide (10.3 vs. 5 percent). Other urban schools serve an English language learner population of 17 percent. With respect to students with disabilities, KIPP schools serve a special education population of 8 percent, compared to 6 percent nationally and 12.9 percent for urban districts.

The NAEP Urban Report provides free- and reduced-price lunch data for the 2003 administration of the 4th-grade reading and mathematics assessment (NCES, 2004a; NCES, 2004b). Nationally, public school students who were eligible for free- or reduced-price lunch scored, on average, 222 on the mathematics assessment (compared to 234 for all students) and 201 on the reading assessment (compared to 216 for all students). Comparatively, urban students who were ineligible for free or reduced-price lunch—more affluent students—scored 244 on the mathematics assessment and 229 on the reading assessment.

The presentation of NAEP findings speak of race/ethnicity and socioeconomic status in a manner that presupposes no interaction between the two. This is, of course, not the case. At the heart of the achievement gap discussion is the union of race/ethnicity and socioeconomic status and its effect on achievement. The analysis that follows does not attempt to isolate these characteristics.

³ (NCES, 2003b)

⁴ (NCES, 2003a); 2000-01 data, Table 52

⁵ The free and reduced-priced lunch eligibility data do not include data from AZ, CT, TN, and WY because those states did not report.

The above data are intended to describe how students on a national scale, who share the characteristics of KIPP students, perform in relation to White, non-Hispanic students and students from the middle and upper classes. The analysis that follows illustrates how KIPP students depart from these national averages.

Methodology

This report details an analysis of a single cohort across 24 KIPP schools, representing approximately 1,800 5th-grade students. Achievement data from three administrations of the SAT 9/10 are analyzed.⁶ One group took the test in the fall of 2003 and the spring of 2004. Another group took the test in the fall of 2003 and the fall of 2004 when they were in the 6th grade. Some students were tested all three times. No effort was made to match the cohorts, but the change in numbers tested was minimal, and we found no demographic shift significant enough to raise concerns about the consistency of the sample from one test administration to the next. Still, we cannot guarantee that the cohorts are identical.

Table 3. Number of Students Tested by Administration and Subject

Number Tested	Fall 2003 (base)	Spring 2004	Fall 2004
Reading	1819	1167	636
Language	1336	891	580
Mathematics	1865	1243	640

All achievement data for all KIPP schools were provided to the authors for analysis in cohort form. EPI was not privy to individual student data. The sample was based on a number of criteria that can be found in a more detailed methodology section in Appendix A. Eighteen schools are in the first group of fall-to-spring administration; 9 schools are in the second group of fall-to-fall administration. Three schools tested in fall of 2003, spring of 2004, and fall of 2004, yielding 24 schools in total administering the Stanford Achievement Test. Our analysis breaks down analysis into fall-to-spring and fall-to-fall testing groups to ensure that there is no double counting and to simplify the discussion.

No screening tests are required for enrollment in a KIPP school. This leaves no reason to believe that students attending KIPP are higher-performing than others in their district when they enter. This analysis does not examine true pre-test data—data from before students started with KIPP—so it is not possible to prove that “creaming” (i.e., taking the best students from other schools) is or is not

⁶ The Stanford Achievement Test 9th and 10th Editions were administered. KIPP provided scores to EPI after transforming all scores to 2002 norms, making scores comparable across editions.

occurring. However, a previous student-level analysis conducted by the New American Schools does consider scores from the 3rd and 4th grades with findings very similar to those here. Fall test results are the only data available to us that can reasonably approximate a pre-test situation and typically fall significantly below the mean (see Appendix B).

It is also impossible, considering this data set, to make comments about the comparability of the findings here to those in a random assignment design. Families do choose KIPP over other institutions. This type of selection, although not indicative of pre-KIPP achievement, does indicate possible special characteristics of KIPP parents that are not captured by this analysis. Further research would need to be done to determine the characteristics of KIPP parents compared to non-KIPP parents in the same districts and what effects those characteristics have.

Although this study design does not allow for claims of causality, we feel confident that the following findings are reflective of what is happening in KIPP schools. Future studies will need to be conducted before we can fully attribute outcomes to the schools.

Findings

The primary findings in this analysis regard academic achievement growth, as measured by the change in test scores from one test administration to the next (e.g., fall-to-spring; fall-to-fall). All results were significant at the .01 level for all subject areas, and test scores were provided as normal curve equivalent (NCEs) scores on the Stanford Achievement Test. Designed to make manipulation (aggregation and averaging) easier, the NCE has a mean of 50, a standard deviation of 21.06, a range from 1-99, and can be manipulated arithmetically. Normal growth in NCEs is zero. That is to say, if a student scores a 50 at the first administration and a 50 at the second administration, he or she demonstrated normal academic growth.

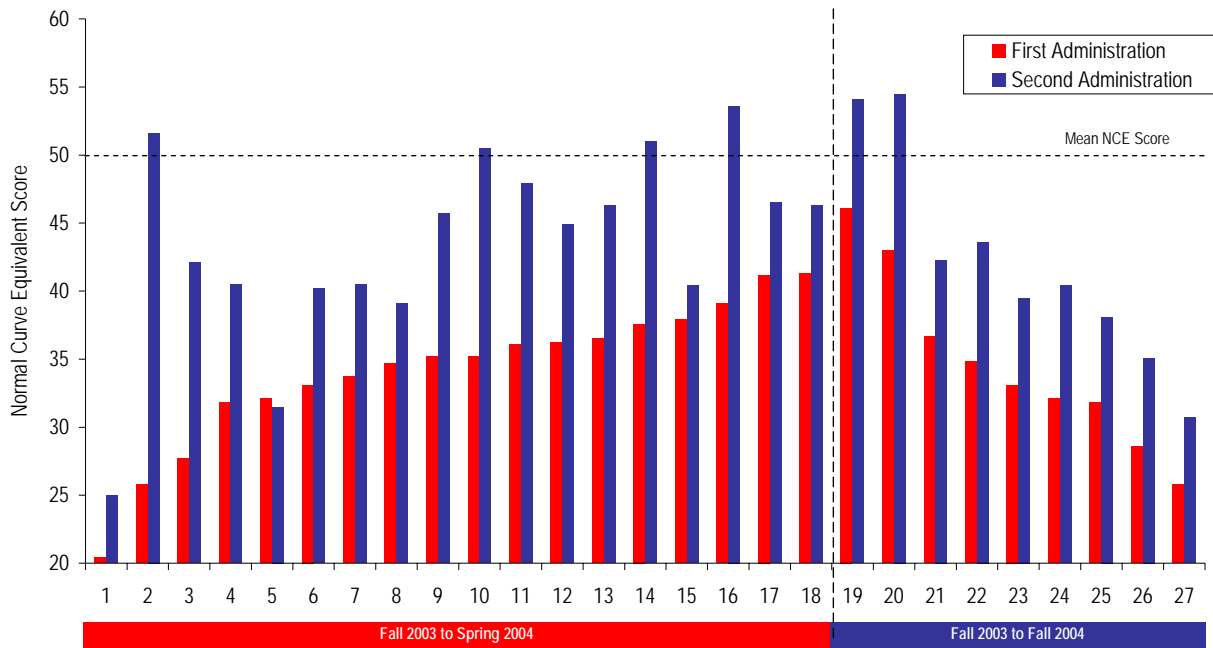
Based on our analysis, KIPP schools post substantially greater gains than what is considered normal. This finding is consistent with the results of a previous study which also describes substantial and significant gains compared to other schools in the same districts studied (Doran and Drury, 2002). The present analysis found that KIPP students scored well below the mean normal curve equivalent score (NCE) on their initial testing in reading, language, and mathematics. However, when tested in

spring 2004 or fall 2004⁷, students at these schools posted significant increases in their ability. The following provides a brief synopsis by test area.

READING

Figure 1 illustrates the reading outcomes of 27 KIPP 5th-grade cohorts in 24 schools. The figure illustrates 27 cohorts, but it should be noted that 3 schools administered the SAT 9/10 in both spring and fall. Each cohort examined scored well below the mean of 50 in fall 2003. For the fall-to-spring testing group, cohorts moved from well below the mean to above the mean, and all but one cohort made increases. This testing group experienced a mean gain of 10.1 points (Table 4; Figure 4) with a maximum gain of 25.8 by 1 cohort. For the fall-to-fall testing group, 2-of-9 cohorts moved above the mean NCE and all made substantial progress. The mean reading gain for the fall-to-fall testing group was 7.4 points with a maximum of 11.5 points (Table 4; Figure 5).

Figure 1. KIPP Reading Normal Curve Equivalent Scores for the 5th Grade on the Stanford Achievement Test by School, 2003-04



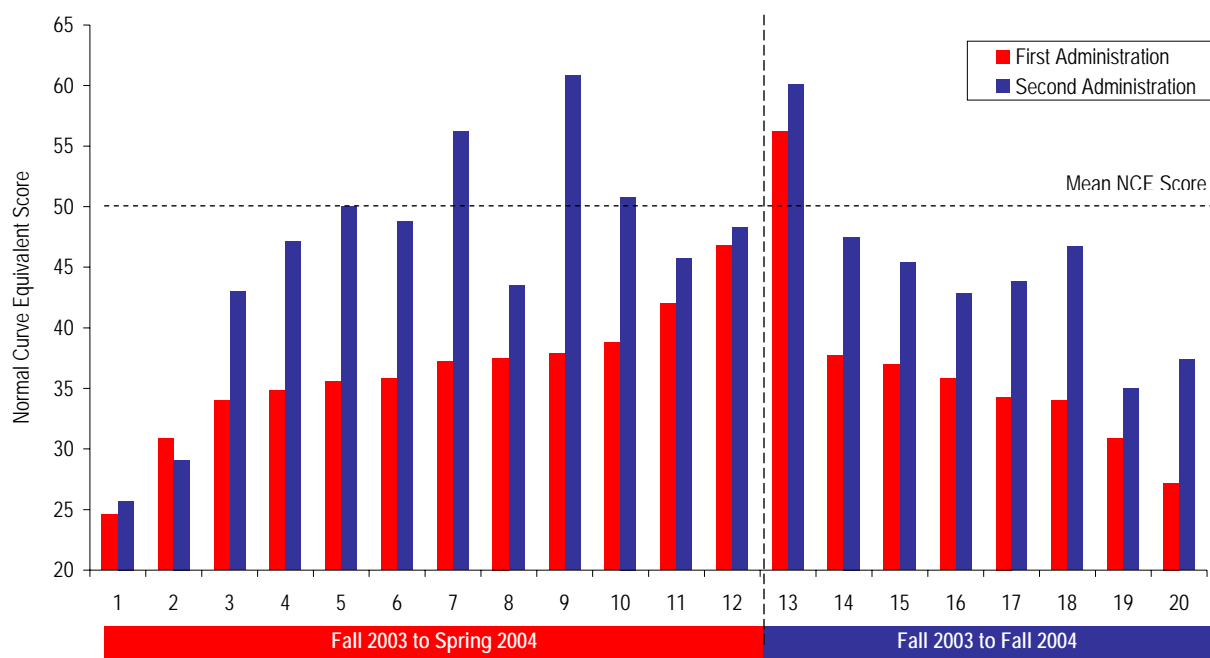
NOTE: Three KIPP schools administered SAT 9/10 tests in both spring and fall 2004, thus summing 27 indicators from 24 schools.

⁷ Some schools did not test in the spring of 2004 but tested in the fall of 2004 when 5th-grade students were in the 6th grade. Three schools tested in both spring and fall of 2004 and these schools are included in both categories.

LANGUAGE

Figure 2 illustrates the language outcomes of 20 KIPP 5th-grade cohorts. Again, three schools administered all tests as with reading. All but one cohort scored well below the mean of 50 in fall 2003. For the fall-to-spring testing group, 4-of-12 5th-grade cohorts moved from well below the mean to at or above the mean in 2004, and all but 1 cohort made increases. This testing group experienced a mean gain of 10.9 points (Table 4; Figure 4) with a maximum gain of 22.9 points. For the fall-to-fall testing group, 1-of-8 cohorts moved above the mean NCE and all made substantial progress. The mean language gain for the fall-to-fall testing group was 9.1 points with a maximum of 10.2 points (Table 4; Figure 5).

Figure 2. KIPP Language Normal Curve Equivalent Scores for the 5th Grade on the Stanford Achievement Test by School, 2003-04



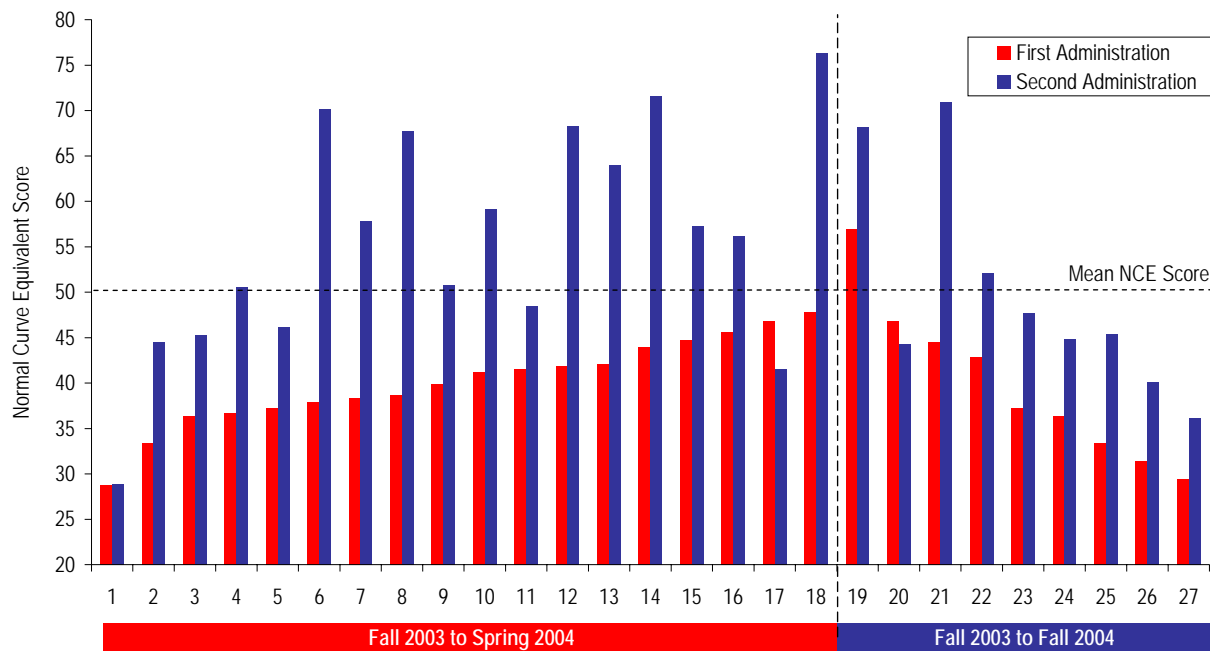
NOTE: Three KIPP schools administered SAT 9/10 tests in both spring and fall 2004, thus summing 20 indicators from 17 schools.

MATHEMATICS

Figure 3 illustrates the mathematics outcomes of 27 KIPP 5th-grade cohorts in 24 schools. Again, readers should understand that schools administered both follow-up spring and fall 2004 tests; thus, the illustration shows 27, not 24, cohorts. The mathematics progress of KIPP 5th-grade cohorts was by far the most impressive among the three test areas. All but 1 cohort scored well below the mean of 50 in fall 2003. For the fall-to-spring testing group, 12 of 18 cohorts—two-thirds—moved

from well below the mean to at or above the mean in 2004. All but one cohort posted an increase. This testing group experienced a mean gain of 17.4 points with a maximum gain of 32.3 points (Table 4; Figure 4). For the fall-to-fall testing group, 3-of-9 cohorts moved above the mean NCE and all but 1 made substantial progress. The mean mathematics gain for the fall-to-fall testing group was 11.6 points with a maximum of 26.4 (Table 4; Figure 5).

Figure 3. KIPP Mathematics Normal Curve Equivalent Scores for the 5th Grade on the Stanford Achievement Test by School, 2003-04



NOTE: Three KIPP schools administered SAT 9/10 tests in both spring and fall 2004, thus summing 27 indicators from 24 schools.

Aside from exceeding the normal growth threshold, KIPP 5th-grade cohorts have increased, on average, almost one-half a standard deviation with several cohorts moving one standard deviation above the mean (see tables in Appendix B for school by school results).

Table 4. Growth Scores by Test Administration

Administration	Subject	N	Max	Min	Mean
Fall-Spring	Reading	18	25.8	-0.6	10.1
	Language	12	22.9	-1.8	10.9
	Mathematics	18	32.3	-5.3	17.4
Fall-Fall	Reading	9	11.5	4.9	7.4
	Language	8	10.2	3.9	9.1
	Mathematics	9	26.4	-2.5	11.6

Figures 4 and 5 simply illustrate the gain scores of KIPP cohorts provided in Table 4. The red area in each figure represents below-normal growth, which in each testing group (spring and fall) and test (reading, language, and mathematics) was realized by only one 5th-grade cohort. The grey area signifies growth above normal up to the mean of the distribution. Finally, the black area represents growth above the mean to the maximum found in this study by any 5th-grade cohort in each testing group and in each test area.

Figure 4. Growth Scores for the 5th Grade, Fall 2003 to Spring 2004

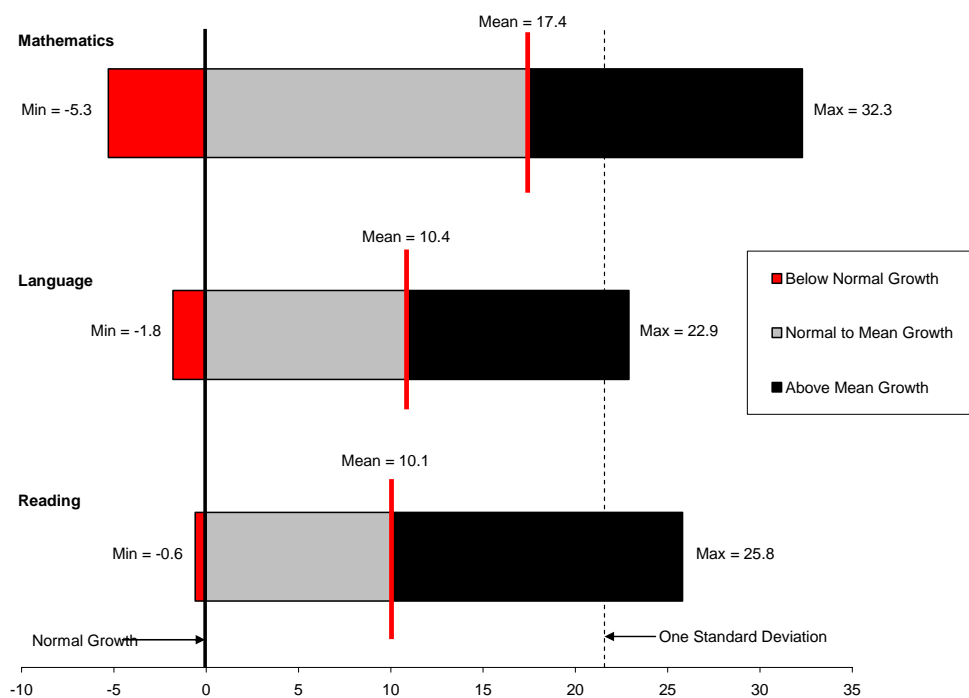
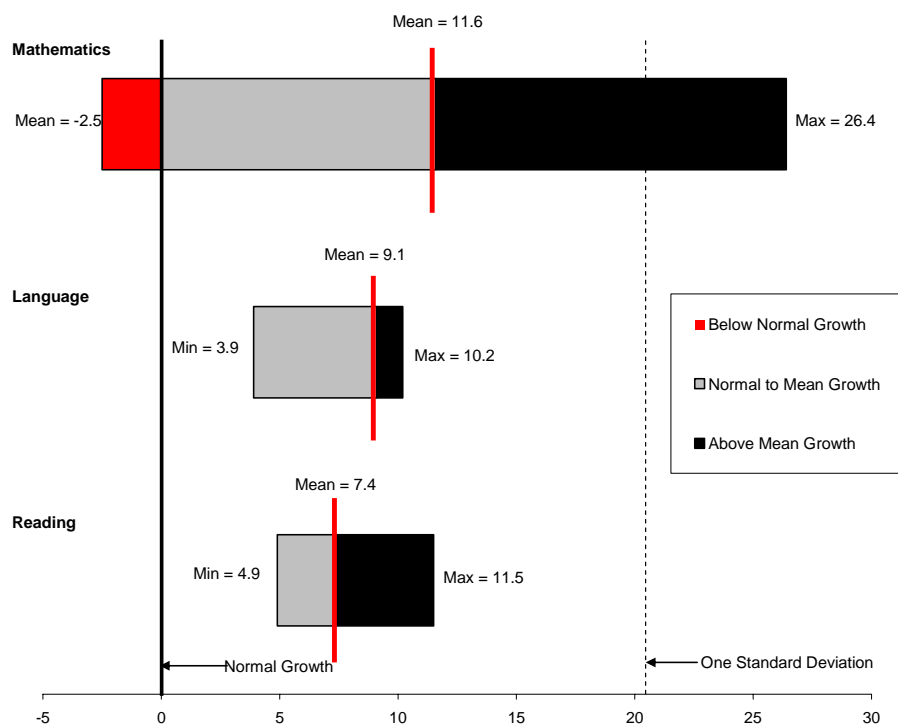


Figure 5. Growth Scores for the 5th Grade, Fall 2003 to Fall 2004



Discussion

The Knowledge Is Power Program (KIPP) has posted large and significant gains on a nationally norm-referenced standardized test. This performance is true across schools and throughout the nation. The fact that KIPP 5th-grade cohorts showed a dramatic increase well above normal growth rates in reading, language, and mathematics is laudable and worthy of continued investigation and practice.

These findings, while impressive, do not allow us to suggest that KIPP schools have necessarily found “the answer” to the educational woes of urban schools. However, the data suggest that these schools are doing something right. Only after continued research will we be able to find out whether the practices utilized by KIPP schools are replicable to aid urban areas in the education of youth.

Thus, we conclude with a proposal of continued research in the following areas:

Multi-Year Student-Level Analysis. This analysis is weakened by its small sample size. Looking closely at individual students over a longer period of time would allow researchers to discuss how many students enjoy the gains discussed here and how students perform over time. Are the majority of gains enjoyed in the 5th grade? At which point do the majority of students begin to perform at or above

grade level in reading, an area of consistent under-performance? What happens with students who do not enter at the 5th grade?

Comparison/Control Groups. KIPP schools are located in urban areas. The students they serve come from established schools. Being able to analyze data from schools serving students most like KIPP students would greatly enrich the findings. How do the other students perform in the 5th grade on similar assessments? What kinds of gains are made over time in the other schools? Ultimately, we need quality comparison groups to substantiate any claims about the program.

Qualitative/Ethnographic. If the data continue to point to KIPP as a consistently successful model, people will want to know what happens inside KIPP schools. In addition to the two types of quantitative analysis described above, qualitative studies inside KIPP schools and comparison schools will be necessary to demonstrate what KIPP does that is so different from other middle schools. Linking this information to the quantitative data will provide a rich look at the progression of KIPP students and the pedagogy that drives this progression.

These are only three suggestions with potential for research in a great many other areas. The findings presented in this report illustrate that students who are generally expected to perform poorly by the larger society according to historical performance in some cases, outperformed the greatest of expectations. The purpose of the analysis was to provide some answers, but in reality, we are left with more questions. The outlook is very positive for KIPP schools and the students and families they serve, and through further research we will be able to provide more definitive statements of outcomes for these students and schools.

References

- ACSFA. (2002). *Empty Promises: The Myth of College Access in America*. Washington, DC: Advisory Committee on Student Financial Assistance.
- Cabrera, A. F., & LaNasa, S. M. (Eds.). (2000). *Understanding the college choice of disadvantaged students*. San Francisco: Jossey-Bass.
- CGCS. (2005). *Beating the Odds V: A City-by-City Analysis of Student Performance and Achievement Gaps on State Assessments*. Washington, DC: Council of the Great City Schools.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of Educational Opportunity*. Washington, D.C.: U.S. Department of Health, Education and Welfare.
- Herrnstein, R., & Murray, C. (1994). *The Bell Curve*. Intelligence and Class Structure in American Life. Northampton, MA: The Free Press.
- Hossler, D., Braxton, J., & Coopersmith, G. (1989). Understanding student college choice. In J. C. Smart (Ed.), *Higher Education: Handbook of Theory and Research* (Vol. V, pp. 231-288). New York: Agathon Press.
- Hossler, D., & Gallagher, K. S. (1987). Studying student college choice: A three-phase model and the implications for policymakers. *College and University*, 62(3), 207-221.
- Jencks, C. (1972). *Inequality; a reassessment of the effect of family and schooling in America*. New York,: Basic Books.
- Jencks, C., & Phillips, M. (1998). *The Black-White Test Score Gap*. Washington, DC: Brookings Institution Press.
- Ogbu, J. U. (1998). Voluntary and involuntary minorities: a cultural-ecological theory of school performance with some implications for education. *Anthropology & Education Quarterly*, 29(2), 155-188.
- Snipes, J., Doolittle, F., & Herlihy, C. (2002). *Foundations for Success: Case Studies of How Urban School Systems Improve Student Achievement*. Washington, DC: Council of the Great City Schools and MDRC.
- Steele, C., & Aronson, J. (1995). Stereotype Threat and the Intellectual Test Performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797.
- NCES (2002). Common Core of Data, "Public/Elementary School Universe Survey," 2001-02, Version 1a, and "Local Education Agency Universe Survey," 2001-02, Version 1a.
- NCES. (2003a). *Digest of Education Statistics, 2002* (No. NCES 2003-060). Washington, DC: National Center for Education Statistics.
- NCES. (2003b). *Overview of Public Elementary and Secondary Schools and Districts: School Year 2001-2002* (No. NCES 2003-411). Washington, DC: National Center for Education Statistics.
- NCES. (2004a). *The Nation's Report Card: Trial Urban District Assessment : Mathematics Highlights 2003* (No. NCES 2004-458). Washington, DC: National Center for Education Statistics.
- NCES. (2004b). *The Nation's Report Card: Trial Urban District Assessment : Reading Highlights 2003* (No. NCES 2004-459). Washington, DC: National Center for Education Statistics.
- Valencia, R. R. (Ed.). (1997). *The Evolution of Deficit Thinking*. Washington, D.C.: The Falmer Press.

Appendix A

Methodology

Sampling

The non-profit KIPP Foundation encourages all KIPP schools to administer a nationally-recognized norm-referenced exam once a year, in order to track academic growth over time. During the 2003-2004 academic year, all 31 KIPP schools in operation took at least one norm-referenced exam.⁸ The Educational Policy Institute received test data for these schools from the KIPP Foundation. Of these schools, EPI used data from only those schools (24) that administered the SAT 9/10 to the 5th-grade cohort in the fall of 2003 *and* also administered the SAT 9/10 to the same cohort in either the spring of 2004 or the fall of 2004 (when in 6th grade). The five schools that used other tests (ITBS and TerraNova) were not included. One final caveat is that EPI received SAT 10 data in February, before KIPP Sunshine Peak Academy (Denver, CO) delivered the data to KIPP Foundation. However, KIPP Sunshine Peak's results were consistent with the gains seen by other KIPP schools: 4.2 NCEs in reading, 8.8 NCEs in math, and 9.0 NCEs in language.

The cohort with the largest number of scores for the SAT 9/10 was the 2011 cohort. This is the group of students who were 5th graders in the fall of 2003 and will graduate in 2011. There was insufficient data in all other cohorts to conduct even the most minimal analysis. Schools without data for this cohort were excluded.

The majority of schools administer the test in the fall and again in the spring. Some administer tests only in the fall. Others administer the test every semester. Schools were divided into fall-spring and fall-fall groups. Schools with missing data were eliminated.

- Reading
 - Fall-Spring (18 schools)
 - Fall-Fall (9 schools)
- Language
 - Fall-Spring (12 schools)
 - Fall-Fall (8 schools)
- Mathematics
 - Fall-Spring (18 schools)
 - Fall-Fall (9 schools)

It should be noted that data were missing not because KIPP failed to share them with EPI, but because the school did not report them to the KIPP Foundation. EPI did not attempt to gather data from individual schools. Missing data, according to the above definition and within the listed parameters, is the reason for the final, substantially smaller, number of schools used in the analysis.

Statistical Method

The analysis of the data was conducted using the Wilcoxon Signed Ranks Test. The choice of this test evolved from a standard t-test to the Wilcoxon.

⁸ There were 38 schools in 2004-05.

The data require a comparison of paired results. The t-test was insufficient because the sample size was too small and the challenge of consistently meeting the normality assumption for all tests (e.g., reading and math) was too great.

The Wilcoxon Signed Ranks Test concentrates on the differences between the two samples and is a nonparametric test requiring no adherence to rules of normality. The W is reported below because of the small sample size.

Table 5. Wilcoxon Test Statistics

	Subject	N	W	z
Fall 2003- Spring 2004	Reading	18	170	3.693
	Language	12	75	2.922
	Mathematics	18	169	3.671
Fall 2003-Fall 2004	Reading	9	45	--
	Language	8	36	--
	Mathematics	9	44	--

For $N > 10$, the test statistics indicate that changes in all scores are significant at the .01 level. For $N < 10$, the tests also indicate significance at the .01 level, but at this size N the W must be compared to an exact sampling distribution. No z is calculated.

Limitations

Several limitations to this analysis require the consideration of readers. First, this is not a randomized study, nor are there control groups. Any comparisons that can be made are restricted to general performance on nationally-standardized tests.

Second, only school-level data were used in this study, not the much preferred student-level data. This was done because student-level data were simply not available and the school aggregate was the only available option for analysis. Although the sample size is small due to this approach, it does represent the majority of schools in KIPP. If student-level data were available, the sample size would exceed 1,800.

Third, the cohorts in this analysis are unmatched. That is, we cannot be sure that the students measured in the first SAT 9/10 at a particular KIPP school were exactly the same students in the subsequent spring or fall cohorts. We assume they are very close to being the same, but because they are unmatched we are unable to make that statement. We considered several ways of checking for consistency across test administrations. Without student-level data and using only one year of demographic data, we settled on reporting the numbers of students tested from one administration of the test to the next (Table 3). There was no major demographic shift in the schools along racial/ethnic and/or socioeconomic lines.

Assumptions were made that test administration dates were sufficiently close to group schools by fall-fall and fall-spring, without compromising the comparability.

Appendix B

Data Tables

Table 6. List of NCE Scores and Number Tested--Reading

Fall 03	N	Spring 04	N	Growth
41.3	87	46.3	84	5.0
41.2	71	46.5	74	5.3
39.1	65	53.6	65	14.5
37.9	67	40.4	66	2.5
37.6	87	51.0	76	13.4
36.5	57	46.3	52	9.8
36.2	85	44.9	82	8.7
36.1	78	47.9	82	11.8
35.2	91	45.7	78	10.5
35.2	86	50.5	85	15.3
34.7	86	39.1	77	4.4
33.7	85	40.5	86	6.8
33.1	70	40.2	66	7.1
32.1	38	31.5	30	-0.6
31.8	68	40.5	59	8.7
27.7	74	42.1	76	14.4
25.8	83	51.6	84	25.8
20.4	29	25.0	23	4.6
Fall 03	N	Fall 04	N	Growth
46.1	86	54.1	87	8.0
43.0	64	54.5	65	11.5
36.7	85	42.3	73	5.7
34.8	51	43.6	54	8.8
33.1	70	39.5	69	6.4
32.1	38	40.4	23	8.3
31.8	68	38.1	56	6.3
28.6	83	35.1	78	6.5
25.8	73	30.7	75	4.9

Table 7. List of NCE Scores and Number Tested--Language

Fall 03	N	Spring 04	N	Growth
46.8	87	48.3	84	1.5
42.0	84	45.7	77	3.7
38.8	85	50.8	81	12.0
37.9	86	60.8	85	22.9
37.5	83	43.5	85	6.0
37.2	84	56.2	84	19.0
35.8	70	48.8	66	13.0
35.6	80	50.0	77	14.4
34.8	80	47.1	76	12.3
34.0	68	43.0	59	9.0
30.9	39	29.1	28	-1.8
24.6	29	25.7	23	1.1
Fall 03	N	Fall 04	N	Growth
56.2	86	60.1	87	3.9
37.7	52	47.5	54	9.8
37.0	85	45.4	72	8.4
35.8	70	42.8	70	7.0
34.3	80	43.8	78	9.5
34.0	68	46.7	56	12.7
30.9	39	35.0	23	4.1
27.2	86	37.4	58	10.2

Table 8. List of NCE Scores and Number Tested--Mathematics

Fall 03	N	Spring 04	N	Growth
47.8	70	76.3	66	28.5
46.8	39	41.5	26	-5.3
45.6	86	56.2	84	10.6
44.7	74	57.3	74	12.6
44.0	86	71.6	85	27.6
42.1	86	64.0	75	21.9
41.9	82	68.3	82	26.4
41.5	68	48.5	66	7.0
41.2	85	59.1	82	17.9
39.9	85	50.8	78	10.9
38.7	96	67.7	78	29.0
38.3	58	57.8	52	19.5
37.9	88	70.1	84	32.2
37.3	89	46.2	86	8.9
36.7	77	50.6	76	13.9
36.4	70	45.3	66	8.9
33.4	68	44.5	59	11.1
28.8	30	28.9	24	0.1
Fall 03	N	Fall 04	N	Growth
56.9	86	68.2	87	11.3
46.8	39	44.3	23	-2.5
44.5	63	70.9	65	26.4
42.9	52	52.1	54	9.2
37.2	86	47.7	73	10.5
36.4	70	44.8	70	8.4
33.4	68	45.4	56	12.0
31.4	85	40.1	53	8.7
29.4	84	36.1	78	6.7



Educational Policy Institute

EPI International

2345 Valle Rio Way
Virginia Beach, VA 23456
(540) 288-2322
info@educationalpolicy.org

EPI United States

2345 Valle Rio Way
Virginia Beach, VA 23456
(540) 288-2322

EPI Canada

20 Maud Street, Suite 300
Toronto, ON M5V 2M5
Canada
(416) 848-0237

EPI Australasia

174 Wingrove St, Fairfield, 3078
Melbourne, Australia
61 3 9486 1334

www.educationalpolicy.org