

Exploring Patterns in School Achievement from KCCT to K-PREP: The Role of School-Level Socioeconomic Status

Prepared for: Kentucky Department of Education
Capital Plaza Tower, 18th Floor
500 Mero Street
Frankfort, KY 40601

Authors: Emily R. Dickinson
Heather Levinson
Arthur A. Thacker

Date: June, 2013



Human Resources Research Organization (HumRRO)
66 Canal Center Plaza, Suite 700, Alexandria, Virginia 22314-1578
Phone: 703.549.3611 | Fax: 703.549.9661 | www.humrro.org

Exploring Patterns in School Achievement from KCCT to K-PREP: The Role of School-Level Socioeconomic Status

Prepared for: Kentucky Department of Education
Capital Plaza Tower, 18th Floor
500 Mero Street
Frankfort, KY 40601

Authors: Emily R. Dickinson
Heather Levinson
Arthur A. Thacker

Date: June, 2013



Exploring Patterns in School Achievement from KCCT to K-PREP: The Role of School-Level Socioeconomic Status

Table of Contents

Background	2
Data Description	3
Analyses and Results	1
Correlational Analysis	1
Effect size Analysis	3
Achievement Gaps	9
Discussion and Conclusions	10
References	12

List of Tables

Table 1. School Grand Mean Scale Scores and Percentage of Students Receiving Free or Reduced Lunch	4
Table 2. Correlations between percentage of students receiving free/reduced lunch and school changes in percent proficient by grade level.....	1
Table 3. Correlations between percentage of students receiving free/reduced lunch and school standardized mean scale score changes by grade level.....	3
Table 4. Grade 3 average score differences by school SES quartiles.....	4
Table 5. Grade 3 average score differences: Effect sizes between the lowest and highest school SES quartiles.....	4
Table 6. Grade 4 average score changes by school SES quartiles.....	5
Table 7. Grade 4 average score changes: Effect sizes between the lowest and highest school SES quartiles.....	5
Table 8. Grade 5 average score changes by school SES quartiles.....	6
Table 9. Grade 5 average score changes: Effect sizes between the lowest and highest school SES quartiles.....	6
Table 10. Grade 6 average score changes by school SES quartiles.....	7
Table 11. Grade 6 average score changes: Effect sizes between the lowest and highest school SES quartiles.....	7
Table 12. Grade 7 average score changes by school SES quartiles.....	8
Table 13. Grade 7 average score changes: Effect sizes between the lowest and highest school SES quartiles.....	8
Table 14. Grade 8 average score changes by school SES quartiles.....	9
Table 15. Grade 8 average score changes: Effect sizes between the lowest and highest school SES quartiles.....	9
Table 16. Effect size differences in mean scale scores for free/reduced lunch and non-free/reduced lunch students.....	10

Exploring Patterns in School Achievement from KCCT to K-PREP: The Role of School-Level Socioeconomic Status

Background

As Kentucky has transitioned from the Kentucky Core Content Tests (KCCT) to the Kentucky Performance Report for Educational Progress (K-PREP), a number of concerns have been raised about the interpretations of K-PREP scores and how they can be used to track the continued progress of Kentucky students and schools. K-PREP is designed to measure new content standards in mathematics, reading, and writing following adoption of the Common Core State Standards through Senate Bill 1 in 2009. Science and social studies standards have not been established within the common core framework, so new items have not been developed for those subjects. K-PREP also includes other design changes such as differences in item weighting and a new reporting scale for all content areas. Based on the content changes in mathematics, reading, and writing and design changes that affect all content areas, the change to K-PREP constitutes a break in the trend line established by the KCCT administrations.

Breaking the KCCT trend line means that direct comparisons with K-PREP should not be made. Students' scale scores are not comparable because of the new scale and their classifications (Novice, Apprentice, Proficient, and Distinguished (NAPD)) are not comparable because the performance standards were changed for K-PREP. Kentucky changed the tests and standards to reflect the newly adopted content standards and the goal of ensuring that all Kentucky students graduate college and career ready. Changing the trend line also helped address concerns regarding score patterns from KCCT. Among the concerns that have been raised is that prior years' scores may have been artificially inflated as teachers "taught to the test" in an effort to help ensure students performed at adequate levels. Students from lower-income families (as measured by eligibility for free or reduced price meals) have historically tended to score lower on assessments. In response to this trend, specific concerns were raised that schools with higher numbers of low-income students would be more likely to use test-driven instruction, resulting in higher scores than what would have been achieved without the test preparation focus (assuming that the impetus to "teach to the test" would be greater for lower performing schools). If this were the case, and Kentucky teachers had developed a "teach to the test" approach to help boost the performance of at-risk students, then it would stand to reason that the introduction of a new test based on new standards would disproportionately impact schools with higher percentages of low income students. The logic is that teachers would not have been able to teach as effectively to the new test, so the score inflation demonstrated in the past would be lost, resulting in larger overall losses for these schools.

We should note that "teaching to the test" refers to teaching specific test item formats, teaching parallel forms of assessments, concentrated drill and practice, or only emphasizing content or algorithms based on the content of the test items. The concern with these test preparation practices is a narrowing of instruction that does not meet rigorous content standards and threatens test score interpretation (Grogan, 2001; Koretz, 2008; Moore, 19994). "Teaching to the test" does not, in this context, mean teaching to the standards, which is expected and encouraged for all schools and teachers. A simple example of teaching the standards might include teaching students several ways to subtract one number from another. If, however, the test

only had subtraction items formatted such that one number was stacked on top of another with the place values aligned, teachers might favor teaching only that format for subtraction. Teaching only a simple algorithm for subtraction, rather than teaching students the deeper math that would allow them to deal with any subtraction problem, however formatted, is an example of “teaching to the test.” It should also be noted that the Common Core State Standards were adopted by Kentucky nearly two years prior to the first K-PREP administration in 2012. The test changed in 2012 to better reflect the standards, but the content was being phased in throughout the state prior to the first administration of K-PREP through professional development for teachers and implementation of standards in classroom instruction.

The purpose of this study is to verify if a pattern exists in which schools with larger percentages of free or reduced lunch students did in fact experience less in terms of overall achievement gains in 2012 than in recent years, and to document the magnitude of that pattern. Additionally, evidence will be examined for differences based on test-preparation practices in schools with higher numbers of low-income students.

Data Description

Four separate student-level data files for 2009 through 2012 were provided by KDE. As the present analysis focuses on grades 3-8 only, all records for those students in grades 9-12 were removed from each file. Student records were also removed if they had been marked for exclusion from accountability calculations (EXC=1), or if they were tested using the alternate assessment (TEST_TYPE= 1 or 2).

Each final student-level data file was aggregated to the school and grade level. The school code used for the aggregation procedure was the code of accountability (CODEOD in 2009 through 2011, whereas there was a separate CODE for each subject area in 2012), which represents the school that the student is accountable to for assessment. This process produced four aggregate files containing the percentage of students receiving free or reduced lunch and the average scale scores for each school by grade level. These files were then merged by school code.

Because schools had a different percentage of students receiving free/reduced lunch each year, an average free-reduced lunch percentage was calculated for each school. Table 1 presents average school mean scale scores for each year as well as the average percentage of students receiving free or reduced lunch.

Table 1. School Grand Mean Scale Scores and Percentage of Students Receiving Free or Reduced Lunch

Grade Level		Reading				Math				Science/Social Studies				% Free/Reduced Lunch
		2009	2010	2011	2012*	2009	2010	2011	2012*	2009	2010	2011	2012*	
3 rd	n	703	703	702	718	703	703	702	718					718
	Mean	53.69	53.59	56.20	208.81	54.22	55.47	56.58	205.52					61.76
	SD	7.10	6.74	6.87	5.97	8.55	8.33	7.85	7.19					19.15
4 th	n	705	705	705	719	705	705	705	719	727	723	719	719	719
	Mean	52.28	54.17	50.73	207.85	51.67	53.91	54.06	206.12	49.64	49.87	49.89	217.60	61.17
	SD	7.09	6.71	6.32	5.75	8.46	8.51	7.88	6.42	7.95	8.16	7.72	6.46	19.40
5 th	n	698	698	698	714	698	698	698	714	722	717	714	714	714
	Mean	49.42	51.20	51.65	207.20	48.30	49.69	49.03	205.74	44.88	42.44	44.47	214.40	60.42
	SD	6.71	7.09	6.95	5.58	8.60	8.84	8.31	6.69	8.36	8.28	8.66	7.02	19.33
6 th	n	398	398	397	409	398	398	397	409					409
	Mean	49.21	50.26	51.27	207.92	47.46	49.10	50.07	206.42					61.28
	SD	6.54	6.25	5.77	5.28	7.66	7.62	6.70	6.83					17.21
7 th	n	313	313	313	328	313	313	313	328	324	326	327	328	328
	Mean	46.75	47.73	47.77	208.54	46.25	47.18	47.70	205.20	45.43	43.83	46.48	214.81	58.99
	SD	6.36	5.54	5.94	5.07	7.71	7.84	7.19	6.49	7.10	7.01	7.04	5.37	17.65
8 th	n	311	311	311	325	311	311	311	325	324	323	324	325	325
	Mean	47.62	49.13	49.04	208.12	42.34	43.06	44.91	206.29	42.09	42.72	44.31	214.59	57.25
	SD	5.29	5.72	5.21	4.96	8.01	7.64	7.47	6.78	7.04	7.09	6.74	7.14	17.24

* 2012 data reflect a new reporting scale. *Note:* Science is assessed in grades 4 and 7. Social Studies is assessed in grades 5 and 8.

Analyses and Results

Correlational Analysis

First, changes in the percentage of students classified as proficient or higher were computed 2009-10, 2010-11, and 2011-12. Proficiency changes were calculated by subtracting the school's percentage of students scoring proficient or higher in the previous year from that of the more current year so that positive values would reflect gains. This was done separately for each subject area at each grade level. Next, Pearson's product-moment correlations were calculated separately for each grade level, indicating the relation between the percentage of schools' students receiving free or reduced lunch and their proficiency gains for each subject area.

It should be noted that no tables are presented to indicate the change in percent proficient from year to year. Such information would be interpretable from 2009 to 2010 and from 2010 to 2011 because the scale and cut scores did not change. However, because the scale and cut scores did change in 2012, the difference in percent proficient would be misleading for the change from 2011 to 2012. Kentucky adopted a more rigorous definition for Proficiency in 2012, and as a result, the percent of students meeting proficiency was lower across the commonwealth. On the other hand, schools with higher percentages of proficient students in 2011 tended to also have higher percentages of proficient students in 2012. So, while the overall percentage of proficient students dropped, the rank order of schools by percent proficient remained very similar between 2011 and 2012. Therefore, while direct differences between percent proficient are not easily interpreted, the correlation between the percentage of schools' students receiving free or reduced lunch and the school-level proficiency gains should be directly comparable for all indicated years.

Table 2 presents these results. While correlations between 2009-10 gains and 2010-11 gains tended not to be statistically significantly different than 0, all correlations between 2011-12 school level proficiency gains and the percentage of students receiving free or reduced lunch were negative, weak to moderate in magnitude, and statistically significant. This suggests that schools serving a higher proportion of students who receive free or reduced lunch tended to experience lower proficiency gains between 2011 and 2012, a pattern not seen in recent years.

Table 2. Correlations between percentage of students receiving free/reduced lunch and school changes in percent proficient by grade level

Grade	Reading			Math			Science/Social Studies		
	09-10	10-11	11-12	09-10	10-11	11-12	09-10	10-11	11-12
3	.045	.083*	-.298**	.084*	.076*	-.263**			
4	.043	-.006	-.249**	.032	.103*	-.325**	.043	.005	-.171**
5	.021	-.025	-.241**	.014	.037	-.149**	.033	-.013	-.093*
6	.061	.096	-.318**	.085	.020	-.223**			
7	.043	.041	-.302**	.084	-.011	-.268**	-.015	.095	-.169*
8	.042	.013	-.257**	.054	.066	-.182**	.098	.013	-.111*

*p < .05; **p ≤ .001. Note: Science is assessed in grades 4 and 7. Social Studies is assessed in grades 5 and 8.

One potential problem with focusing on proficiency gains is that proficiency classifications are based on cut scores that were reset based on the new assessment and reporting scale introduced in 2012. Assuming that the scores are normally distributed, moving the cut score could produce very large changes depending on the magnitude of the change and the location of the cut score in the distribution. To explore whether the observed pattern of correlations was an artifact of changes to the proficiency cut scores, similar analyses were conducted using scale score data. Because 2012 marked a change in the assessment and reporting scale, school-level mean scale scores were first converted to standardized z-scores with a mean of zero and a standard deviation of one. Mean gains were then calculated by subtracting the standardized mean of the earlier year from that of the more recent year, such that positive values would reflect a positive change in standardized average scale score across the two years. These mean gain scores are shown on Table 3.

Table 2. Average standardized mean gain scores by grade level

Grade	Reading		Math		Science/Social Studies		
	Mean	SD	Mean	SD	Mean	SD	
3	09-10	.000	.744	.000	.738		
	10-11	-.003	.806	-.001	.830		
	11-12	.005	.872	.004	.953		
4	09-10	.000	.756	.000	.818	.068	.705
	10-11	.000	.744	.000	.759	-.096	.634
	11-12	.003	.868	.002	.898	-.002	.784
5	09-10	.000	.759	0.000	0.728	-0.117	0.737
	10-11	0.000	0.727	0.000	0.751	0.011	0.717
	11-12	0.004	0.883	0.004	0.899	-0.017	0.814
6	09-10	0.000	0.827	0.000	0.840		
	10-11	-0.004	0.788	-0.003	0.774		
	11-12	0.010	0.882	0.005	0.900		
7	09-10	0.000	0.763	0.000	0.755	-0.130	0.622
	10-11	0.000	0.710	0.000	0.678	0.187	0.562
	11-12	0.009	0.813	0.020	0.714	0.001	0.648
8	09-10	0.000	0.718	0.000	0.701	0.263	0.586
	10-11	0.000	0.647	0.000	0.652	-0.052	0.535
	11-12	0.012	0.741	0.026	0.772	0.036	0.648

Note: Science is assessed in grades 4 and 7. Social Studies is assessed in grades 5 and 8.

Finally, Pearson's product-moment correlations were calculated separately for each grade level, indicating the relation between the percentage of schools' students receiving free or reduced lunch and their standardized mean scale score changes.

As Table 3 shows, while changes in standardized school means tended to be not statistically significantly related to the percentage of students receiving free or reduced lunch between 2009-10 and 2010-11, changes in standardized school means between 2011-12 and free/reduced lunch percentages were consistently negative and statistically significantly

correlated. In other words, though the relationship wasn't particularly strong, schools with higher percentages of students receiving free or reduced lunch tended to experience less positive change, or more negative change, in standardized mean scale scores between 2011 and 2012.

Table 3. Correlations between percentage of students receiving free/reduced lunch and school standardized mean scale score changes by grade level

Grade	Reading			Math			Science/Social Studies		
	09-10	10-11	11-12	09-10	10-11	11-12	09-10	10-11	11-12
3	.027	.068	-.207**	.094*	.025	-.177**			
4	.027	-.003	-.189**	.028	.072	-.246**	.055	-.04	-.253**
5	.035	-.065	-.197**	.078*	-.034	-.097*	-.001	-.014	-.100*
6	.008	.054	-.220**	.111*	-.055	-.109*			
7	.022	.012	-.280**	.096	-.057	-.253**	.050	.059	-.286**
8	.046	.026	-.184**	-.005	.072	-.118*	.145*	.027	-.278**

*p < .05; **p ≤ .001. Note: Science is assessed in grades 4 and 7. Social Studies is assessed in grades 5 and 8.

Effect size Analysis

To further explore the relation between schools' percentage of students receiving free/reduced lunch and changes in school means, schools were divided into quartiles based on free/reduced lunch percentages. First, the distribution of change scores was analyzed within quartiles across years to determine if it would be appropriate to compare average change scores. Across the three years in which change scores were calculated, the distributions of change scores within each quartile were similar, with some schools experiencing positive change and some experiencing negative change. Removal of outlier schools (standardized mean changes greater than 3 or less than -3) did not affect the mean change values in any systematic way. Based on this, it was decided that it would be appropriate to compare average change scores.

Next, differences in average score change between the lowest and highest quartiles based on percentage of students receiving free or reduced lunch were used to calculate an effect size (Cohen's d). Effect sizes reflect differences between means on a standardized metric. These effect sizes were then converted to the 2012 K-Prep reporting scale for ease of interpretation. Overall, schools with higher percentages of students receiving free or reduced lunch tended to have higher average change scores from 2009-10 and 2010-11, but tended to have lower average change scores in 2011-12. Tables 4 through 15 present the average gains by SES quartiles and the effect size comparisons.

Table 4. Grade 3 average score differences by school SES quartiles

Grade	Quartile		Reading Average Score Differences			Math Average Score Differences		
			09-10	10-11	11-12	09-10	10-11	11-12
03	1	Mean	-.002	-.07	.24	-.10	.01	.23
		SD	.552	.618	.606	.541	.625	.697
	2	Mean	-.04	-.06	.10	-.06	-.07	.12
		SD	.656	.736	.742	.779	.720	.774
	3	Mean	-.01	-.02	-.05	.11	-.09	-.03
		SD	.743	.785	.993	.759	.822	1.001
	4	Mean	.04	.15	-.26	.05	.14	-.30
		SD	.960	1.009	.993	.902	1.066	1.172

Notes: Quartile 1 reflects the lowest 25% of the distribution of percentage of students receiving free/reduced lunch.

Table 5. Grade 3 average score differences: Effect sizes between the lowest and highest school SES quartiles

Grade 3	Reading Average Score Differences			Math Average Score Differences		
	09-10	10-11	11-12	09-10	10-11	11-12
Effect Size	-.06	-.26	.61	-.19	-.15	.56
2012 Scale Score Points	.36	1.55	-3.64	1.37	1.08	-4.03

Notes: Effect size calculated between lowest and highest quartiles of percentage of students receiving free/reduced lunch. Effect sizes have been converted to their equivalent in 2012 scale score points for illustrative purposes only.

Table reads: Average reading gains from 2009 to 2010 were roughly 0.36 points (on the 2012 reporting scale metric) higher for schools in the highest quartile compared to the lowest quartile in percentage of students receiving free/reduced lunch. Average reading gains from 2011 to 2012 were roughly 3.64 points lower for schools in the highest quartile compared to the lowest quartile in percentage of students receiving free/reduced lunch.

Table 6. Grade 4 average score changes by school SES quartiles

Grade	Quartile		Reading Average Score Changes			Math Average Score Changes			Science Average Score Changes		
			09-10	10-11	11-12	09-10	10-11	11-12	09-10	10-11	11-12
04	1	Mean	-.05	.06	.24	-.05	-.01	.32	-.02	-.03	.30
		SD	.610	.566	.690	.601	.518	.651	.508	.528	.609
	2	Mean	-.02	-.07	.07	-.04	-.07	.04	.12	-.10	.04
		SD	.609	.660	.870	.619	.662	.848	.570	.539	.722
	3	Mean	.04	-.09	-.09	.03	-.06	-.05	.05	-.18	-.06
		SD	.848	.804	.836	.934	.840	.848	.808	.683	.719
	4	Mean	.02	.11	-.19	.05	.13	-.30	.12	-.07	-.28
		SD	.897	.882	.994	1.014	.925	1.075	.850	.742	.934

Notes: Quartile 1 reflects the lowest 25% of the distribution of percentage of students receiving free/reduced lunch.

Table 7. Grade 4 average score changes: Effect sizes between the lowest and highest school SES quartiles

Grade 4	Reading Average Score Changes			Math Average Score Changes			Science Average Score Changes		
	09-10	10-11	11-12	09-10	10-11	11-12	09-10	10-11	11-12
Effect Size	-.08	-.06	.51	-.12	-.19	.70	-.19	.07	.73
2012 Scale Score Points	.46	.35	-2.93	.77	1.22	-4.49	-1.2	-.45	-4.72

Note: Effect size calculated between lowest and highest quartiles of percentage of students receiving free/reduced lunch. Effect sizes have been converted to their equivalent in 2012 scale score points for illustrative purposes only.

Table reads: Average reading gains from 2009 to 2010 were roughly .46 points (on the 2012 reporting scale metric) higher for schools in the highest quartile compared to the lowest quartile in percentage of students receiving free/reduced lunch. Average reading gains from 2011 to 2012 were roughly 2.93 points lower for schools in the highest quartile compared to the lowest quartile in percentage of students receiving free/reduced lunch.

Table 8. Grade 5 average score changes by school SES quartiles

Grade	Quartile		Reading Average Score Changes			Math Average Score Changes			Social Studies Average Score Changes		
			09-10	10-11	11-12	09-10	10-11	11-12	09-10	10-11	11-12
05	1	Mean	-.07	.06	.32	-.11	.04	.21	-.16	-.01	.14
		SD	.639	.640	.705	.604	.654	.734	.676	.566	.613
	2	Mean	.02	-.03	-.14	.04	-.03	-.09	-.02	-.02	-.05
		SD	.692	.686	.841	.571	.658	.818	.688	.656	.723
	3	Mean	.08	-.03	-.07	.08	-.02	-.05	-.10	.04	-.03
		SD	.827	.844	.899	.719	.796	.918	.759	.791	.838
	4	Mean	-.03	.00	-.11	.00	.00	-.08	-.17	.02	-.12
		SD	.854	.727	.991	.941	.873	1.063	.808	.825	1.01

Notes: Quartile 1 reflects the lowest 25% of the distribution of percentage of students receiving free/reduced lunch.

Table 9. Grade 5 average score changes: Effect sizes between the lowest and highest school SES quartiles

Grade 5	Reading Average Score Changes			Math Average Score Changes			Social Studies Average Score Changes		
	09-10	10-11	11-12	09-10	10-11	11-12	09-10	10-11	11-12
Effect Size	-.05	.09	.51	-.14	.04	.31	.003	-.03	.32
2012 Scale Score Points	.28	-.50	-2.85	.94	-.27	-2.07	-.02	.21	-2.25

Note: Effect size calculated between lowest and highest quartiles of percentage of students receiving free/reduced lunch. Effect sizes have been converted to their equivalent in 2012 scale score points for illustrative purposes only.

Table reads: Average reading gains from 2009 to 2010 were roughly .28 points (on the 2012 reporting scale metric) lower for schools in the highest quartile compared to the lowest quartile in percentage of students receiving free/reduced lunch. Average reading gains from 2011 to 2012 were roughly 2.85 points lower for schools in the highest quartile compared to the lowest quartile in percentage of students receiving free/reduced lunch.

Table 10. Grade 6 average score changes by school SES quartiles

Grade	Quartile		Reading Average Score Changes			Math Average Score Changes		
			09-10	10-11	11-12	09-10	10-11	11-12
06	1	Mean	-.13	-.01	.35	-.25	.12	.09
		SD	.589	.653	.808	.521	.575	.627
	2	Mean	.12	-.06	-.05	.07	-.03	.09
		SD	.671	.639	.726	.595	.570	.610
	3	Mean	-.03	.03	-.10	.07	-.07	.04
		SD	.959	.751	.844	1.03	.805	.906
	4	Mean	.01	.03	-.13	.07	-.01	-.21
		SD	.999	1.058	1.061	1.104	1.053	1.283

Notes: Quartile 1 reflects the lowest 25% of the distribution of percentage of students receiving free/reduced lunch.

Table 11. Grade 6 average score changes: Effect sizes between the lowest and highest school SES quartiles

Grade 6	Reading Average Score Changes			Math Average Score Changes		
	09-10	10-11	11-12	09-10	10-11	11-12
Effect Size	-.17	-.04	.51	-.39	.15	.30
2012 Scale Score Points	.90	.21	-2.69	2.66	-1.02	-2.05

Note: Effect size calculated between lowest and highest quartiles of percentage of students receiving free/reduced lunch. Effect sizes have been converted to their equivalent in 2012 scale score points for illustrative purposes only.

Table reads: Average reading gains from 2009 to 2010 were roughly .90 points (on the 2012 reporting scale metric) higher for schools in the highest quartile compared to the lowest quartile in percentage of students receiving free/reduced lunch. Average reading gains from 2011 to 2012 were roughly 2.69 points lower for schools in the highest quartile compared to the lowest quartile in percentage of students receiving free/reduced lunch.

Table 12. Grade 7 average score changes by school SES quartiles

Grade	Quartile		Reading Average Score Changes			Math Average Score Changes			Science Average Score Changes		
			09-10	10-11	11-12	09-10	10-11	11-12	09-10	10-11	11-12
07	1	Mean	-.003	-.11	.38	-.07	-.03	.32	-.14	.08	.30
		SD	.607	.569	.788	.526	.538	.566	.388	.454	.462
	2	Mean	-.03	.12	-.08	-.04	.08	-.04	-.15	.25	-.002
		SD	.751	.558	.660	.780	.557	.635	.629	.519	.541
	3	Mean	-.04	.02	-.04	.05	-.03	.05	-.15	.19	-.02
		SD	.827	.821	.887	.763	.812	.700	.592	.558	.678
	4	Mean	.09	-.07	-.25	.03	-.05	-.29	-.07	.22	-.30
		SD	.874	.890	.823	.930	.812	.856	.847	.707	.791

Notes: Quartile 1 reflects the lowest 25% of the distribution of percentage of students receiving free/reduced lunch.

Table 13. Grade 7 average score changes: Effect sizes between the lowest and highest school SES quartiles

Grade7	Reading Average Score Changes			Math Average Score Changes			Science Average Score Changes		
	09-10	10-11	11-12	09-10	10-11	11-12	09-10	10-11	11-12
Effect Size	-.12	-.05	.77	-.20	.04	.84	-.11	-.24	.91
2012 Scale Score Points	.61	.25	-3.90	1.30	-.26	-5.45	.59	1.29	-4.89

Note: Effect size calculated between lowest and highest quartiles of percentage of students receiving free/reduced lunch. Effect sizes have been converted to their equivalent in 2012 scale score points for illustrative purposes only.

Table reads: Average reading gains from 2009 to 2010 were roughly .61 points (on the 2012 reporting scale metric) higher for schools in the highest quartile compared to the lowest quartile in percentage of students receiving free/reduced lunch. Average reading gains from 2011 to 2012 were roughly 3.90 points lower for schools in the highest quartile compared to the lowest quartile in percentage of students receiving free/reduced lunch.

Table 14. Grade 8 average score changes by school SES quartiles

Grade	Quartile		Reading Average Score Changes			Math Average Score Changes			Social Studies Average Score Changes		
			09-10	10-11	11-12	09-10	10-11	11-12	09-10	10-11	11-12
08	1	Mean	-.05	-.04	.17	-.02	-.05	.14	.16	-.11	.26
		SD	.642	.536	.766	.508	.467	.584	.474	.454	.553
	2	Mean	.06	-.001	-.02	.06	-.001	.05	.24	-.01	.04
		SD	.769	.632	.715	.651	.538	.595	.637	.472	.539
	3	Mean	-.11	.10	-.07	.01	-.07	.03	.25	-.03	-.02
		SD	.657	.599	.693	.808	.668	.845	.527	.610	.682
	4	Mean	.10	-.07	-.03	-.07	.14	-.16	.44	-.07	-.16
		SD	.785	.824	.790	.828	.918	1.050	.664	.615	.773

Notes: Quartile 1 reflects the lowest 25% of the distribution of percentage of students receiving free/reduced lunch.

Table 15. Grade 8 average score changes: Effect sizes between the lowest and highest school SES quartiles

Grade 8	Reading Average Score Changes			Math Average Score Changes			Social Studies Average Score Changes		
	09-10	10-11	11-12	09-10	10-11	11-12	09-10	10-11	11-12
Effect Size	-.20	.03	.26	.07	-.26	.34	-.48	-.07	.63
2012 Scale Score Points	.99	-.15	-1.29	-.47	1.76	-2.31	3.43	.50	-4.50

Note: Effect size calculated between lowest and highest quartiles of percentage of students receiving free/reduced lunch. Effect sizes have been converted to their equivalent in 2012 scale score points for illustrative purposes only.

Table reads: Average reading gains from 2009 to 2010 were roughly .99 points (on the 2012 reporting scale metric) higher for schools in the highest quartile compared to the lowest quartile in percentage of students receiving free/reduced lunch. Average reading gains from 2011 to 2012 were roughly 1.29 points lower for schools in the highest quartile compared to the lowest quartile in percentage of students receiving free/reduced lunch.

Achievement Gaps

One possible explanation for schools with higher percentages of students receiving free or reduced lunch demonstrating lower standardized mean scale score changes would be that the score gap between students receiving free or reduced lunch and those not grew wider in 2012 than in past years. To explore this issue, effect size (Cohen's d) differences between the mean scale scores of free/reduced lunch and non-free/reduced lunch students were calculated separately by grade level for each assessment year from 2009 through 2012. The negative values presented in Table 16 indicate that students receiving free or reduced lunch have tended to score lower than their counterparts on the state assessment since 2009. However, there also appears to be a marked increase in the gap in 2012, as the magnitude of the effect sizes are larger than in any other year across all grade levels.

Table 16. Effect size differences in mean scale scores for free/reduced lunch and non-free/reduced lunch students

	Reading				Math				Science/Social Studies			
	2009	2010	2011	2012	2009	2010	2011	2012	2009	2010	2011	2012
Grade 3	-0.57	-0.58	-0.60	-0.68	-0.59	-0.56	-0.61	-0.68				
Grade 4	-0.58	-0.57	-0.60	-0.70	-0.59	-0.62	-0.59	-0.70	-0.60	-0.58	-0.65	-0.71
Grade 5	-0.57	-0.56	-0.59	-0.70	-0.62	-0.63	-0.62	-0.70	-0.66	-0.66	-0.67	-0.74
Grade 6	-0.62	-0.63	-0.60	-0.72	-0.64	-0.62	-0.62	-0.73				
Grade 7	-0.58	-0.62	-0.63	-0.72	-0.64	-0.67	-0.68	-0.74	-0.62	-0.68	-0.67	-0.72
Grade 8	-0.58	-0.62	-0.63	-0.72	-0.63	-0.66	-0.67	-0.72	-0.67	-0.66	-0.71	-0.76

Note: Science is assessed in Grades 4 and 7. Social Studies is assessed in Grades 5 and 8.

Discussion and Conclusions

First, some important caveats must be considered. K-PREP is based on what has been described as a more rigorous set of content standard and it was expected that proficiency levels would be lower as students are held to higher standards. Also, as K-PREP is reported on a different scale than KCCT, direct comparisons of average scale scores cannot be made. In spite of the changes to the assessment that are expected to impact comparisons at the student level, we would not anticipate major shifts in the relative standing of schools. School-level scores across the years were correlated in the .6 to .7 range. By standardizing school means and looking at gains in that metric, we are able to address differences in reporting scales, though that does not address all the differences in the two assessments. Results should therefore be interpreted with caution.

The results presented above indicate a distinct pattern in which the 2012 K-PREP administration coincided with lower overall changes in student proficiency and average scale scores among schools with higher percentages of low-income students, as well as increases in the overall score gap between students eligible for free or reduced lunch and their higher income counterparts.

The purpose of this study was to document and verify the existence of this pattern and estimate the magnitude of the effect. These results do not, however, constitute specific evidence of score inflation or of the instructional approaches used by teachers. In fact, some of the evidence runs counter to the score inflation explanation. The pattern was apparent in mathematics and reading with assessments on new content standards with new items. However, the same pattern was observed in science and social studies, which were assessed on the same standards as prior years with the same items. This suggests there is a systematic difference in the results this year, but this difference may not be based on the change in standards or inappropriate test-preparation methods.

While this study cannot substantiate that score inflation resulted in the pattern of higher income schools improving more than lower income schools in 2012, the fact that such a consistent pattern exists warrants further study. On average there are differences in school score changes by school concentration of low-income students, but there is also considerable

variability. Some schools with a high percentage of students eligible for free or reduced lunch experienced positive standardized mean scale score changes between 2011 and 2012 while other schools with low concentrations experienced negative changes. This suggests that further investigation is warranted to identify the school-level characteristics that may explain that variability. It may be necessary to visit schools to generate suspect characteristics to investigate. If score inflation is a substantial contributor to this pattern, it may be only one of several contributing factors. The persistence of the pattern across science and social studies points to additional factors beyond score inflation alone.

It is also recommended that these analyses be conducted again using 2013 results. An alternate explanation for the pattern we see could be that schools knew that the trend line was going to be broken. If so, they may not have done everything they had done in prior years to ensure that students scored as well as possible (including score inflation practices) in the hopes of setting a lower “baseline” score. It seems far-fetched that schools would knowingly reduce students’ scores in the hopes of showing larger gain scores in later years, but they may not have felt as much pressure to improve in 2012. It is well documented that schools engage in many types of test preparation and motivational activities in the weeks prior to the accountability assessment, even though the benefit of these activities is unclear based on motivational research (e.g. Deci & Ryan, 2000; Lepper, Corpus, & Iyengar, 2005; Thacker, Koger, Hoffman & Koger, 1999). If schools chose to do fewer of these activities this year than previously, and if these activities were more prevalent at lower-scoring schools, that might help account for the observed pattern. If this were the case, we would expect a reversal of the pattern in 2013.

References

- Deci, E.L., & Ryan, R.M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*, 227-268.
- Grogan, M. (2001). Positive and not-so-positive effects of preparing students to pass Standards of Learning Tests in Virginia’s high school classrooms. *National Association of Secondary School Principals, NASSP Bulletin, 85*(622), 5-14.
- Koretz, D.M. (2008). *Measuring up: what educational testing really tells us*. Harvard University Press. Cambridge, MA.
- Lepper, M., Corpus, J., & Iyengar, S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *Journal of Educational Psychology, 97*, 184-196.
- Moore, W. P. (1994). The devaluation of standardized testing: One district’s response to a mandated assessment. *Applied Measurement in Education, 7*, 343-367.
- Thacker, A. A., Koger, L. E., Hoffman, R. G., & Koger, M. E. (1999). *The Transition from KIRIS to CATS: Instruction, Communication, and Perceptions at 20 Kentucky Schools*. (HumRRO Report DFR-WATSD-99-23). Alexandria, VA: Human Resources Research Organization.