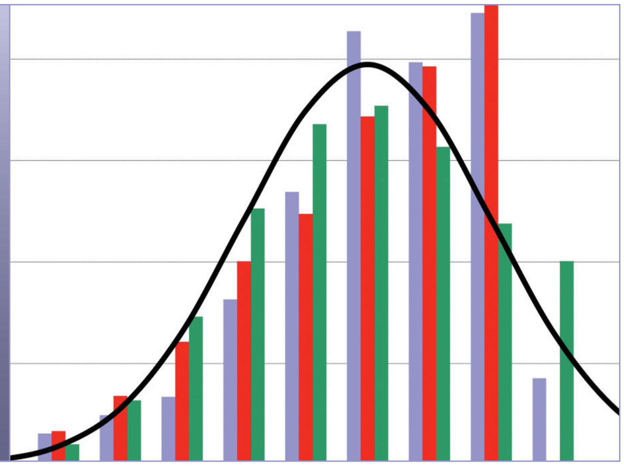
ACT Research Report Series

2012 (4)



Examining the Relationship of QualityCore® Implementation and Student Achievement in Algebra I

Edwin Ndum
Jeff Allen
Teri Fisher

July 2012

ACT®

For additional copies, write:
ACT Research Report Series
P.O. Box 168
Iowa City, IA 52243-0168

Examining the Relationship of QualityCore® Implementation and Student Achievement in Algebra I

Edwin Ndum
Jeff Allen
Teri Fisher

Table of Contents

Abstract

We assessed the relationship of level of QualityCore® implementation and student achievement in Algebra I. The study is based on 1,291 9th grade students in three large metropolitan school districts who took EXPLORE in the fall and the QualityCore Algebra I end-of-course assessment (EOC) in the spring. During the same period, the students' 41 Algebra I teachers completed questionnaires measuring their use of QualityCore's tools and strategies for Algebra I. We found that level of teachers' QualityCore implementation had a positive and statistically significant relationship with student achievement, measured by performance on the QualityCore Algebra I EOC controlling for performance on the fall EXPLORE tests. When different domains of implementation were considered, the *Teaching Practices* and *Collaboration with Colleagues* domains had the strongest relationship with student achievement. Survey results also indicate how QualityCore use varied across components and that QualityCore was well-received by teachers.

Examining the Relationship of QualityCore® Implementation and Student Achievement in Algebra I

Introduction

Years of ACT research show that it is not the number of courses a student takes in high school, but what happens in those courses that matters most (ACT, 2005; ACT & The Education Trust, 2005). Despite educational reforms and improvement initiatives, the percentage of students ready for all four college courses has only increased slightly over the last ten years. In 2011, only one in four students had at least a 50% chance of earning a B or higher college grade in all four first-year courses at a typical postsecondary institution: English Composition, College Algebra, Social Sciences, and Biology (ACT, 2011).

In response to this lack of college readiness, ACT's research and development team partnered with the nation's leading educators to develop QualityCore as a tool for raising the quality and intensity of core high school courses (ACT & The Education Trust, 2005). QualityCore currently supports twelve core high school courses: English 9, 10, 11, and 12; Algebra I, Geometry, Algebra II, Precalculus, Biology, Chemistry, Physics, and U.S. History. QualityCore resources include the following components:

- *Instructional resources* consist of a combination of course standards, test blueprints, and model instructional units that help educators customize instruction to the needs of their students.
- *Formative item pools and test builders* are pools of formative items that can be used to create customized quizzes and temporal benchmark assessments. The use of the formative item pools saves time and provides timely feedback to students.

- *Professional development* resources and workshops help educators expand their teaching skills and learn about effective practices of other teachers.
- *End-of-course assessments* (EOCs) with constructed response and multiple-choice options help educators evaluate student achievement gains in each course.
- *Score reports and progress reporting* provide comparison of students' performance on the EOCs at the local, state, and national levels. They also include evaluation of students' progress towards college readiness in a given course.

With the introduction of any school improvement initiative or new instructional program, school leaders and policymakers want to know if progress is being made in raising student achievement. Moreover, school leaders want to know what aspects of the program are working or not working, and how much teachers are utilizing the program. In response, this study examined the relationship between the levels of teachers' implementation of QualityCore Algebra I and student achievement (measured by the QualityCore Algebra I EOC) for 9th grade students. A hypothetical example of a teacher with a high level of implementation is a teacher who

- worked extensively with his or her colleagues to modify course objectives and lesson plans to address QualityCore course standards;
- adopted and regularly used QualityCore resources that provide systematic approaches to examining student work and modifying lessons to make them more rigorous;
- regularly used the QualityCore formative item pool for interim assessments and classroom instruction; and
- evaluated his or her student's performance on the EOC and used the results to set goals for next year's students.

The study was conducted in three large metropolitan school districts with high concentrations of lower-achieving students, many living in poverty. Many of the schools were designated as needing improvement for not making adequate yearly progress (AYP) towards proficiency, as mandated by the No Child Left Behind Act (NCLB).

The 9th grade Algebra I teachers underwent professional development (PD) to introduce them to the resources offered by QualityCore and how to integrate the resources into their instruction and collaboration. Most of the training took place in the summer before the 2009-2010 academic year. Training topics included defining rigor and relevance of high school courses, analyzing depth of knowledge of test items, analyzing course objectives to ensure alignment, research-based instructional strategies, developing quality in-course assessments, using constructed response items, and creating a course syllabus. Follow-up training was offered during the school year to reinforce the summer training and offer additional support for problems encountered throughout the school year. Topics included techniques for scaffolding instruction to match instruction, examining student work, and revising lesson plans. The last day of follow-up training was only available through online training in two of the districts; in the other district teachers could participate either online or through the traditional face-to-face training.

In one district, all Algebra I teachers were invited to the PD; in the other two districts only selected teachers underwent training. (We did not collect information on how teachers were selected for training). During the same academic year, a sample of Algebra I teachers completed questionnaires about their implementation of QualityCore's instructional practices, tools, and strategies. Although the districts implemented other QualityCore courses (English 9 and Biology), we were able to collect implementation data only on Algebra I.

Research Questions Addressed

The primary research question addressed by the study is:

1) To what extent did level of implementation of QualityCore Algebra I correspond to student achievement in Algebra I?

Secondary research questions include:

2) Which aspects of QualityCore implementation (domains) were most related to student achievement in Algebra I?
3) Which components of QualityCore did teachers use the most?

Methodology

We used a longitudinal study design with measurements of QualityCore implementation occurring between students' initial measure of academic achievement (EXPLORE) and students' end-of-course measure of achievement. The study was carried out by administering surveys to teachers to measure their level of QualityCore implementation. The survey response data was then merged to their students' test scores from fall 2009 (EXPLORE) and spring 2010 (QualityCore).

Measures of QualityCore Implementation

Survey construction and administration. A survey was constructed by ACT staff in fall 2009 to measure different aspects of teachers' use of QualityCore resources. The survey content was based on QualityCore training materials, with a focus on the materials used for the professional development offered to teachers in the three large school districts where the study took place. The survey items covered the use of specific QualityCore tools and resources (e.g., the Formative Item Pool, the Model Instructional Unit, the Template to Examine Assignments for Rigor & Relevance), as well as teaching practices that were not specific to QualityCore but that targeted the professional development and/or are reinforced by the QualityCore program (e.g., meeting with colleagues to review student work; designing a course syllabus; classroom routines such as summarizing lessons, and providing feedback to students).

The first survey was administered in November 2009 and the second survey was administered in May 2010. In the fall, teachers were invited to participate in the survey using emails with links to the web-based survey. In the spring, paper and pencil versions of the survey were delivered in person to teachers with postage-paid return envelopes in an effort to improve survey response rates. The contents of the fall and spring surveys were not identical, but common domains of QualityCore implementation were measured by the two surveys. The fall survey included 49 items and covered more items related to teaching practices and collaboration that were less directly related to the QualityCore training objectives. We shortened the spring survey to encourage more survey respondents. The spring survey included 21 items and was focused more on use of QualityCore-specific resources.

In all, 28 teachers completed the fall survey, 41 teachers completed the spring survey, and 13 teachers completed both the fall and spring versions. Thus, 56 teachers completed at least one survey, and we estimate that this represents 37% of all teachers in the three districts who implemented QualityCore Algebra I in 2009-2010. (An exact count of teachers who implemented QualityCore Algebra I was not available from the districts, but we estimated the count at 152 based on PD training roster data and information provided by some schools.) Among these 56, 41 teachers had students that took EXPLORE in the fall and the QualityCore Algebra I EOC in the spring. The 41 participating teachers were from 19 high schools in 3 large school districts. The sample contained 1,291 grade 9 students (with EXPLORE scores, QualityCore scores, and with teachers who took one or more implementation surveys); on average, there were about 32 students per teacher. This is not an estimate of average class size, however, because teachers could teach multiple sections of Algebra 1. Table 1 summarizes the samples of teachers and students with EXPLORE scores used in the study. The primary analyses are based on the sample of teachers and students with full data; secondary analyses use the entire sample of surveyed teachers.

Table 1

Study Samples

Timing of implementation survey	Data available	N schools	N teachers	N students
Fall 2009	Total	19	28	
	With student data	16	18	507
Spring 2010	Total	17	41	
	With student data	17	35	1,117
Combined Fall & Spring	Total	26	56	
	With student data	19	41	1,291

Implementation domains. The survey items were classified into one of five domains: *Collaboration with Colleagues, Educator Resources, Formative Items, Teaching Practices, and QualityCore Adoption.* The domain classifications were based on descriptions of QualityCore professional development (ACT, 2011) and were created with the goal of distinctively and exhaustively capturing the various aspects of implementation. Below, we describe the domains along with their connections to the overall goals of raising the intensity and quality of Algebra I course. The survey items under each domain for the fall and spring surveys are presented in Appendix A.

Collaboration with Colleagues. This domain measured the amount of time teachers spent with colleagues examining student work, planning lessons, executing instructional strategies, and evaluating student progress. It measured the effort and time teachers spend working together towards developing Algebra I instructional strategies.

Educator Resources. Educator Resources measured the use of specific QualityCore resources such as the template to examine assignments for rigor and relevance, depth of knowledge analysis, and worksheets to examine student work and determine expectations for high quality performance. In addition, it measured the use of teaching strategies from QualityCore's Educator's Toolbox. This domain assessed the extent that teachers learned to use QualityCore Algebra I materials to reflect on, augment, and enhance their own instructional materials.

Formative Items. The QualityCore professional development was intended to improve a teacher's ability to "develop the capacity to analyze formative and summative test items for depth of knowledge and compare those items to differentiate level of (student) thinking required" (ACT, 2011). The Formative Items domain focused on the use of formative constructed-response and multiple choice items from the QualityCore Formative Item Pool (FIP). This domain assessed the extent that teachers used the FIP throughout the school year.

Teaching Practices. Teaching Practices covered a variety of classroom and other practices supported by QualityCore training, model instructional units, and course syllabi. It included items related to classroom routines and organization, as well as some activities teachers performed outside of the classroom in preparation for lessons.

QualityCore Adoption. QualityCore Adoption measured the extent that teachers and their colleagues had "bought in" to QualityCore and had ingrained QualityCore into their teaching and planning activities. It surveyed the level of familiarity with and the use of QualityCore instructional resources, such as model instructional units, as well as the extent that existing course standards were compared with QualityCore course objectives. This domain also measured expectations of and enthusiasm for the QualityCore program.

Survey items and domain scoring. The survey items measured agreement, frequency of behaviors, or time spent. The three sets of survey response options were:

- *Agreement*: NA = Not Applicable, 1 = Strongly Disagree, 2 = Moderately Disagree, 3 = Slightly Disagree, 4 = Slightly Agree, 5 = Moderately Agree, and 6 = Strongly Agree.
- *Frequency of behaviors*: NA = Not Applicable, 0 = Never (None), 1 = Seldom, 2 = Sometimes, 3 = Often, 4 = Frequently, and 5 = Daily.
- *Time spent*: 0 = No Time, 1 = One Hour, 2 = Two Hours, 3 = Three Hours, 4 = Four Hours, 5 = Five Hours, 6 = Six to Ten Hours, and 7 = Eleven or More Hours.

Item responses were assigned numeric values corresponding to the order of the response options. Items with responses of "Not applicable (NA)" were assigned a missing value, while those with "Never" and "None" were assigned a 0. Although most of the prompts were positively framed (e.g., "*I plan on using this year's QualityCore end-of-course exam results to set a baseline for future years' student achievemen*t"), some items, like, "*I don't have the time right now to use QualityCore elements in my work,"* were negative. Responses to items that were negatively structured were reverse-scored.

Using the fall and spring survey data separately, we computed QualityCore implementation domain scores by taking the average of the items classified under each domain, resulting in ten possible scores for each teacher. Each of the ten scores was then standardized to have mean 0 and standard deviation 1. Then, we calculated overall scores for the five domains by taking the average of each pair of standardized spring and fall scores. The aggregate implementation score (combining fall and spring and all domains) was defined as the average of the ten fall and spring domain scores. Overall fall (spring) implementation scores were obtained by averaging the five fall (spring) domain scores. By using standardized scores, we forced the domain scores to carry equal weights in determining the overall measures of QualityCore implementation. For teachers who only completed the fall (spring) survey, the combined implementation measures (reflecting implementation throughout the year) were set to equal the standardized fall (spring) implementation component. The resulting data included five scores measuring domains of implementation (Collaboration with Colleagues, Educator Resources, Formative Items, Teaching Practices, and QualityCore Adoption) and three aggregate measures of implementation – fall, spring, and combined (fall and spring).

The aggregate variables measured the overall level of QualityCore Algebra I implementation, without regard to the various aspects (domains) of implementation. Unlike the domain-specific measures outlined above, the aggregate measures did not convey knowledge of what aspects of implementation were associated with changes in student achievements. However, the aggregate measures are believed to be more reliable because they are based on more survey items.

Measures of Academic Achievement

Fall EXPLORE test. The EXPLORE tests of educational development were administered to the 9th graders early in the fall. EXPLORE measures student development in the curriculum areas of English, mathematics, reading, and science. EXPLORE is most commonly used for students in 8th and 9th grades and focuses on knowledge and skills usually attained by 8th grade. The EXPLORE scale scores in the four subject areas, which are reported on a scale from 1 to 25, were used as measures of initial academic achievement level. The reliabilities of the EXPLORE scale scores are 0.85-0.87 (English), 0.84 (mathematics), 0.83-0.86 (reading), and 0.79-0.84 (science) (ACT, 2007). These values give the degree of consistency in the test scores (ACT, 2007). Reliabilities closer to 1.0 indicates greater consistency or lesser error in test measurements while a value closer to zero reflect little or no consistency, or a higher potential for error in the test scores.

Spring QualityCore test. The QualityCore Algebra I EOC was administered to the 9th graders in the spring. The assessment was aligned to ACT Course Standards (ACT & The Education Trust, 2005) and included problem-based items embedded in contexts that were accessible and relevant to high school students. Scores on the QualityCore Algebra I EOC range from 125 to 175, with reliability estimated at 0.75 (ACT, 2010a). The EOC scores were used as measures of end of course academic achievement level.

Table 2 lists the variables used for analysis, classified according to the level—student or teacher—and when the data was collected.

Table 2

Analysis Variables

Variable	Range	Level	Data Collection
EXPLORE scores			
EXPLORE English	1-25	Student	Fall
EXPLORE Mathematics	1-25	Student	Fall
EXPLORE Reading	1-25	Student	Fall
EXPLORE Science	1-25	Student	Fall
QualityCore Algebra I score	125-175	Student	Spring
Implementation scores			
Fall Implementation	1.94-5.13	Teacher	Fall
Spring Implementation	2.04-5.06	Teacher	Spring
Aggregate Implementation	2.04-5.10	Teacher	Fall and/or Spring
Collaboration with Colleagues	0.50-5.50	Teacher	Fall and/or Spring
Educator Resources	1.00-5.54	Teacher	Fall and/or Spring
Formative Items	0.50-5.50	Teacher	Fall and/or Spring
Teaching Practices	2.20-5.57	Teacher	Fall and/or Spring
QualityCore Adoption	1.33-6.00	Teacher	Fall and/or Spring

Note: Implementation score ranges before standardization are given.

Statistical Modeling

Because students taking Algebra I were nested within teachers, we used a hierarchical linear model that accounted for unobserved teacher effects on student achievement (Goldstein, 2002; Raudenbush & Bryk, 2002; Snidjers & Bosker, 1999). Because we anticipated unobserved teacher effects on student achievement, we specified a random intercept model. A significant variation in teacher intercepts (p-value < 0.05) would indicate variation in student achievement across teachers that is not fully explained by prior achievement and teachers' QualityCore implementation. The hierarchical linear model was used to test the relationship between the different measures of implementation and student achievement.

Student achievement was measured by QualityCore Algebra I EOC score, controlling for students' initial achievement level, as measured by the four EXPLORE scale scores. By regressing the EOC score on measures of prior academic achievement (the four EXPLORE subject test scores), other independent variables that are predictive can be regarded as contributing to student achievement. We included the group mean EXPLORE Mathematics score for each teacher as an additional covariate to capture possible peer effects (Angrist & Lang, 2002): Students surrounded by higher-achieving peers tend to show greater academic achievement. Although the analysis included EXPLORE English, Reading, and Science, only the group mean EXPLORE Mathematics was used as a teacher-level covariate to capture each group's prior mean mathematics level, which is most applicable to the dependent variable (QualityCore Algebra I score). Mean EXPLORE scores in the other subject areas were not included as predictors because we did not think they would help explain the variation in the dependent variable.

To test the primary research question ("To what extent did level of implementation of QualityCore Algebra I correspond to student achievement in Algebra I?"), the aggregate measure of implementation was used as an independent variable in the hierarchical linear model. The primary analysis measured the relationship of QualityCore Algebra I level of implementation to student achievement. Other analyses tested for effects of different aspects of implementation, as measured by the domain implementation scores.

For the ease of computation and interpretation, all predictors were grand-mean centered (Koenig and Lissitz, 2001) before inclusion in the model. Initially, we fit a model with the aggregate (combining fall and spring) implementation measure. Then we fit separate models for the fall and spring implementation measures. Next, we fit models using the five domain measures of implementation. Each domain measure was entered into the model individually, and then a joint model with all five measures was tested. Independent variables were considered statistically significant if the parameter estimate significance test resulted in a p-value of less than 0.05.

Summarizing Survey Data

To address the third research question (*Which components of QualityCore did teachers use the most*?), we assessed the survey response data from all responding teachers, including some that were not included in the regression analyses because their students did not take EXPLORE or the EOC assessment. To summarize the survey data, item response data were coded to numeric values corresponding to the order of the response options (e.g., 1=strongly disagree, 2=moderately disagree, etc.). We calculated each item's mean and standard deviation.

For the items that used the agreement scale, we also reported the percent of teachers agreeing with the statement (slightly agree to strongly agree). For items that asked teachers to report the number of days (out of the last 5 regular instructional days) that they performed a certain activity, we reported the percentage of teachers reporting 3 or more days of performing the activity. For items that asked teachers to report the number of hours (during the last 5 regular instructional days) that they performed a certain activity, we reported the percentage of teachers reporting 3 or more hours of performing the activity.

Results

The demographic composition of the sample is given in Table 3. The vast majority of students in the study sample were African American/Black (85%); Hispanic (6.9%), Asian (3.0%), and Caucasian/White students (2.6%) were less represented. Compared to the population of students enrolled in elementary and secondary schools in the U.S., the sample had a much larger concentration of African American students, and a much smaller concentration of

Caucasian/White and Hispanic students. The sample was nearly evenly split by gender (51% female and 49% male).

Table 3

Student Demographics

Ethnicity	**N (%)**	**Population %[2]**
Race/Ethnicity		
African American/Black	1,096 (85.0)	16.0
American Indian, Alaskan Native	11 (0.8)	1.4
Asian	38 (3.0)	4.4
Caucasian/White	34 (2.6)	57.8
Hispanic	89 (6.9)	20.4
Multiracial	16 (1.2)	-
Other	1 (0.1)	-
Prefer Not To Respond/Missing	4 (0.3)	-
Gender		
Male	632 (49.0)	50.2
Female	657 (51.0)	49.8
Total[1]	1,289 (100.0)	100

[1]Due to rounding, sums may not equal 100.
[2]Source: U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey (SASS), "Public School Data File," 2007-08.

Table 4 (see page 16) presents the descriptive statistics for the variables used in the analyses. Twice as many teachers participated in the spring survey (35) as did in the fall survey (18). The overall measure of level of implementation ("Aggregated Implementation") was positively correlated with end-of-course achievement (r=.08) and mean prior achievement (r=.16). The positive correlation with mean prior achievement suggests that teachers with higher-achieving students made greater use of QualityCore's resources. The fall and spring measures of implementation were highly correlated (r=.68), suggesting that teachers' level of implementation was consistent across the two semesters. Among the implementation domains, *Collaboration* *with Colleagues* (r=.10) and *Teaching Practices* (r=.08) had the highest correlations with end-of-course achievement, while *QualityCore Adoption* had no correlation (r=.00). *Formative Items* (r=.22) had the highest correlation with mean prior achievement level, suggesting that teachers were more likely to use QualityCore's FIP with higher-achieving students. The correlation among the five domains of implementation was lowest for the relationship between *Teaching Practices* and *QualityCore Adoption* (r=.03) and highest for *Teaching Practices* and *Educator Resources* (r=.76).

The academic achievement data suggest that the students sampled were generally lower-achieving in comparison to students nationally. The mean QualityCore Algebra I score was 143 (Table 4, see page 17); a score of 143 is at the 34th percentile nationally. The mean EXPLORE Mathematics score was 13.2; a score of 13 is at the 21st percentile nationally for students tested in fall of 9th grade (ACT, 2007). On average, the entering 9th grade students were over one standard deviation below the EXPLORE 8th grade College Readiness Benchmark score of 17. The correlation of initial mathematics achievement level (EXPLORE Mathematics) and the spring EOC assessment score (QualityCore Algebra I) was 0.38.

Among the five domains used to define the various aspects of implementation, *Collaboration with Colleagues* had the smallest mean rating of 3.0—representing a slight disagreement on the level, or about 3 days a week, of collaborative work dedicated to the students, while *Teaching Practices* resulted in the highest mean rating of 4.5 (indicating slight-moderate agreement on statements of teaching practice, or a frequency of 4-5 days per week of various teaching practices). The standard deviations of the five domains ranged from 0.7 (*Teaching Practices*) to 1.3 (*Formative Items*), suggesting that teachers did not vary much in their responses to the items assessing teaching practices, but varied more in their use of the FIP.

Table 4

Summary Statistics and Correlations

Variable	N Students	N Teachers	1	2	3	4	5	6	7	8	9	10	11
1. QualityCore Alg. I EOC Score	1,291	41	1.00										
2. EXPLORE Mathematics	1,291	41	0.38*	1.00									
3. Mean EXPLORE Mathematics	1,291	41	0.22*	0.39*	1.00								
4. Fall Implementation	507	18	0.11*	-0.01	-0.03	1.00							
5. Spring Implementation	1,117	35	0.07*	0.10*	0.30*	0.68*	1.00						
6. Aggregated Implementation	1,291	41	0.08*	0.04	0.16*	0.94*	0.97*	1.00					
7. Collaboration with Colleagues	1,291	41	0.10*	0.01	0.10*	0.67*	0.66*	0.66*	1.00				
8. Educator Resources	1,291	41	0.06	0.02	0.10*	0.88*	0.80*	0.81*	0.42*	1.00			
9. Formative Items	1,291	41	0.04	0.09*	0.22*	0.80*	0.73*	0.73*	0.19*	0.39*	1.00		
10. Teaching Practices	1,291	41	0.08*	-0.04	-0.03	0.82*	0.54*	0.63*	0.46*	0.76*	0.14*	1.00	
11. QualityCore Adoption	1,291	41	0.00	0.03	0.10*	0.61*	0.70*	0.67*	0.20*	0.34*	0.70*	0.03	1.00
Mean			143.0	13.2	13.3	4.1	3.8	3.8	3.0	4.4	3.3	4.5	3.8
Standard Deviation			3.5	3.5	1.5	0.8	0.8	0.7	1.2	1.1	1.3	0.7	0.9

* p-value < 0.05

Hierarchical Modeling Results

Relationship of level of implementation and student achievement. Table 5 contains the results of the hierarchical linear model assessing the relationship of level of implementation and student achievement in Algebra I. The aggregated implementation measure was significantly predictive of end-of-course student achievement, controlling for prior achievement (EXPLORE scores) and the group mean EXPLORE Mathematics score. A one point increase in the overall fall and spring implementation level was associated with a 0.66 point increase in the QualityCore Algebra I score (beta=0.66, SE=0.27, p-value=0.020). This result addresses the primary research question (*To what extent did level of implementation of QualityCore Algebra I correspond to student achievement in Algebra I*?). A one standard deviation increase in implementation corresponded to a 0.13[1] standard deviation increase in end-of-course achievement level.

Table 5

Combined Fall and Spring Implementation Model Results

Effect	Estimate	SE	P-value	Effect size
Intercept	142.86*	0.18	<0.0001	
Mean EXPLORE Math	0.04	0.12	0.7227	0.02
EXPLORE Math	0.21*	0.03	<0.0001	0.21*
EXPLORE English	0.12*	0.04	0.0008	0.12*
EXPLORE Reading	0.05	0.04	0.1467	0.05
EXPLORE Science	0.15*	0.04	0.0001	0.12*
Aggregated Implementation	0.66*	0.27	0.0201	0.13*
Variance of intercepts	0.97*	0.35	0.0025	
Residual variance	9.57*	0.38	<0.0001	

N=1,291 students, 41 teachers; R = 0.50; *p-value < 0.05

[1] The effect size estimate of 0.13 is derived as the parameter estimate associated with the aggregated implementation measure (0.66) multiplied by the standard deviation of the aggregated implementation measure (0.7), and divided by the standard deviation of the Algebra I EOC test score (3.5) (0.13 = 0.66*0.7/3.5).

Beyond the QualityCore implementation measure, there was other unexplained variation across teachers in student performance on the EOC assessment: The variance of the teacher intercepts was estimated at 0.97 (p-value<0.01); this is evidence of variability in teacher effects on student achievement in Algebra I. Mean prior achievement level was not significantly predictive of EOC performance. As expected, EXPLORE math (beta=0.21, p-value<0.001) was predictive, and students' prior achievement in English and science also helped predict EOC performance.

A measure of intra-class correlation coefficient (Hedges & Rhoads, 2011) was computed as 0.09 (0.97 / [0.97 + 9.57]). Thus, there was 9% similarity in the QualityCore Algebra I achievement among students taught by the same teacher (with student variation in EXPLORE scores and mean EXPLORE math score accounted for). This can also be viewed as the percentage of total variation in the QualityCore Algebra I performance associated with differences between teachers (Singer, 1998).

To get a sense of how much of the teacher variation was attributable to QualityCore implementation differences, we compared the fit of the model with and without the implementation variable. From the model without the implementation variable to that with it, the variance between teachers decreased from 1.06 to 0.97, a 0.09 decrease. Thus, although there was a significant relationship between level of implementation and achievement, it explained only 9% of the explainable variation in QualityCore Algebra I performance between teachers.

Although the aggregate measure of implementation was significantly predictive of EOC performance, when broken down by semester, only the fall implementation measure had a significant relationship. A one-point increase in the fall implementation measure resulted in a 0.94 point increase in QualityCore Algebra I score (Table 6, beta=0.94, p-value<0.05). In contrast, the spring implementation measure was not significantly predictive of Algebra I score, controlling for individual and group prior achievement (EXPLORE scores). The spring implementation effect was estimated at 0.28 (Table 7), with a standard error of 0.27 (p-value>0.05). The results for the other predictor variables and variance components from the models testing the fall and spring implementation measures are similar to those observed in Table 5 for the aggregate implementation measure.

Table 6

Fall Implementation Model Results

Effect	Estimate	SE	P-value	Effect size
Intercept	143.00*	0.27	<0.0001	
Mean EXPLORE Math	-0.13	0.20	0.5067	-0.05
EXPLORE Math	0.23*	0.05	<0.0001	0.24*
EXPLORE English	0.12*	0.06	0.0669	0.11
EXPLORE Reading	0.02	0.06	0.7504	0.02
EXPLORE Science	0.12*	0.06	0.0371	0.11*
Fall Implementation	0.94*	0.35	0.0172	0.19*
Variance of intercepts	0.88*	0.52	0.0455	
Residual variance	9.18*	0.59	<0.0001	

N=507 students, 18 teachers; R = 0.47; *p-value < 0.05

Table 7

Spring Implementation Model Results

Effect	Estimate	SE	P-value	Effect size
Intercept	142.86*	0.18	<0.0001	
Mean EXPLORE Math	0.17	0.12	0.1578	0.07
EXPLORE Math	0.19*	0.04	<0.0001	0.18*
EXPLORE English	0.10*	0.04	0.0087	0.10*
EXPLORE Reading	0.10*	0.04	0.0094	0.09*
EXPLORE Science	0.14*	0.04	0.0006	0.12*
Spring Implementation	0.28	0.27	0.3070	0.05
Variance of intercepts	0.68*	0.29	0.0092	
Residual variance	9.58*	0.41	<0.0001	

N=1,117 students, 35 teachers; R = 0.50; *p-value < 0.05

Implementation domains and student achievement. To gain more insight into the relationship between implementation and achievement in student achievement, we tested whether the five implementation domain measures were predictive of student achievement. This analysis helps us understand which aspects of QualityCore implementation were most related to student achievement. We first entered each domain measure into the model individually (without the other four domain measures). These models are identical to the combined fall and spring implementation model (Table 5), except that each domain measure replaces the aggregate implementation measure. Then, we fit a model that included all five domain scores as predictors. The two sets of analyses let us examine each domain score's relationship with student achievement, with and without consideration of scores from the other domains.

Presented in Appendix B, Tables B-1 though B-5 contain the parameter estimates for each of the five implementation domains assessed individually. While the estimates for all five domains are positive, only the estimates for *Collaboration with Colleagues* and *Teaching Practices* are significantly predictive of EOC performance, conditioning on individual and group EXPLORE scores. A one point increase in the *Teaching Practices* measure was associated with a 0.52 point increase in the QualityCore Algebra I score (p-value<0.05). A one point increase in the *Collaboration with Colleagues* measure was associated with a 0.46 point increase in the QualityCore Algebra I score (p-value<0.05). The other domains showed positive but non-significant relationships. These results inform the second research question (*Which aspects of QualityCore implementation (domains) were most related to student achievement in Algebra I?*).

Table 8 below contains the output of the hierarchical linear model of QualityCore Algebra I EOC scores regressed on all five implementation domain measures, controlling for individual and group EXPLORE Mathematics scores. When the implementation domain measures were considered jointly in the same model, only *Teaching Practices* was significantly predictive of EOC performance, conditioning on individual and group EXPLORE scores. A one point increase in the *Teaching Practices* measure was associated with a 0.65 point increase in the QualityCore Algebra I score (SE=0.29, p-value<0.05). Domains showing a positive but non-significant relationship include *Collaboration with Colleagues* (beta=0.24), *Formative Items* (beta=0.04), and *QualityCore Adoption* (beta=0.26). *Educator Resources* had a negative but non-significant relationship (beta=-0.37).

Table 8

Domains of Implementation, Full Model Results

Effect	Estimate	SE	P-value
Intercept	142.89*	0.19	
Mean EXPLORE Math	0.09	0.13	0.4943
EXPLORE Math	0.21*	0.03	<0.0001
EXPLORE English	0.12*	0.04	0.0009
EXPLORE Reading	0.05	0.04	0.1513
EXPLORE Science	0.15*	0.04	0.0002
Collaborate with Colleagues	0.24	0.27	0.3935
Educator Resources	-0.37	0.30	0.2359
Formative Items	0.04	0.30	0.8908
Teaching Practices	0.65*	0.29	0.0319
QualityCore Adoption	0.26	0.37	0.4781
Variance of intercepts	0.98*	0.37	
Residual variance	9.57*	0.38	

N=1,291 students, 41 teachers; R = 0.50; *p-value < 0.05

Further assessment of the implementation domains model showed that transitioning from the model without to a model with the five domains of implementation, the between-teachers variance decreased from 1.04 to 0.98 (Table 8), a 0.06 point decrease. Thus, the five domain measures accounted for 6% of the explainable variation (Snidjers and Bosker, 1999) in QualityCore Algebra I performance between teachers. In addition, the intra-class correlation was 0.09 (0.98 / [0.98 + 9.57]), so 9% of the total variation in QualityCore Algebra I performance was associated with differences between teachers.

Survey Results Measuring QualityCore Use

Beyond the investigation of QualityCore implementation's relationship to student achievement, the survey results provided some insights into what components of QualityCore were used most by teachers. Additionally, the results provided information on teachers' perceptions of QualityCore's tools and overall effectiveness. As described earlier, the study included 41 teachers who had been surveyed and had students with EXPLORE and QualityCore Algebra I EOC data. In all, 56 teachers were surveyed —15 teachers did not have students with EXPLORE or EOC assessment scores. In Appendix A, we summarize the survey responses for these 56 teachers. Across survey items, the maximum possible sample size was 69, as 13 of the 56 teachers took both the fall and spring survey and we counted both sets of responses. We organized the results by implementation domain, and focused on the items that refer specifically to QualityCore tools and resources or the QualityCore program.

Collaboration with Colleagues. The vast majority of teachers (90%) reported talking to colleagues about instructional strategies appropriate for specific QualityCore objectives (course standards). However, teachers only "slightly agreed," on average, that teachers in their department used a common QualityCore terminology around teaching, assessment, and student work. Only about one-half of the teachers reported that they used QualityCore's Worksheet to Examine Student Work to determine expectations for high-quality student work.

Educator Resources. On average, teachers "moderately agreed" that they used QualityCore Depth of Knowledge Analysis (to adapt tasks to the needs of students), know how to use most of the resources in QualityCore's Educator Toolbox, and know how to examine lessons for level of rigor and coherence with QualityCore course objectives. Teachers only "slightly agreed" that they used QualityCore's Template to Examine Assignments for Rigor & Relevance.

Formative Items. On average, teachers reported using constructed-response items from the FIP 2.8 days (out of 5 regular instructional days). Similarly, they reported using multiple choice items from the FIP 2.3 days. Teachers "slightly agreed" that they frequently used items from the FIP during the year, and 21% agreed that they have not yet used items from the FIP to construct classroom-based assessments. Teachers "slightly agreed," on average, that they found opportune times to incorporate items from the FIP in their classroom.

Teaching Practices. The majority of teachers (81%) agreed that they celebrated students' progress towards QualityCore's course standards by exhibiting student work in their classroom. The vast majority (95 to 98%) of teachers agreed with statements describing QualityCore-targeted teaching practices, including: following a process for modifying lesson plans after examining student work, being able to describe how scaffolding was related to different depth of knowledge levels, and being able to describe the importance of a course syllabus in a rigorous high school course.

QualityCore Adoption. Nearly all teachers (94%) reported that they had taken steps to design their curriculum so that it was aligned with the knowledge and skills necessary for college. On, average teachers slightly to moderately agreed (mean responses of 4.6) that they would recommend that other schools implement QualityCore courses and that they were excited about implementing QualityCore; teachers were slightly less likely to indicate that their colleagues were excited about implementing QualityCore (mean response of 4.2). Most teachers (89%) agreed that they had observed positive results from QualityCore, although the mean response (4.4) suggests that teachers tended to only "slightly agree." Most teachers reported that they had revised many of their lesson plans to intentionally include QualityCore (mean agreement response of 4.5) and that they often thought about how to incorporate the Model Unit into their instruction (mean of 4.6). About one quarter of the respondents agreed that they did not have the time to use QualityCore elements in their current work.

Other survey items related to *Adoption* suggested less QualityCore buy-in from teachers. Only 58% of the teachers agreed that they planned on using the EOC assessment results to set a baseline for future years. Only 50% agreed that they had discussed (with colleagues) sections that needed to be added to the course to meet QualityCore course standards. Just over half of the teachers agreed that QualityCore did not align very well with the state's performance standards; as discussed later this finding was confounded by state-mandated EOC assessments that were used for school accountability.

Discussion

Relationship of Implementation and Student Achievement

We examined the relationship between QualityCore Algebra I level of implementation and student achievement. Level of implementation was measured by teachers' self-reported utilization of QualityCore Algebra I resources and adherence to related practices. Achievement was measured by QualityCore Algebra I EOC scores controlling for prior EXPLORE scores. We found that, on average, an increase in level of implementation was significantly associated with increased performance on the QualityCore Algebra I EOC assessment.

The study results are consistent with prior studies that have suggested that raising the achievement level of students before they enter high school is likely to be more effective in improving college readiness than other interventions (Sawyer, 2008). The estimated effect of level of implementation was 0.13 (meaning that a "large" increase in level of implementation was related to a 0.13 standard deviation increase in EOC performance). To help put this estimate into perspective, a large (one standard deviation) increase in EXPLORE Mathematics score was related to a 0.38 standard deviation increase in EOC performance. Thus, the implementation effect size was about one-third the size of that for prior mathematics achievement. From this perspective, level of implementation appears to be an important contributor to student achievement.

However, we cannot make a strong argument from this study that greater levels of implementation caused greater achievement. There could be unobserved variables related to both greater implementation and greater student achievement. In addition, as discussed later, there are limitations to the study that further temper the argument of a causal effect of implementation on student achievement. Future research could strengthen the causal argument. For example, the QualityCore Algebra I EOC could be administered to students whose teachers did not participate in the study. A comparison of EOC scores with the scores of the students in this study, adjusted for prior achievement and other background characteristics, would be a stronger test of the QualityCore implementation effect.

We also considered the fall and spring implementation measures separately. Although the aggregate fall and spring implementation was significantly predictive of student achievement, we found that only the fall implementation measure was significant in the term-specific models. One possible explanation for this result is that the fall survey included more items (49) than the spring survey (21) (see Appendix A), and thus should have been a more reliable measure of implementation. If, in fact, a relationship truly exists between level of QualityCore implementation and student achievement, we would be less likely to observe a statistical relationship with a less reliable measure of implementation. Another possible explanation is the differences in fall and spring survey content. The fall survey included more items that were not directly related to the QualityCore course (and less likely to be attributed to QualityCore training and implementation) but likely connected to student achievement. Examples of such items included “I talk with teachers about student progress in grade levels other than those that I teach,” “Teachers in our department do not collaborate on instructional planning,” and “I have a hard time getting my students to take responsibility for their own learning.” It is possible that the fall survey better measured constructs related to student achievement, resulting in a significant effect on EOC student achievement of the fall and aggregated QualityCore implementation.

Implementation was also measured for five domains, defined to measure different aspects of the tools, strategies, or practices of QualityCore Algebra I. A goal of the domain classifications was to be able to assess how each component contributed in raising the quality and intensity of the Algebra I course. While all domains had a positive relationship with student achievement, only *Teaching Practices* and *Collaboration with Colleagues* were statistically significant at the 0.05 significance level. The survey content for these two domains included more items that were not directly related to QualityCore resources, but perhaps more related to general teaching practices and positive school atmosphere (e.g., collaborative lesson planning) that were supported by QualityCore training and resources. It is possible that these general factors were more important than greater use of the specific resources offered by QualityCore (e.g., greater use of the FIP).

QualityCore Use and Perceptions

The survey results shed light on the extent that QualityCore resources were used, as well as teachers’ perceptions of QualityCore. For example, collaborative review of student work is a practice that is covered in QualityCore training and supported with QualityCore resources, but—

relative to other practices—was not reported as much by teachers. On average, teachers only slightly agreed that teachers in their department believed that instruction was improved by collaboratively looking at student work. Only 61% agreed that they and their colleagues came to a consensus as to what constituted high-quality student work, and only 54% agreed that they and their colleagues used the QualityCore Worksheet to Examine Student Work.

In interpreting the survey results, it is important to keep in mind the timing of the survey (2009-2010 school year) and the standards, testing, and accountability systems in place at that time. In all three large school districts that were part of the study, 9th grade students were required to take the state's EOC assessment in 9th grade Mathematics. In at least one district, the state EOC assessment accounted for a significant portion of students' course grade. In all districts, the state EOC assessment was used for school accountability. Thus, the QualityCore EOC assessment was likely perceived as less important (by teachers and students) than the state EOC assessment.

Also, 56% of the teachers agreed that QualityCore did not align very well with the state's performance standards. In one of the study districts, the 9th grade math course was not called "Algebra I," but resembled an integrated Mathematics course instead. In that district, the course standards did not match well with those of traditional Algebra I courses. Accordingly, 67% of the teachers agreed that QualityCore Algebra I did not align very well with the state's performance standards. In comparison, in the other two districts that did have an Algebra I course for 9th graders, only 22% of the teachers agreed that QualityCore Algebra I did not align very well with the state's performance standards. Because most states have since adopted the Common Core State Standards and because QualityCore's course standards align well with the

Common Core State Standards (ACT, 2010b), it is likely that current and future studies of QualityCore implementation would suggest stronger alignment.

Study Limitations

The study has some important limitations that should be acknowledged. First, the number of items used for the fall survey greatly outnumbered those used in the spring, with only 13 items used for both the fall and spring surveys. The two surveys differed in both content and reliability, and cannot be considered repeated measurements of the same implementation construct. Ideally, the same survey would have been administered during both terms so that reliability would be more uniform and better assessed. Moreover, using the same survey in both terms would have provided a better way to measure change in implementation across the year.

Another limitation is that more teachers were surveyed in the spring than in the fall. A more accurate measure of the relationship between implementation and achievement could have been achieved with the same teachers surveyed in fall and spring. For future studies, we would recommend a unified and reconstructed questionnaire, building off of the items provided in Appendix A. Only 41 teachers were included in the primary analyses. Because the effect of any one-year program or intervention (including QualityCore Algebra I) on student achievement is likely to be small, a larger sample size of teachers and students is needed in order to draw more conclusive results and to have adequate power to detect effects of different aspects of implementation.

As noted above, the study was conducted in three large school districts with high percentages of lower-achieving students, many living in poverty. The adoption of QualityCore was funded by school improvement grants, and was part of a comprehensive strategy for improvement in the three districts. The use of QualityCore, and the effect of QualityCore on student achievement in Algebra I, could both be influenced by these contextual factors, and limit the generalization of the findings. The vast majority of students (85%, see Table 3) in the sample were African American, which could also affect the generalization of the findings.

The measurement of implementation relied on teachers' self-report. With self-report, social desirability bias is a concern; teachers have been shown to over-report positive changes in classroom practices, which is attributed to a desire to appear favorably in relation to one's peers (Kopcha & Sullivan, 2006). Teachers were told in the survey instructions that their responses would remain anonymous, which can help reduce this bias (Scaeffer, 2000). Still, it is possible that teachers over-reported their use of QualityCore resources and teaching practices, and this may have weakened the observed relationship between level of implementation and student achievement.

In addition to the reliance on self-reported levels of implementation, there is also a concern that survey non-response could have affected the results. We estimate that 37% of the teachers implementing QualityCore Algebra I participated in one or more surveys; it is possible that the non-participants had different levels of implementation and that the relationship between implementation and student achievement was different for the non-participants. If the relationship between level of implementation and student achievement was weaker for the non-participants, the survey non-response bias would result in this study reporting an inflated relationship between level of implementation and student achievement.

The study assessed the relationship of level of QualityCore Algebra I implementation and student achievement. QualityCore implementation was conceptualized as a "treatment" measured on a scale; the study did not, however, directly test whether QualityCore adoption was related to improved student achievement (e.g., QualityCore adoption versus no use of QualityCore). The study found evidence of a positive relationship between level of implementation and student achievement, but future studies should examine the impact of QualityCore adoption on students' college and career readiness. A recent large-scale study has documented positive improvements in college readiness associated with QualityCore adoption, with larger improvements observed for schools that participated in professional development on formative instructional practices (Battelle for Kids, 2012). Future research should continue to examine the effects of QualityCore adoption and level of implementation on school-wide achievement and growth.

References

ACT. (2005). *Courses Count: Preparing Students for Postsecondary Success*. Iowa City, IA: Author.

ACT. (2007). EXPLORE Technical Manual. Iowa City, IA: Author.

ACT. (2010a). QualityCore Technical Manual. Iowa City, IA: Author.

ACT. (2010b). The Alignment of Common Core and ACT's College and Career Readiness System. Iowa City Iowa: Author.

ACT. (2011). QualityCore Professional Development. Retrieved from http://www.act.org/qualitycore/development.html

ACT & The Education Trust. (2005). *On Course for Success: A close look at selected high school courses that prepare all students for college*. Iowa City, IA: Authors.

Angrist, J., & Lang, K. (2002). How Important Are Classroom Peer Effects? Evidence from Boston's Metco Program, NBER Working Paper 9263.

Battelle for Kids. (2012). *Ohio Value-Added Schools*. Viewed June 6, 2012 at http://www.battelleforkids.org/Impact/Value_Added_Initiatives/High_School_VA.html.

Goldstein, H. (2002). Multilevel statistical models (3rd ed.). New York: Wiley.

Hedges, V. L., & Rhoads, C. H. (2011). Correcting an analysis of variance for clustering. *British Journal of Mathematical and Statistical Psychology*, 64, 20–37.

Koenig, J. A., & Lissitz, R. W. (2001). The Effects of Centering Methods on Parameter Estimates and Variance Explained in Multilevel Models. *Reports-Evaluative, 142*.

Kopcha, T. J., & Sullivan, H. (2006). Self-presentation bias in surveys of teachers' educational technology practices. *Education Technology Research and Development*, 55, 627-646.

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models. Thousand Oaks, CA: Sage Publications.

Sawyer, R. (2008). Benefits of Additional High School Course Work and Improved Course Performance in Preparing Students for College, *ACT Research Report Series*, 2008-1.

Scaeffer, N. C. (2000). Asking questions about threatening topics: A selective overview. In A. Stone, J. Turkkan, J. Jobe, C. Bachrach, H. Kurtzman, & V. Cain (Eds.), *The science of self report: Implications for research and practice*. Mahwah, NJ: Erlbaum.

Singer, D. J. (1998). Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Achievement Models. *Journal of Educational and Behavioral Statistics*, Vol. 23, No. 4. (Winter, 1998), pp. 323-355.

Snidjers, T. A. B., & Bosker, R. J. (1999). Multilevel modeling: An introduction to basic and advanced multilevel modeling. London: Sage Publications.

Appendix A

Survey Items and Statistics

Collaboration with Colleagues						
Over the past 5 regular instructional days, how often have you performed each of the following activities? 0=No time, 1=1 hour, 2=2 hours, 3=3 hours, 4=4 hours, 5=5 hours, 8=6-10 hours, 15=11 hours or more						
Item	**Fall**	**Spring**	**N**	**% $\geq$ 3**	**M**	**SD**
Collaboratively "Looking at Student Work"	X	X	62	23	1.4	1.7
Collaboratively planning lessons	X	X	62	27	1.9	1.9
Please rate the level of your agreement with each of the following statements. 1=Strongly Disagree, 2=Moderately Disagree, 3=Slightly Disagree, 4=Slightly Agree, 5=Moderately Agree, 6=Strongly Agree, Not Applicable						
Item	**Fall**	**Spring**	**N**	**%Agree**	**M**	**SD**
My colleagues and I talk about instructional strategies that would be appropriate for teaching specific QualityCore objectives.	X		20	90	4.9	1.2
I talk with teachers about student progress in grade levels other than those that I teach.	X		20	80	4.6	1.5
My colleagues and I have trouble deciding on ways to change our instruction based on collaborative review of student work. **[R]**	X		18	39	2.9	1.6
I have discussed my course syllabus with other teachers who teach the same course.	X		20	95	5.5	0.9
In my department, teachers believe that instruction for all students is improved by collaboratively looking at student work.	X		20	70	4.1	1.4
I rarely talk to teachers in departments other than Math about course objectives and students progress. **[R]**	X		20	30	2.7	1.9
Teachers in my department use a common QualityCore terminology to describe teaching, assessment, and student work.	X		19	68	3.9	1.7

Collaboration with Colleagues						
Please rate the level of your agreement with each of the following statements. 1=Strongly Disagree, 2=Moderately Disagree, 3=Slightly Disagree, 4=Slightly Agree, 5=Moderately Agree, 6=Strongly Agree, Not Applicable						
My colleagues and I have a scheduled time when we meet to plan instruction and analyze student work.	X		20	75	4.5	1.6
Teachers in our department do not collaborate on instructional planning. **[R]**	X		20	20	2.2	1.7
Item	**Fall**	**Spring**	**N**	**%Agree**	**M**	**SD**
My colleagues have asked me about the results I have observed using new practices I learned in QualityCore training.	X		19	58	3.7	1.8
At faculty meetings, we regularly discuss student successes.	X		20	75	4.3	1.6
My colleagues and I do not follow a process when examining student work. **[R]**	X		19	32	2.7	1.6
My colleagues and I come to a consensus as to what constitutes high-quality student work.	X		18	61	4.1	1.6
Teachers in my department are not given much time to collaboratively plan. **[R]**	X		19	53	3.4	1.9
At my school, experienced faculty mentor to less experienced faculty.	X		19	89	4.7	1.3
My colleagues and I use a QualityCore Worksheet to Examine Student Work and determine expectations for high-quality student work.	X	X	56	54	3.5	1.7
Number of items	*18*	*3*				

R indicates reverse-score item.

Educator Resources						
Please rate the level of your agreement with each of the following statements. 1=Strongly Disagree, 2=Moderately Disagree, 3=Slightly Disagree, 4=Slightly Agree, 5=Moderately Agree, 6=Strongly Agree, Not Applicable						
Item	**Fall**	**Spring**	**N**	**%Agree**	**M**	**SD**
I use QualityCore Depth of Knowledge analysis to adapt tasks to the needs of my students.	X	X	59	81	4.6	1.4
I use QualityCore's Template to Examine Assignments for Rigor and Relevance to help me modify assignments.	X	X	59	63	3.7	1.8
I know how to use most of the resources in the QualityCore Educator's Toolbox.		X	39	77	4.8	1.4
I know how to examine lessons for level of rigor and coherence with QualityCore Course Objectives.		X	40	88	4.9	1.2
Over the past 5 regular instructional days, how often have you used the following teaching strategies from QualityCore's Educator's Toolbox? 5=Daily (5 days), 4=Frequently (4 days), 3=Often (3 days), 2=Sometimes (2 days), 1=Seldom (1 day), 0=Never (0 days), NA						
Item	**Fall**	**Spring**	**N**	**% ≥ 3**	**M**	**SD**
Formative Assessment (e.g., Index Cards, K-W-L, Clickers, 3-2-1 Assessment, Traffic Light Icons)	*X		22	45	2.4	1.6
Problem-Based Inquiry (e.g., Systems Analysis, Essential Questions, Question Wall)	*X		22	73	3.4	1.9
Collaborative Discourse (e.g., Whiteboard Activities, Dialogic Journals, Group Work, Four Corners)	*X		23	74	3.3	1.5
Assessing Prior Knowledge (e.g., Background Knowledge Probe, Entrance and Exit Slips)	*X		23	70	3.3	1.7
Practice, Review, and Revision (e.g., Focused Listing, Three-Minute Review, Advance Organizer)	*X		23	70	3.5	1.6
Sense-Making and Wrap-up (e.g., Muddiest Point, Minute Paper, Journals and Learning Logs)	*X		23	57	2.7	1.6
Number of items	**3*	*4*				

*Note that the items related to frequency of use of resources from QualityCore's Educator's Toolbox were treated as a single item using the sum of the items (after standardization).

Formative Items						
Over the past 5 days regular instructional days, how much time have you spent on the following activities? 5=Daily (5 days), 4=Frequently (4 days), 3=Often (3 days), 2=Sometimes (2 days), 1=Seldom (1 day), 0=Never (0 days), Not Applicable						
Item	**Fall**	**Spring**	**N**	**% $\geq$ 3**	**M**	**SD**
I used constructed-response items from the QualityCore formative item pool.	X	X	63	57	2.8	1.6
I used multiple choice items from the QualityCore formative item pool.	X	X	63	43	2.3	1.3
Please rate the level of your agreement with each of the following statements. 1=Strongly Disagree, 2=Moderately Disagree, 3=Slightly Disagree, 4=Slightly Agree, 5=Moderately Agree, 6=Strongly Agree, Not Applicable						
Item	**Fall**	**Spring**	**N**	**%Agree**	**M**	**SD**
I have not yet used items from the QualityCore Formative Item Pool to construct classroom-based assessments. **[R]**	X		19	21	2.8	1.7
This year, I frequently used items from the QualityCore Formative Pool.		X	40	63	3.8	1.8
I find opportune times to incorporate QualityCore formative assessment items in my classroom.		X	39	77	4.1	1.7
Number of items	*3*	*4*				

R indicates reverse-score item.

Teaching Practices						
Over the past 5 regular instructional days, how much time have you spent on the following activities? 5=Daily (5 days), 4=Frequently (4 days), 3=Often (3 days), 2=Sometimes (2 days), 1=Seldom (1 day), 0=Never (0 days), Not Applicable						
Item	**Fall**	**Spring**	**N**	**% $\geq$ 3**	**M**	**SD**
I provided my students with a warm-up (or hook) to review previous lessons or to anticipate a new lesson.	X	X	64	95	4.6	0.9
I provided rationale to my students regarding the focus and purpose of the day's lesson.	X	X	64	95	4.4	1
I used scaffolding during the class period.	X		23	87	3.9	1.1
I concluded my lesson with a summary.	X		23	87	3.9	1
I instructed my students to summarize their learning with a wrap-up activity (closing, exit slip, journal, etc.)	X	X	64	70	3.4	1.4
Over the past 5 regular instructional days, how much time have you spent on the following activities? 0=No time, 1=1 hour, 2=2 hours, 3=3 hours, 4=4 hours, 5=5 hours, 8=6-10 hours, 15=11 hours or more						
Item	**Fall**	**Spring**	**N**	**% $\geq$ 3**	**M**	**SD**
Independently grading assignments/tests and "Looking at Student Work"	X		22	68	3.9	2.2
Independently planning lessons	X	X	63	63	4.0	3.2
Please rate the level of your agreement with each of the following statements. 1=Strongly Disagree, 2=Moderately Disagree, 3=Slightly Disagree, 4=Slightly Agree, 5=Moderately Agree, 6=Strongly Agree, Not Applicable						
Item	**Fall**	**Spring**	**N**	**%Agree**	**M**	**SD**
I celebrate students' progress towards the QualityCore course standards by exhibiting student work in my classroom.	X		21	81	4.8	1.6
My course syllabus does not include a personal statement. **[R]**	X		17	35	2.6	2
I often adapt assignments to ensure that they have sufficient rigor and relevance .	X		21	95	5.4	0.9

Teaching Practices						
Please rate the level of your agreement with each of the following statements. 1=Strongly Disagree, 2=Moderately Disagree, 3=Slightly Disagree, 4=Slightly Agree, 5=Moderately Agree, 6=Strongly Agree, Not Applicable						
When I give an assessment, I can describe exactly how it measures progress towards meeting course objectives.	X		21	86	4.9	1.2
I have a hard time getting my students to take responsibility for their own learning. **[R]**	X		21	86	4.8	1.2
My students don't really understand the purpose behind each day's lesson. **[R]**	X		21	38	2.8	1.6
I often give feedback in the form of sticky-notes to a student with the expectation that the student will revise their work.	X		20	70	4.0	1.4
When examining student work, I record comments for individual students.	X		21	90	5.0	1.3
I follow a process for modifying lesson plans after examining student work.		X	40	95	5.3	1.1
I can describe how scaffolding is related to different levels of Depth of Knowledge.		X	40	98	5.3	1
I can describe the importance of a course syllabus in a rigorous high school course		X	40	95	5.3	1.1
Number of items	*15*	*7*				

R indicates reverse-score item.

QualityCore Adoption						
Please rate the level of your agreement with each of the following statements. 1=Strongly Disagree, 2=Moderately Disagree, 3=Slightly Disagree, 4=Slightly Agree, 5=Moderately Agree, 6=Strongly Agree, Not Applicable						
Item	**Fall**	**Spring**	**N**	**%Agree**	**M**	**SD**
I often think about how I can incorporate features of the QualityCore Model Instructional Unit into my instruction.	X		20	80	4.6	1.7
I have revised many of my existing lesson plans to intentionally include QualityCore resources into my teaching.	X	X	60	78	4.5	1.4
I plan on using this year's QualityCore end-of-course exam results to set a baseline for future years' student achievement.	X	X	52	58	3.7	2
I don't have the time right now to use QualityCore elements in my work. **[R]**	X		19	26	2.5	1.5
I have taken steps to design my curriculum so that it is aligned with the knowledge and skills necessary for college.	X		18	94	5.2	0.9
My colleagues and I have discussed sections that need to be added to the course to meet the standards of QualityCore objectives.	X		18	50	3.7	1.8
I have observed positive results from implementing QualityCore in my classroom.	X		18	89	4.4	1.2
My colleagues are excited about implementing QualityCore in the classroom.	X		20	75	4.2	1.4
I would recommend that other schools implement QualityCore courses.	X		18	94	4.6	1.1
I am excited about implementing QualityCore in the classroom.	X		20	85	4.6	1.4
QualityCore doesn't align very well with our state's performance standards. **[R]**		X	39	56	3.9	1.9
Number of items	*10*	*3*				

R indicates reverse-score item.

Survey Totals						
Item	**Fall**	**Spring**	**N**	**%Agree**	**M**	**SD**
Number of items	*49*	*21*				

Appendix B

Results of Hierarchical Linear Models

Table B-1

Partial model results: Collaboration with Colleagues

Effect	Estimate	SE	P-value	Effect Size
Intercept	142.91*	0.19	<0.0001	
Mean EXPLORE Math	0.02	0.12	0.8357	0.01
EXPLORE Math	0.21*	0.03	<0.0001	0.21*
EXPLORE English	0.12*	0.04	0.0012	0.11*
EXPLORE Reading	0.05	0.04	0.1595	0.05
EXPLORE Science	0.15*	0.04	0.0002	0.12*
Collaboration with Colleagues	0.46*	0.21	0.0375	0.11*
Variance of intercepts	0.92*	0.35	0.0043	
Residual variance	9.60*	0.39	<0.0001	

*p-value < 0.05

Table B-2

Partial model results: Educator Resources

Effect	Estimate	SE	P-value	Effect Size
Intercept	142.83*	0.19	<0.0001	
Mean EXPLORE Math	0.06	0.12	0.6412	0.02
EXPLORE Math	0.21*	0.03	<0.0001	0.21*
EXPLORE English	0.10*	0.04	0.0009	0.12*
EXPLORE Reading	0.05	0.04	0.1544	0.05
EXPLORE Science	0.15*	0.04	0.0001	0.12*
Educator Resources	0.30	0.18	0.1191	0.09
Variance of intercepts	1.04*	0.38	0.0028	
Residual variance	9.59*	0.39	<0.0001	

*p-value < 0.05

Table B-3

Partial model results: Formative Items

Effect	Estimate	SE	P-value	Effect Size
Intercept	142.84*	0.19	<0.0001	
Mean EXPLORE Math	0.05	0.12	0.7051	0.02
EXPLORE Math	0.21*	0.03	<0.0001	0.21*
EXPLORE English	0.12*	0.04	0.0011	0.11*
EXPLORE Reading	0.05*	0.04	0.1471	0.05
EXPLORE Science	0.15*	0.04	0.0001	0.12*
Formative Items	0.26	0.21	0.2153	0.07
Variance of intercepts	1.09*	0.39	0.0023	
Residual variance	9.59*	0.39	<0.0001	

*p-value < 0.05

Table B-4

Partial model results: Teaching Practices

Effect	Estimate	SE	P-value	Effect Size
Intercept	142.85*	0.18	<0.0001	
Mean EXPLORE Math	0.10	0.11	0.3852	0.04
EXPLORE Math	0.21*	0.03	<0.0001	0.21
EXPLORE English	0.12*	0.04	0.0007	0.12
EXPLORE Reading	0.05	0.04	0.1510	0.05
EXPLORE Science	0.15*	0.04	0.0002	0.12
Teaching Practices	0.52*	0.18	0.0057	0.15
Variance of intercepts	0.86*	0.32	0.0037	
Residual variance	9.58*	0.39	<0.0001	

*p-value < 0.05

Table B-5

Partial model results: QualityCore Adoption

Effect	Estimate	SE	P-value	Effect Size
Intercept	142.81*	0.20		-0.04
Mean EXPLORE Math	0.06	0.12	0.6489	0.02
EXPLORE Math	0.21*	0.03	<0.0001	0.21
EXPLORE English	0.12*	0.04	0.00010	0.12
EXPLORE Reading	0.05*	0.04	0.1313	0.05
EXPLORE Science	0.15*	0.04	0.0001	0.12
QualityCore Adoption	0.22	0.25	0.3808	0.05
Variance of intercepts	1.10*	0.39		0.09*
Residual variance	9.59*	0.39		0.77*

*p-value < 0.05