The Relationship between Lesson Progress in Direct Instruction Programs and Student Test Performance

Technical Report 2014-1

Jean Stockard
Office of Research and Evaluation
NATIONAL INSTITUTE FOR DIRECT INSTRUCTION
JULY 17, 2014

Contents

List of Figures

List of Tables

The Relationship between Lesson Progress in Direct Instruction Programs and Student Test Performance[1]

Executive Summary

As the so-called "high stakes" tests have proliferated, students' performance on state developed and other forms of assessment has become a major concern of school officials. Because the results of these tests often have serious financial and political implications, schools may devote many hours to "test preparation," hoping that this will increase students' scores. Yet, policy makers and parents have expressed concern over this practice, suggesting that it can result in invalid measures of student achievement, divert attention from crucial subject matter, and provide less than optimal models for students. Clearly, school officials face a dilemma. They are under intense political and financial pressure to have high test scores. Yet, they also wish to promote their students' future academic success and provide appropriate behavioral models.

This report presents data that support an alternative to extensive test preparation as the means to higher test scores. The alternative is systematic progress, at mastery, through the Direct Instruction (DI) curricula, *Reading Mastery Signature Edition (RMSE)* and *Connecting Math Concepts: Comprehensive Edition (CMCCE)*. The analysis shows that progress at mastery through these curricula results in students having scores on state assessments and a nationally normed achievement test that are significantly higher than what their earlier scores would predict. Moreover, the value added by progress through the curriculum is markedly higher than the average effect reported for test preparation programs.

Methods

Most curricular programs use a "spiral approach," briefly visiting and revisiting a series of topics. In contrast, Direct Instruction programs incorporate mastery learning, sometimes described as analogous to a stair-stepped progression (Engelmann, 2014b). Each step, or lesson, in the program builds on previous learning, and the programs include explicit instructions for determining students' mastery and appropriate placement. This report examines the extent to which students' cumulative progress through the programs is related to their scores on measures of achievement that are external to the DI programs: two different state assessments (the Texas Assessments of Academic Readiness [STAAR] and the Transitional Colorado Assessment Program [TCAP]) and the Northwest Evaluation

[1] The author gratefully acknowledges the helpful comments of Shep Barbash, Carrie Beck, Christina Cox, Kurt Engelmann, and Caitlin Rasplica on earlier drafts of this report. All conclusions and opinions in this document are, however, the sole responsibility of the author.

Association's, Measures of Academic Progress [MAP], for which national norms are available).

Data from the 2012-2013 school year were obtained from charter schools in Texas, Colorado, and an east coast inner city and involved a range of grades (K to 5 in Texas, grade 3 in Colorado, and grades 1-4 for the inner city school). There was substantial variation both within and between schools in students' progress through the curriculum and students' demographic characteristics of students, including both race-ethnicity and levels of poverty. All of the schools received technical support from the National Institute for Direct Instruction and had strong procedures to help ensure that students were placed at points within the DI programs where they were at mastery.[2]

Findings

Results indicate that students who were at or approaching mastery at their assigned grade level in *RMSE* and *CMCCE* had a strong probability of scoring at the proficient level on state assessments and were much more likely than other students to score above the national mean on the MAP. Far fewer students who were at lower levels in the program were found to be proficient. (See Figure A.) All of the comparisons were statistically significant. The associated effect sizes were large and generally several times the level used to denote educationally important effects (averaging 1.12 across all comparisons for reading and .95 for mathematics). Similar results appeared with the MAP data. (See Figure B.) Almost all of the differences were statistically significant, even though some of the sample sizes were relatively small; and the associated effect sizes were substantial (averaging .93 for reading and .83 for math). When students have mastered the curriculum for their grade (or, in several analyses, only the material partly through their grade level), they have a very high probability of passing their state assessment and are much more likely than other students to score above the national mean on standardized tests. The author knows of no documentation in the literature of another curricular program having such a strong association of progress through the program and assessment scores.

Most importantly, the strong relationship of lesson progress to achievement scores persisted when students' prior achievement was controlled. Students' previous achievement measures are usually a very strong predictor of later scores on achievement tests. Yet, the analysis reported in this paper found that students' cumulative progress through *RMSE* and CMCCE provided significant "value added" to students' earlier achievement scores. The average effect sizes associated with the value added for each site, assessment, and subject are shown in Figure C. All are positive and all but one is greater than .20. Most of the individual estimates of value added were statistically significant, even with relatively small

[2] It is, of course, possible that some students were not truly at mastery for their designated point in the curriculum. However, to the extent that this occurred, the estimates of the effects of RMSE and CMCCE on students' test scores are conservative in nature.

samples and strong controls. In other words, *students' progress in their DI programs added significantly to their achievement scores beyond what they would be expected to score given their earlier performance on the assessment*. The findings are especially noteworthy given the replication of results across three different sites and several grade levels.

Implications for Policy and Practice

The author has not been able to locate studies that document the effect size associated with test preparation programs for students in the elementary grades. Studies have, however, documented the effect of test preparation on SAT and PSAT scores of high school students. Summaries of this work report only a small impact on students' scores, with effect sizes ranging from .06 to .07 for the verbal tests and .11 to .20 for mathematics (see Briggs, 2001; Dominguez and Briggs, 2009). These effects are substantially smaller than those associated with the value added by progress through *RMSE* and *CMCCE.*

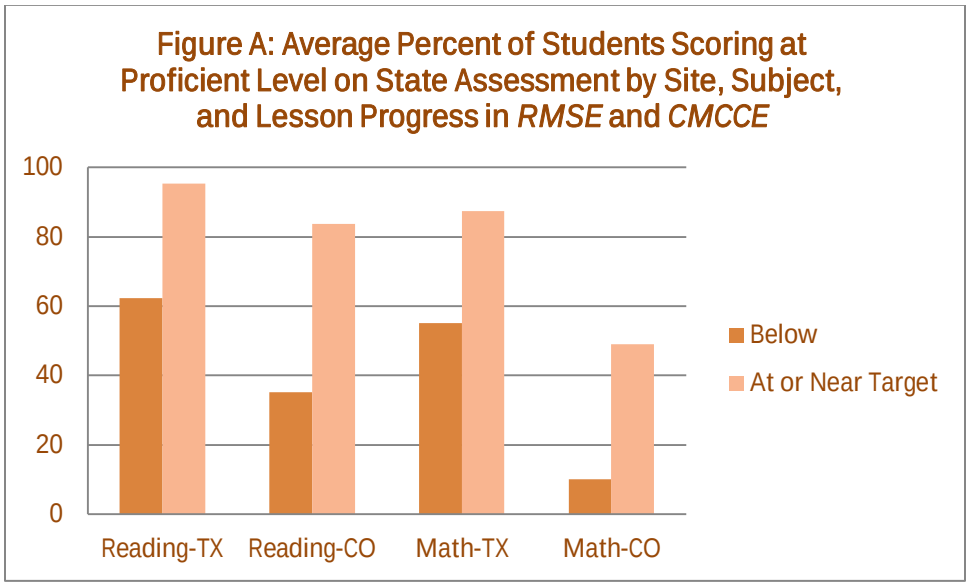
The author suggests that some elements of test preparation strategies, such as helping students understand technical details of examination formats, might be appropriate adjuncts to study of the curriculum, especially for students whose performance on standardized assessments does not conform to their placement in their curriculum or their performance on day-to-day classroom assessments. However, given the strong association of lesson placement and assessment scores described in this document, these students are, undoubtedly, a small minority.

The discussion specifically cautions against using elaborate test preparation as a way to raise test scores of students who are substantially behind their peers. The results of this paper suggest that such an approach would be relatively ineffective and could result in students being even further behind. The DI curriculum includes specific, highly successful, strategies to help students catch up to grade level and use of class time in the curriculum would no doubt be more effective in the long run in raising students' scores.

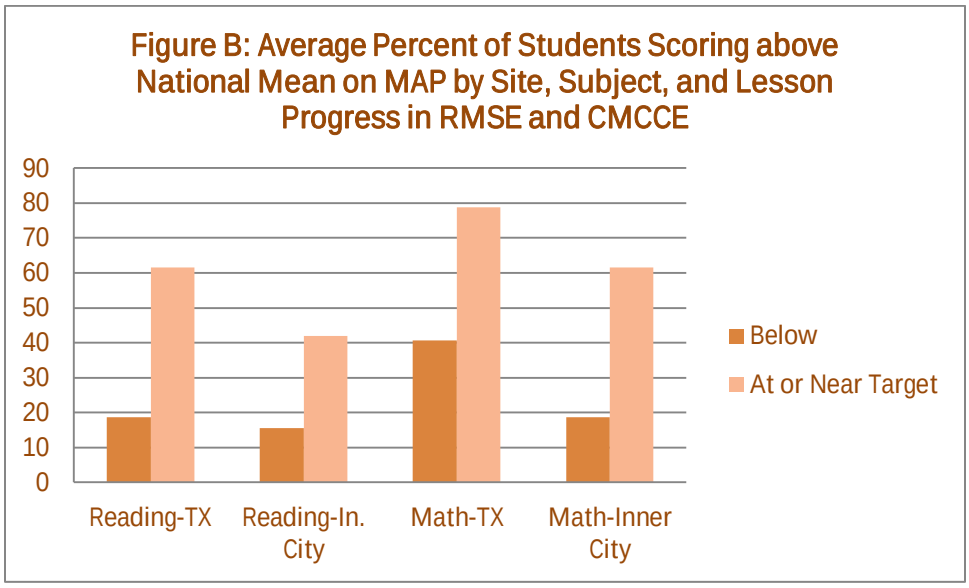
The report also emphasizes that the conclusions are based on students' progress through *RMSE* and *CMCCE* at mastery, when they have thoroughly learned the material, as demonstrated by the tests included within the programs. Simply exposing students to the DI curriculum, without ensuring that students have learned the content, would, most likely not produce the results described within this report.

Taken together, the results described in this document suggest that an effective alternative to extensive test preparation could simply be to ensure that students make adequate progress, at mastery, through *Reading Mastery* and *Connecting Math Concepts*. In contrast to the learning that occurs with test preparation, students engaged in the subject matter through these programs build a broad array of content knowledge that prepares them for

continued academic success as well as being more likely to pass state assessments and have higher scores on national exams.



Note: The values in the figure represent the average percentage across the grades within each of the sites. As explained in the body of the text above, to accommodate varying rates of lesson progress, the definition of "at or near target" varied across sites and grades. For Texas it was defined as the last lesson for a given grade for grades K-4 for reading and for grades K-3 for mathematics. For grade 5 in reading and grade 4 in mathematics, it was defined as being 90% through the grade level material. For grade 5 in mathematics it was defined as being 60% through the grade level material. For the Colorado school "at or near target" was defined as being at any point in grade 3 material for both reading and mathematics.



Note: The values in the figure represent the average percentage across the grades within each of the sites. For Texas the definition of "at or near target" was the last lesson for a given grade for grades K-4 for reading and for grades K-3 for mathematics. For grade 5 in reading and grade 4 in mathematics, it was defined as being 90% through the grade level material. For grade 5 in mathematics it was defined as being 60% through the grade level material. For the inner city school "at or near target was defined as being at the last lesson of the grade level for all grades for reading. For mathematics it was defined as being at the last lesson of the grade level or beyond for first grade, at the mid-point of the grade level lessons for second and third graders and at any place within the grade level material for fourth grade.

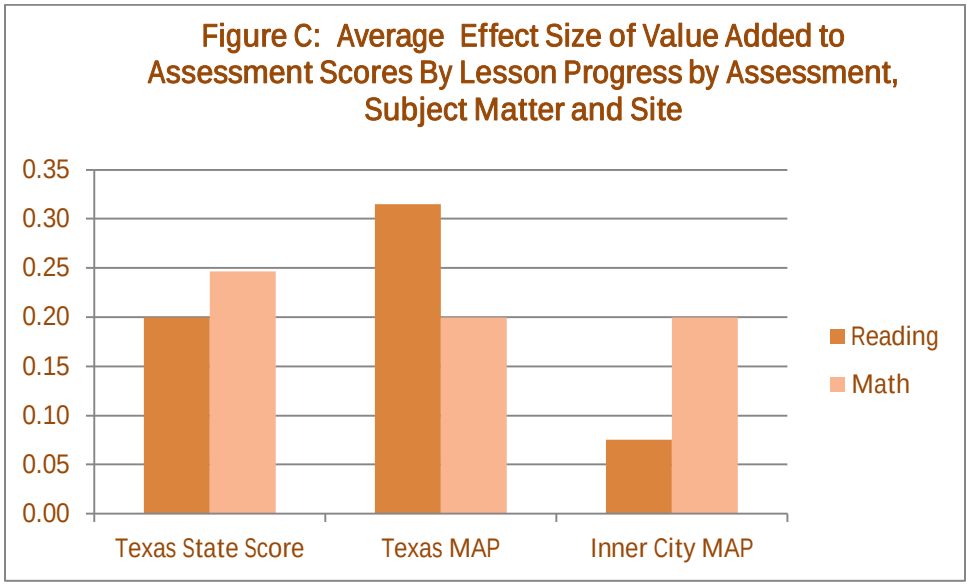


Figure C: Average Effect Size of Value Added to Assessment Scores By Lesson Progress by Assessment, Subject Matter and Site
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Texas State Score
Texas MAP
Inner City MAP
Reading
Math

The Relationship between Lesson Progress in Direct Instruction Programs and Student Test Performance

Students' performance on state developed and other types of assessments is a major concern of school officials. The so-called "high stakes" tests that have become increasingly common in recent years have focused the attention of administrators, teachers, and parents on students' performance. In preparing students for these exams schools may devote many hours to "test preparation," diverting time from regular classroom instruction to provide skills that are thought to help students do better on an exam.

Policy makers and parents have expressed concern over this attention to test taking skills and preparation for specific exams. On the policy side, such extensive test preparation diminishes the probability that the assessment results give a valid measure of students' achievement. Put quite simply, teaching to the test can deprive parents and the public of a true measure of their students' learning. From a parents' perspective, the reallocation of instructional time decreases the attention given to substantive instruction. Parents, and the public at large, may fear that students are learning test taking skills and specific items that are likely to be tested at the expense of learning the full range of subject matter in the curriculum. In addition, they may worry about the ethical and moral message that is imparted with extensive test preparation, as school officials actively promote a practice that some clearly regard as cheating (e.g. Engelmann, 2014a, p. 20).

Obviously, school officials face a dilemma. They are under intense political and financial pressure to have high test scores. Yet, they also wish to promote their students' learning and to provide appropriate ethical models. This report examines data that support an alternative to extensive test preparation as the means to higher test scores – progress in the mastery learning based Direct Instruction programs.

A substantial literature has documented the efficacy of Direct Instruction programs, concluding that students using these curricula have higher levels of achievement than those in other programs. (See Adams & Engelmann, 1996; Borman, Hewes, Overman, & Brown, 2003; Coughlin, 2014; Liem & Martin, 2013; Hattie, 2009; Przychodzin, Marchand-Martella, Martella, & Azim, 2004; and Schieffer, Marchand-Martella, Martella, Simonsen, & Waldron-Soler, 2002 for meta-analyses and summaries of this literature.) In contrast to approaches that use a "spiral" approach to curriculum design, all DI programs incorporate mastery learning. They are structured in a cumulative manner, so that each new lesson to which a

student is exposed builds on previous knowledge. Previous learning is systematically reinforced and frequent in-program assessments help teachers know if their students have learned the material (Engelmann, 2014b). This cumulative nature of the curriculum and the tests of student mastery are considered key reasons that the programs are so effective and efficient. The programs are extensively tested during their development to ensure that, when properly presented, all children can learn the material.

Given the cumulative nature of the DI curriculum and the assessments of students' mastery, it is reasonable to suggest that *students who have progressed further in the programs, and thus mastered the curricular material, would have higher levels of achievement and be more likely to pass state assessments and exceed national norms on standardized achievement tests*.

The analysis below supports this hypothesis. Data from three different sites show a strong relationship between students' test scores and their progression through two Direct Instruction programs – *Reading Mastery Signature Edition (RMSE)* and *Connecting Math Concepts: Comprehensive Edition (CMCCE)*. Students who were on or close to grade level in these programs had a strong probability of passing the state assessments and were much more likely than other students to score higher than the national mean on a normed achievement test. Those who were behind grade level in the programs had substantially lower scores and were unlikely to pass the state assessments. These results continued to appear when strong controls were included for prior levels of achievement, indicating that progress in the programs during the academic year, and not prior learning, produced significant "value added" to students' tests scores.

Before providing details of the results, it is important to emphasize that the association of progress through the programs and test scores is based on the assumption that students have mastered, or thoroughly learned, the material. The programs include extensive instructions for testing such mastery and ensuring that students are not placed in material that is either too difficult or too easy. Each of the schools in the analysis had extensive procedures to ensure that students were appropriately placed and had mastered the material. To the extent that students' mastery of the material was less than perfect, the estimates provided in this analysis are conservative in nature, potentially underestimating the actual correlation between students' progress in the curriculum at mastery and their scores on the assessments. However, simply exposing students to the DI curriculum, without ensuring that students had mastered the content, would, most likely, not reproduce the results reported here. In other words, replication of the strong association of progress through the curriculum and test scores depends upon appropriate implementation of the

programs and, specifically, ensuring that students learn the material at mastery. (See Engelmann, 2014b, for an extensive discussion of the importance of mastery learning.)

The remainder of this report provides details on the findings. The first section describes the methodology that was used. Succeeding sections describe the relationship between students' progress through the programs and their achievement scores, focusing first on reading and then on mathematics. The final section summarizes the results and discusses the implications for schools that want to prepare their students for success on examinations. Figures are used to summarize the findings in the main body of the report. An appendix includes extensive supporting tables.

Methodology

All data analyzed in this report came from the 2012-13 academic year and were provided by charter schools that use Direct Instruction programs as their core curriculum. The schools differed in their racial-ethnic composition, their students' level of poverty, and their experience using Direct Instruction. All three sites received technical support for implementation from the National Institute for Direct Instruction (NIFDI), including extensive guidance in assessing students' mastery of the curriculum and determining appropriate placements for students given their prior learning.

Sites in the Analysis and Measures of Achievement

A group of charter schools in Texas provided data on lesson progress and results on 1) the State of Texas Assessments of Academic Readiness (STAAR) for students in grades 3 to 5 and 2) a test with national norms, the Northwest Evaluation Association's Measures of Academic Progress (MAP) for those in grades K to 5 (NWEA, 2011a, b). Over 80 percent of the students in the system qualified for free or reduced meals, and about 40 percent were classified as having Limited English Proficiency (LEP). The system was in its second year of implementation of DI programs in 2012-13.

A small charter school in the state of Colorado provided data on lesson progress and scores on their state assessment, the Transitional Colorado Assessment Program (TCAP), for students in grade 3 (CDOE, 2011). About half of the students were Hispanic, and close to 20 percent had Limited English Proficiency. Slightly more than half (54%) qualified for free or reduced lunch. The school was in its first year of implementation of DI programs in 2012-13.

A charter school in an inner city of a large metropolitan area in the eastern United States provided data on lesson progress and MAP scores for students in grades K to 4. All of the students were African American and about 90 percent qualified for free or reduced meals. This school had been using DI programs as the core curriculum for over ten years.

Measuring Progress through *Reading Mastery* and *Connecting Math Concepts*

As noted briefly above, all DI programs are designed to ensure cumulative mastery of a subject as students proceed through the curriculum. At each lesson, students have the basic knowledge needed for more advanced learning and do not have to repeat material that was learned at earlier stages. In addition, the earlier learning is systematically reinforced. Components within the program provide ways of testing for mastery and specific instructions for what teachers should do when students have not achieved that level. The analysis presented in this report utilizes this cumulative nature of the programs by examining the relationship of students' progress through the lessons to their scores on each measure of achievement.

Table 1 shows the data used to calculate the measure of cumulative lesson progress in *Reading Mastery*. The measure can theoretically range from 1 (the first lesson in the kindergarten program) to the final lesson in the program. For instance, there are 160 lessons in *Reading Mastery Signature Edition (RMSE)* in the Kindergarten level of the program. A student beginning *RMSE Grade 1* (the start of first grade material) would then be at the 161st lesson of the series. *RMSE Grade 1* also has 160 lessons, so a student at the end of *RMSE Grade 1* would have completed 320 lessons. A fifth grader in grade-level

Table 1
Cumulative Lesson Progress Scores in Reading Mastery Signature Edition (RMSE)

Level	Number of Lessons	Starting Cumulative Lesson	Midway Cumulative Lesson	Ending Cumulative Lesson
K	160	1	120	160
1	160	161	240	320
2	145	321	392	465
3	140	466	535	605
4	120	606	665	725
5	120	726	785	845

Note: Students in the upper grades who are significantly behind grade level may be placed in the DI program *Corrective Reading*, designed for older students and moving at a faster pace than *Reading Mastery*. *Corrective Reading B1* was assumed to be equivalent to *RMSE1*, *CRB2* was assumed to be equivalent to *RMSE 2*. However, the pace of progress was assumed to be twice as fast as in *RMSE*. For instance, the placement at lesson 65 in *CRB2* was given a cumulative lesson number of 320, equivalent to the end of *RMSE2*.

material (*RMSE Grade 5*) would be at the 726th lesson at the start of the year and the 845th lesson at the end of the year. In the discussion below, reference to "on grade level" or "meeting a grade level goal" or "target" indicates that a student was nearing, at, or beyond the ending lesson for their grade level at the close of school in the spring.

Table 2 shows the data used to calculate the measure of cumulative lesson progress in *Connecting Math Concepts: Comprehensive Edition (CMCCE)*. The measure can theoretically range from 1 (the first lesson in the kindergarten program, Level A) to 755, the final lesson in the program. For instance, there are 120 lessons in *CMCCE-A*. A student beginning *CMCCE-B* would then be at the 121st lesson of the program. CMCCE-B has 125 lessons, so a student at the end of *CMCCE-B* would have completed 245 lessons. A fifth grader in grade-level material (*CMCCE-F*) would be at the 636th lesson at the start of the year and at lesson 755 at the end of the year.

Table 2
Cumulative Lesson Progress Scores in Connecting Math Concepts Comprehensive Edition (CMCCE)

Level	Number of Lessons	Starting Cumulative Lesson	Midway Cumulative Lesson	Ending Cumulative Lesson
A (K)	120	1	60	120
B (1st)	125	121	183	245
C (2nd)	130	246	310	375
D (3rd)	130	376	440	505
E (4th)	130	506	570	635
F (5th)	120	636	695	755

Lesson Progress in *Reading Mastery* and Achievement

This section of the report examines the relationship between students' cumulative progress in *RMSE* and their assessment scores. The first sub-section summarizes descriptive statistics on students' progress in *RMSE* in the three sites. The following sub-sections describe the relationship of this lesson progress to students' reading scores on the two state assessments, the STAAR and the TCAP, and the MAP, which has well developed national norms. Two analytic approaches are used. The first simply looks at the relationship of students' success on the exams, defined as reaching established benchmarks or surpassing national norms, to their placement in the programs. The second focuses on scale scores on

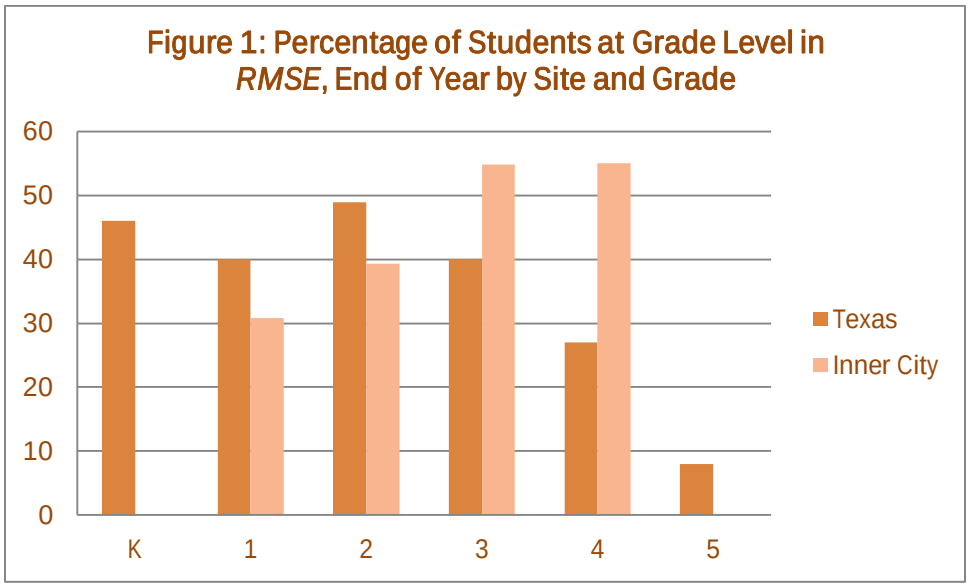
the exams. Multivariate analyses adjust for students' prior achievement giving an estimate of the "value added" to test scores by lesson progress that is independent of prior scores.

Lesson Progress in *RMSE* at Three Sites

The students in the analysis varied in the extent to which they had progressed to the end of their grade level program at the conclusion of the school year. Variations were apparent across the sites and between grades within the two sites with more than one grade in the analysis (Texas and the inner city site). These variations appear to be related both to the demographic characteristics of the students and the extent to which teachers, administrators, and students had experienced Direct Instruction. Research indicates that students' timely progress through DI programs, and thus their academic achievement, is higher when teachers have more experience in the program and when students have been exposed to it for a greater proportion of their school years (Engelmann & Engelmann, 2004; Stockard, 2011a,b; Vitale and Joseph, 2008).

The lowest level of lesson progress was in the Colorado school, which was in its first year of implementation. Only 13 percent of the school's third graders were at grade level in reading at the end of the year, 38 percent were within one grade of being on level (i.e. at some point in third grade material at the end of the year), but about half were more than one grade level behind (in second grade material or lower). (See Table A-2.)

Lesson progress varied substantially across grades in the other two sites, as shown in Figure 1. In the Texas schools, which were in their second year of implementation, students in the



Note: Thirteen percent of the Colorado students were at grade level in *RMSE* at the end of the year.

lower grades were substantially more likely than those in the higher grades to be on grade level at the end of the school year, no doubt reflecting their exposure to the program for a greater proportion of their school career. Forty to fifty percent of the Texas students in kindergarten to third grade were at grade level. In contrast, only about 25 percent of fourth graders and less than 10 percent of the fifth graders, who began the program much later in their school careers, were on target to finish their grade level program at mastery. (See Table A-1.)

The pattern differed for students in the inner city site, in which DI had been implemented for the longest period of time, but which also had the highest level of poverty. The percentage of students on grade level in *RMSE* increased over the grade levels, ranging from less than a third of the first graders to over half of the fourth and fifth graders. This may suggest that the students in this school, who began with extraordinarily large deficits in language development and school readiness, were gradually catching up with their appropriate grade placement. (See Figure 1 above and Table A-3.)

Importantly, there was substantial variation in lesson progress in each of the sites, with some students more likely than others to be at grade level. This provides the basis for a sound test of the hypothesis that progress through the curriculum is associated with higher test scores.

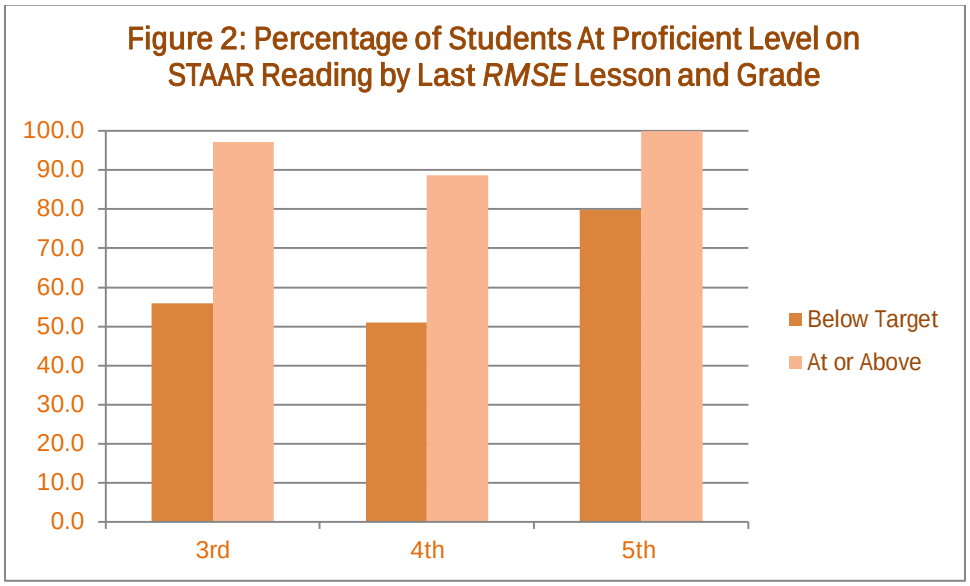
Lesson Progress and Meeting Normative Standards in Reading

The No Child Left Behind Act requires that all public schools report the percentage of students who reach a state-established proficiency level on their state's measure of achievement. The first sub-section below examines the relationship between lesson progress in *RMSE* and the probability that students surpassed the state established benchmarks for the STAAR and the TCAP. Many schools also use scores from nationally normed achievement tests, such as the MAP, as an additional assessment. These tests provide information on national averages to which students' performance can be compared. The second sub-section looks at the relationship between lesson progress and meeting or surpassing national averages on the MAP.

Lesson Progress and State Established Benchmarks –Figure 2 reports the percentage of students who scored at the proficient level or higher on the reading portion of the STAAR, comparing those that were at or beyond the last *RMSE* lesson for their grade and those who were at lower levels.[3] The results are clear cut. At each grade students who had adequate lesson progress, defined as mastering material at the end of the year that was close to, at,

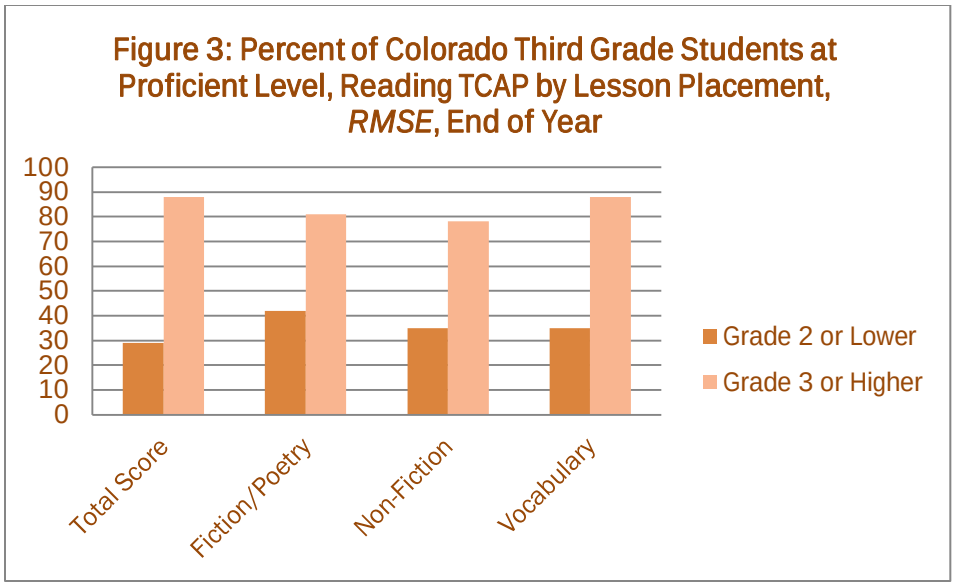
[3] Because so few fifth grade students had completed their grade level program, the group of students at the targeted level includes those who were within 90 percent of the final lesson.

or beyond their assigned grade placement, were far more likely to score at the proficient level on the STAAR. All of the relationships were statistically significant, and the associated effect sizes ranged from 0.92 to 1.25 far beyond the traditional criterion of educationally important effects (.25, Tallmadge, 1977). The relationship is especially striking for fifth graders. All of the students who were at or near grade level in *RMSE* scored at the proficient level or higher on the STAAR. At the other grades, about 90 percent or more of the students at grade level in *RMSE* scored at the proficient level. (See Table A-4 for detailed data.)

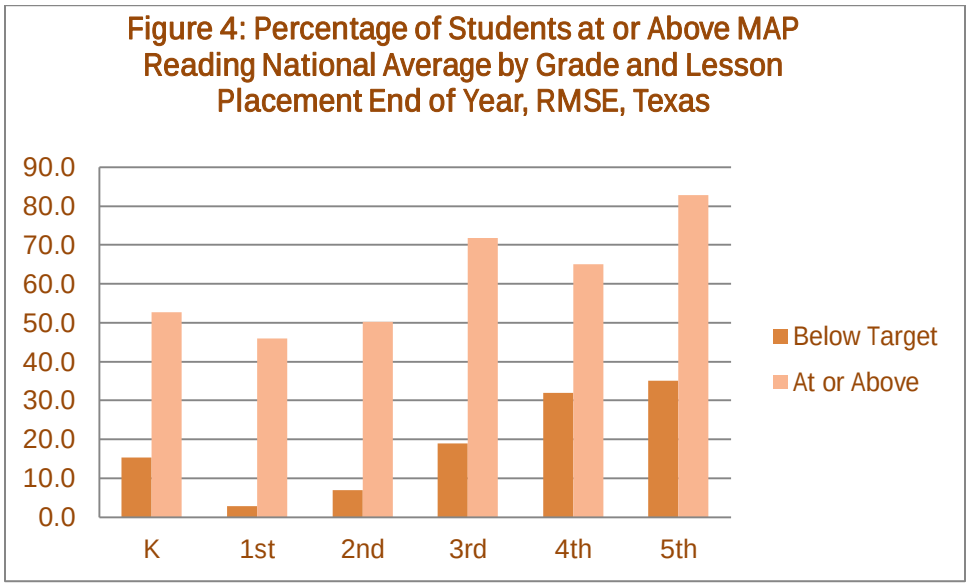


Note: Because so few fifth grade students had completed their grade level program, the group of students at the targeted level at that grade includes those who were within 90 percent of the final lesson. At 3rd and 4th grade the "target" represents the last lesson for the grade level.

Figure 3 reports data on the percentage of students in the Colorado school who scored at the proficient level or higher on the total score of the reading portion of the TCAP as well as various sub-parts of the exam. As noted above, relatively few students in this school were at grade level in *RMSE* at the end of the school year. Thus, to increase the number of students in the comparison, data in Figure 3 contrast students who were at any point in grade 3 material or higher by the end of the year (a maximum of only one year behind grade level) and those who were in grade 2 level or lower, It is clear that those who were closer to grade level were much more likely to pass the state exam. Almost 90 percent of the students who had mastered material at some point in grade 3 by the end of the year had passing scores on the total reading assessment, but less than a third of those who were below this level had passing scores. All of the results were highly significant even though the sample size is relatively small. The associated effect sizes ranged from .88 to 1.52. (See Table A-4 in the Appendix.)



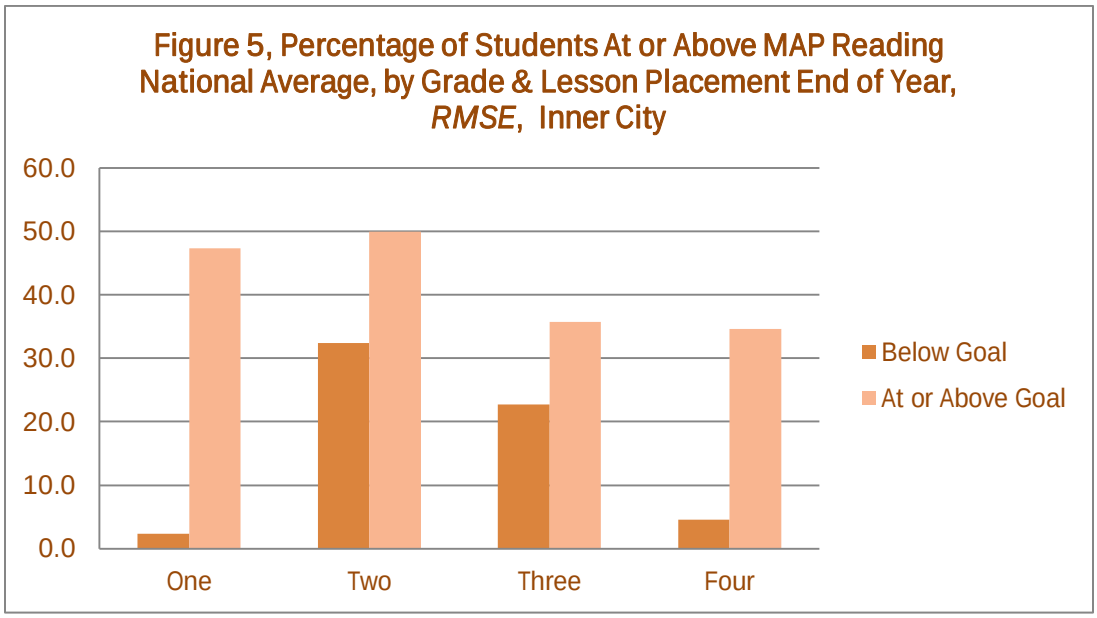
Lesson Progress and Scoring above the National Mean on the MAP – Data on scores on the nationally normed MAP were available for both the Texas and inner city schools. Figure 4 reports the percentage of students in the Texas schools who had MAP reading scores at or above the national mean for those who were at or near grade level in *RMSE* at the end of the school year and those who were below grade level. By definition, scoring above the



Note: Because so few fifth grade students had completed their grade level program, the group of students at the targeted level at that grade includes those who were within 90 percent of the final lesson. At 3rd and 4th grade the "target" represents the last lesson for the grade level.

national mean is a substantially higher bar than scoring at a state defined proficiency level. Yet, as with the data on state assessments, the results are clear cut. At each grade students who were on pace with mastering lessons at their assigned grade level were far more likely to score above the national mean. All of the relationships were statistically significant, and the associated effect sizes ranged from .71 to 1.29. (See Table A-5,) Again, the relationship was especially striking for fifth graders. Of those who had mastered *RMSE* material at or near their grade level over 80 percent scored above the national mean on the MAP, 30 percent more than would be expected simply by chance.

Figure 5 reports similar data for the high poverty inner city school. The students in this school were less likely than those in the Texas schools to score above the national mean, but differences between those with varying levels of lesson progress were similar in the two settings. Inner city students at grade level were much more likely to approach the national mean than those below grade level. The differences were much stronger for results in first and fourth grade than for those in second and third, although all of the effect sizes exceeded the usual criterion for educational importance, ranging from .29 to 1.39. (Table A-5.) The results in the inner city school were quite similar to those at the Texas site for students in grades 1 and 2. But, in grades 3 and 4, relatively fewer of the inner city students who were at grade level in *RM* surpassed the national mean. Further work, using panel data, will investigate these differences, with special attention given to areas such as differential rates of mobility, continuous enrollment, and scores in earlier years.



Note: Students in the "at or above goal" group were at or beyond the last lesson at their assigned grade level by the end of the school year.

Value Added by Lesson Progress in *Reading Mastery*

The data presented in Figures 2 through 5 do not adjust for students' past achievement, In other words, it could be possible that students had higher scores on the exams simply because they began the school year with more skills and not because of how far they progressed during the year. Regression analyses were used to address this issue. These analyses provide estimations of the extent to which progress in *RMSE* "added value" to students' achievement beyond the scores that they would be expected to receive given their prior levels of achievement.[4] Students' spring scores were first predicted by their previous achievement scores, and then their cumulative lesson progress in *RMSE* was added to the equation. This allows one to see the extent to which progress in the program added to students' achievement scores beyond what would be expected given their earlier performance.

To add greater precision to the analyses, the measures were continuous in nature. The dependent measures were the continuous measure of students' assessment scores, rather than whether or not they had reached an established benchmark. The measure of lesson progress was also continuous in nature, rather than the categories of meeting or not meeting a targeted goal. For the analyses of MAP scores the measure of previous achievement was students' MAP scores in the fall. For the analyses of STAAR scores for fourth and fifth graders the measure of previous achievement was STAAR scores in the previous school year. For third graders the measure of previous achievement was their fall MAP scores. (The STAAR is only given in grades 3 to 5, so they had no STAAR data from their second grade year.)[5] (See Tables A-6 to A-9.)

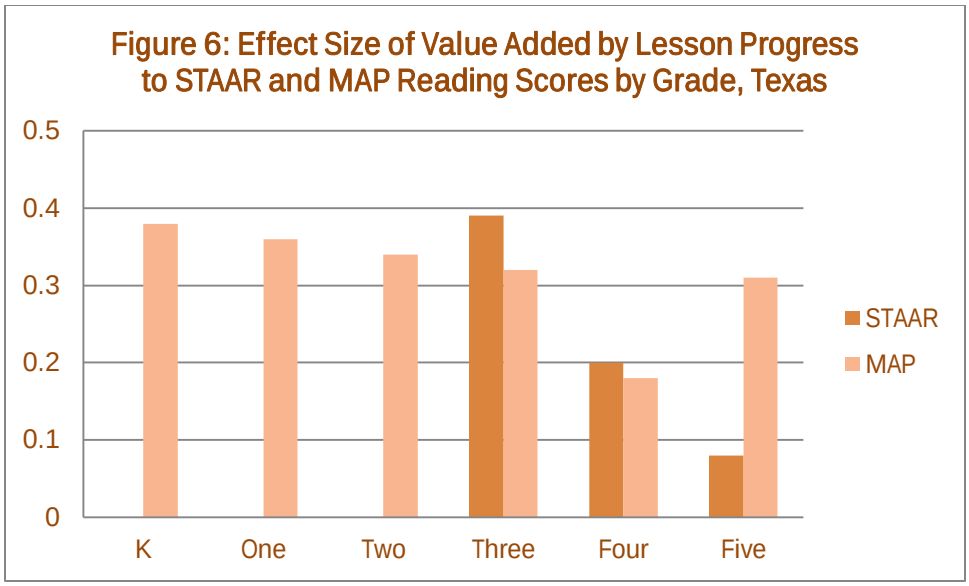
The data given in Figure 6 summarize the results for the STAAR and MAT assessments from the Texas schools. The reported values are standardized regression coefficients and can be interpreted as effect sizes, for they indicate the association between lesson progress and achievement in standard deviation units once prior test scores were equalized (controlled). All of the effects are positive and most surpass the usual criterion (.25) of educational importance. All but one of the effects was statistically significant at well beyond the .001 level.[6] (See Tables A-6 and A-7.)

---

[4] Free or reduced lunch and LEP status were also used as controls for the Texas data, but had very little impact on the results. See details in the appendix.

[5] No measure of previous achievement was available for the Colorado schools. The appendix includes results of a value added analysis for this site that used the students' *RMSE* lesson at the beginning of the year as the measure of prior learning. The results parallel those found with the other sites. The estimates of value added were generally highly significant and the average estimate of value added was approximately twice that found in the other sites. Given the different measure of prior learning and the much larger estimate of effect, these results are not reported here, but are included in the appendix.

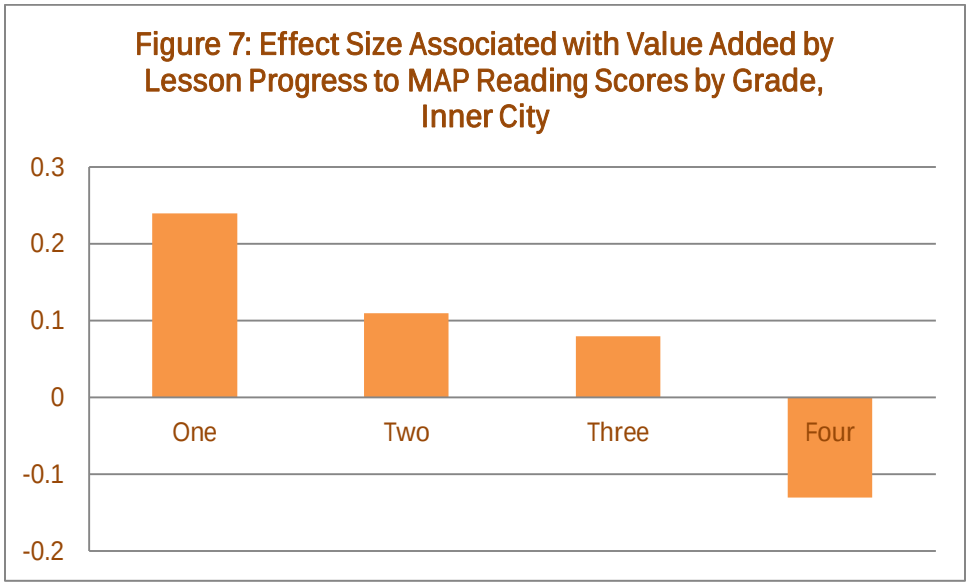
[6] The exception is the coefficient for the STAAR score for reading for fifth graders, where the t-value of 1.65 is associated with a probability level (two-tail) of .10.



Note: For the MAP analyses the effects are net of students' fall MAP scores. For the STAAR analyses the results for grades 4 and 5 are net of students' STAAR score in the previous year and for grade 3 they are net of the fall MAP score. The STAAR was only administered to students in grades 3 to 5.

Interestingly, the estimates of value added were substantially smaller for the very high poverty inner city school. While the estimate for grade one approached the level of educational importance, the estimated effects were smaller for each of the subsequent grades and the estimate for fourth grade was negative. (See Figure 7 and Table A-8.) Further research will examine the reasons underlying this impact, using panel data

.



Note: The effects reported in this figure are net of students' fall MAP scores.

Summary

Taken together, the results, from three different sites with different demographic characteristics, show a strong relationship between students' progress through *Reading Mastery* and their performance on both state assessments and a nationally normed achievement test. Students who were at or approaching mastery by the end of the school year of grade level material in *RMSE* were very likely to pass their state assessments. On average, the probability that students in the Texas schools would score at the proficient level on the reading portion of the STAAR was .95 for those who had mastered grade level material, but only .62 for those who were at lower levels. Students in the Colorado school who were at mastery of material at any point within their grade level curriculum had almost a 90 percent probability of scoring at the proficient level on their state assessment. In contrast, those who were more than one year behind grade level had only a .29 probability of doing so. Similar results appeared with the nationally normed MAT achievement test scores, with those who had mastered grade level material being much more likely to score at or above the national mean.

These results were confirmed by the "value added" analysis, which included strong controls for students' prior achievement. On average, the value added to students' spring test scores, above what they would have been expected to score based on their previous learning and performance, was .21 of a standard deviation. In other words, *the results presented in this section indicate that students' cumulative progress, at mastery, in RMSE increased their spring test scores by more than a fifth of a standard deviation beyond their previous achievement levels.*

Lesson Progress in *CMCCE* and Mathematics Achievement

This section describes results regarding the relationship of lesson progress in *Connecting Math Concepts: Comprehensive Edition (CMCCE)* and achievement scores in mathematics. The analysis parallels that used with the measures of reading achievement, looking first at lesson progress at mastery in *CMCCE*; then at the relationship of students' success on the exams, defined as surpassing national norms or reaching established benchmarks, and their progress in the programs; and third, examining scale scores (continuous measures) and using multivariate analyses to adjust for students' previous achievement to calculate the "value added" to test scores by progress through the *CMCCE* program. The results parallel those reported for reading with strong association of students' progress at mastery through *CMCCE* and their test scores.

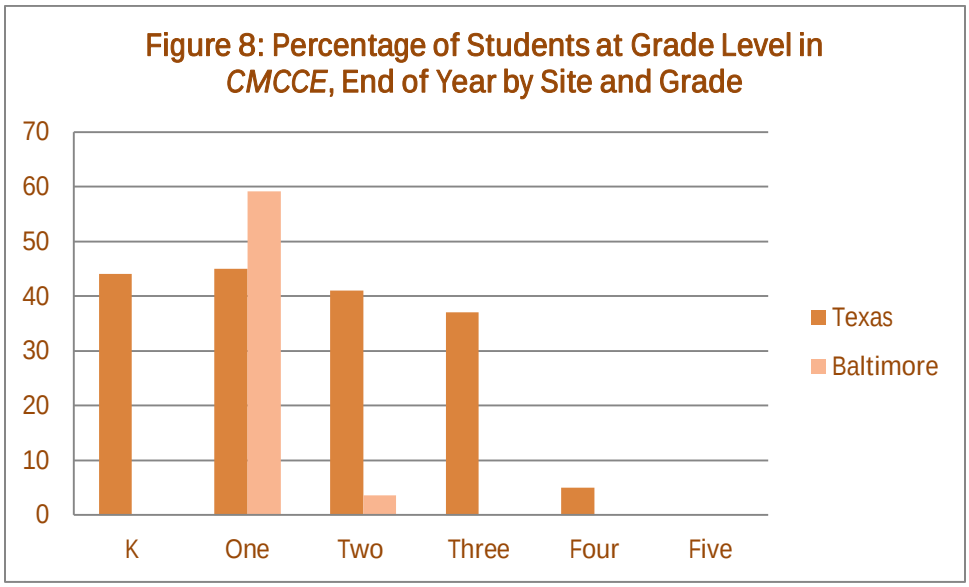
Lesson Progress in *Connecting Math Concepts* at Three Sites

The students in the analysis varied in the extent to which they had progressed to the end of their grade level program in mathematics by the conclusion of the school year. Variations in

progress were apparent across the sites and between grades within the two sites with multiple grades in the analysis. As with reading, the variations appear to be related to both the demographic characteristics of the students and the length of time that DI programs had been used in the schools. In addition, students generally made less progress in mathematics than in reading.

The lowest level of lesson progress was in the Colorado school, which was in its first year of implementation. About half of the students began the year in second grade material and the remainder started with third grade material. However, only one of the students who began in second grade material finished that level and moved to the third grade level. None of the students who began in third grade material finished the lessons for that grade level. The most advanced students had mastered material through the midpoint of the third grade curriculum by the end of the year. (See Table A-10.)

Figure 8 gives data for the Texas and inner city schools, reporting the percentage of students nearing or reaching the end of the grade level program by the spring of 2013. As in reading, lesson progress for the Texas students was stronger in the earlier grades. Over a third of the Texas students in grades K-3 were at mastery at grade level in *CMCCE* by the end of the year, but only 5 percent of the fourth graders and none of the fifth graders had reached that point. (See Table A-12.) As noted for the reading results, this likely reflects the fact that the higher grade students began the program in their later years of elementary



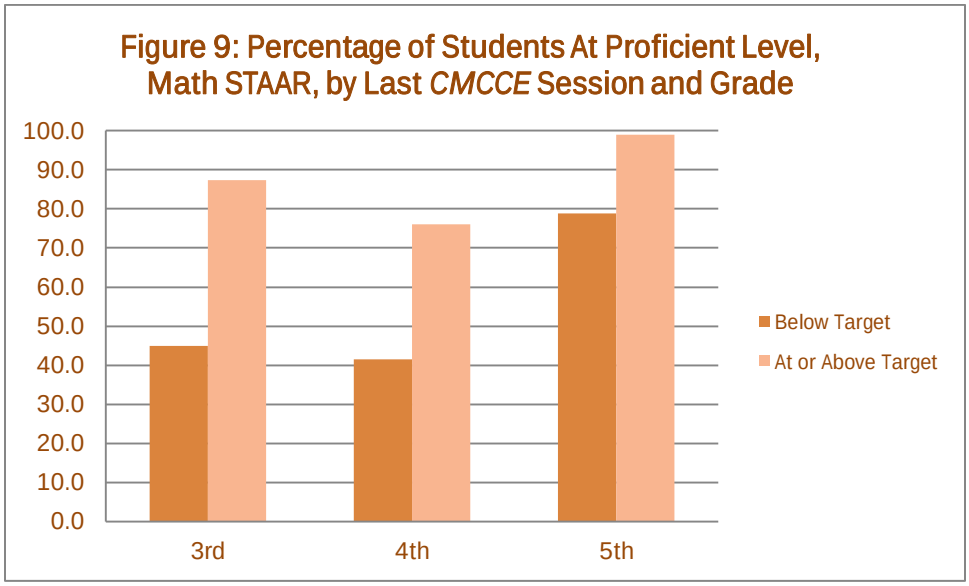
Note: Data were available for grades K to 5 for the Texas schools and grades 1 to 4 for the inner city school. The bars represent the percentage of students who had mastered the last lesson for the assigned grade level by the end of the school year. None of the Colorado students reached the last lesson in their grade level program.

school and thus had greater amounts of material to learn to approach mastery at grade level. The pattern across grade levels was similar for the inner city school. Over half of the first graders were at grade level, but less than five percent of the second graders and none of the third or fourth graders had mastered their grade level material. (See Table A-11.)

Lesson Progress and Meeting Normative Standards in Mathematics

The first sub-section below examines the relationship between lesson progress in *CMCCE* and the probability that students surpassed the state established benchmarks for the mathematics portions of the STAAR and the TCAP. The second sub-section looks at the relationship between lesson progress and meeting or surpassing national averages on the mathematics portion of the MAP.

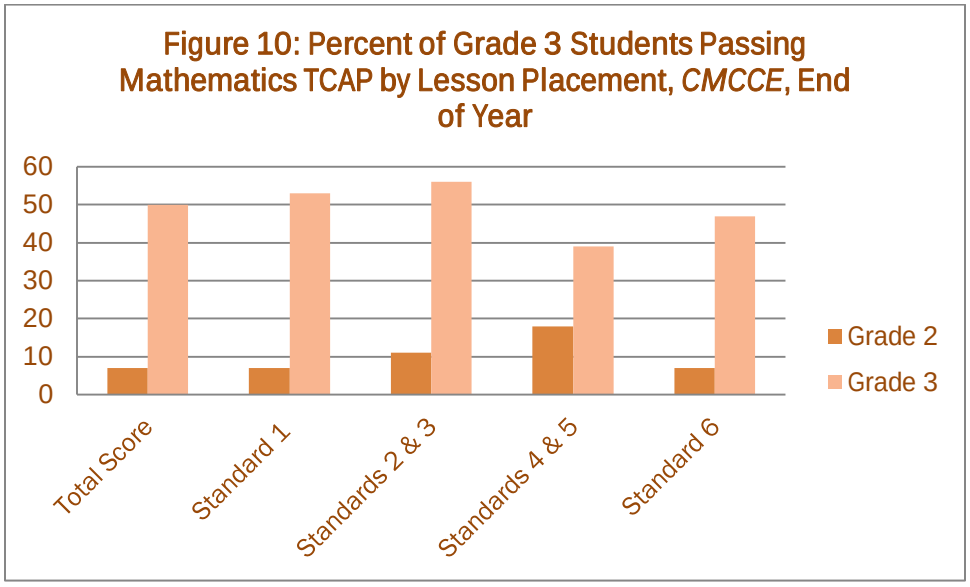
Lesson Progress in CMCCE and State Established Benchmarks –Figure 9 reports the percentage of students who scored at the proficient level on the STAAR mathematics test within two categories: those that did not master the last *CMCCE* lesson for their grade and those that were near, at or beyond the last lesson. (For fifth grade, because no student was on a trajectory to meet the end of year lesson, the distinction is between those who were 60% or more of the way through the program and those below that point. For fourth graders the distinction is between those who were at least 90% through the program and others.) The results are clear cut and very similar to those found with reading. At each grade students who were at or approaching grade level in *CMCCE* were far more likely to score at



Note: For fourth graders, the comparison is between those who were at least 90% through the program and others; for fifth graders the comparison is between those who were at least 60% through their grade level program and others.

the proficient level on the mathematics portion of the STAAR. (See Table A-13.) All results were statistically significant and the associated effect sizes were large, ranging from .75 to 1.02.

Figure 10 shows the percentage of students who passed the TCAP for those who had mastered some grade 3 material by the end of the year (thus less than one year behind grade level) and those who were only in grade 2 material (and thus more than one year behind their assigned grade level). It is clear that those who were closer to grade level in *CMCCE* were much more likely to pass the state exam. Half of the students who were in grade 3 material by the end of the year had passing scores on the total mathematics assessment, but only two (7%) of those who were below this level had passing scores. The patterns were similar across all of the subscales, although slightly less marked with results related to standards 4 and 5, which involve geometry and measurement. All of the results, except those with standards 4 and 5, were highly significant even though the sample size is relatively small, and the effect sizes were again substantial in size. (See Table A-14.)

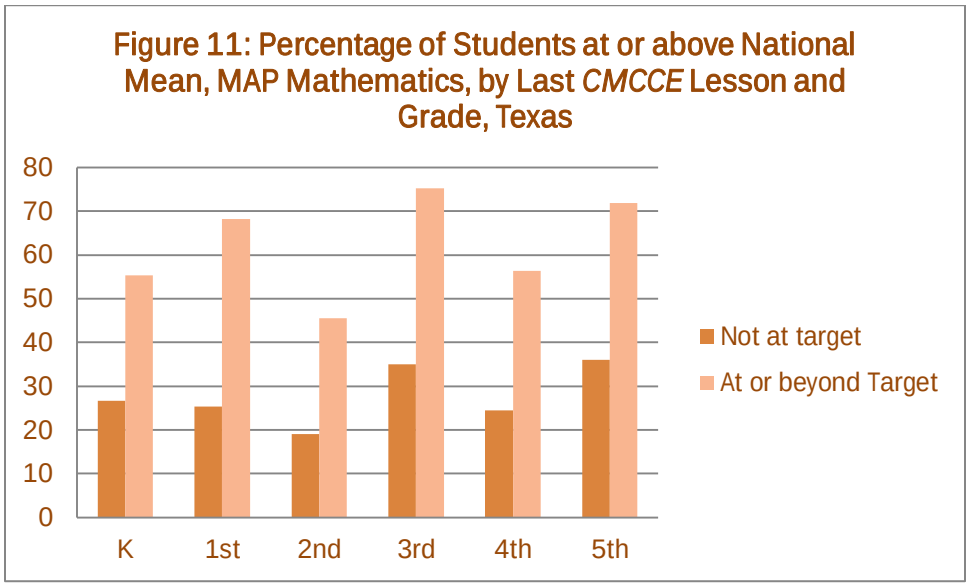


Note: No students had reached the end of the third grade level program by the end of the year. Some were still in second grade material (labeled grade 2 in the graph) and the others were between the start and middle of the third grade material (labeled grade 3).

Lesson Progress and Scoring Above the National Mean on the MAP –Figure 11 reports the percentage of students in the Texas schools who had MAP mathematics scores that were at or above the national norms, comparing those who were approaching, at, or beyond mastery of the last lesson for their grade with other students. Again, the results are clear cut. At each

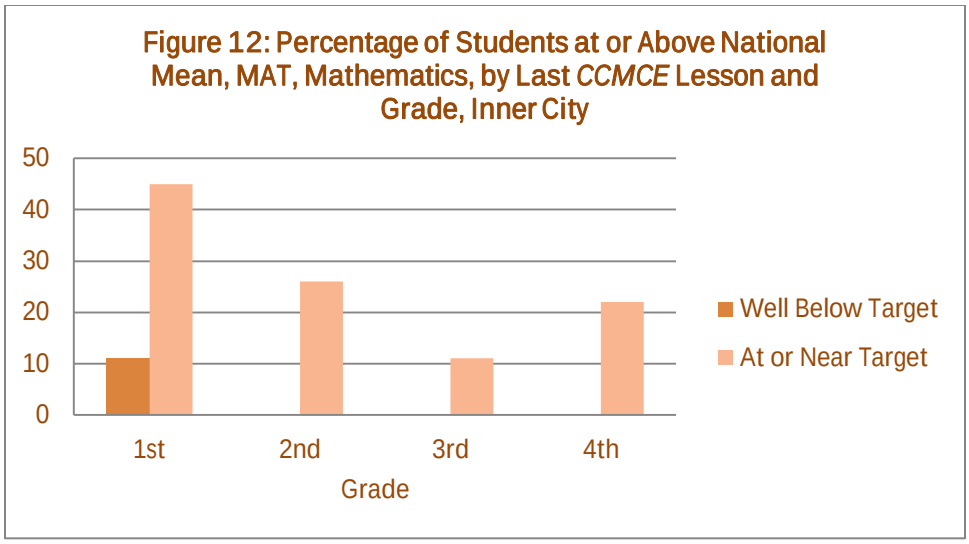
grade students who had progressed further through the program were far more likely to score above the national average on the MAP.

Students in the inner city school had substantially lower scores on the MAP. Those in first grade, the only group where substantial numbers of students had finished the grade level program, scored the highest. Almost half (45%) of the first graders who were at grade level scored at or above the national mean in contrast to 11% of those who were below grade level. The only students in the higher grades who exceeded the national mean were those who had mastered higher levels of the *CMCCE* program (beyond the mid-point of the year's



Note: For fifth grade, because no student was on a trajectory to meet the end of year lesson, the distinction is between those who were 60% or more of the way through the program and those below that point and for fourth grade the results distinguish those who had completed 90% of the program and those below that point.

curriculum for 2nd and 3rd graders and at any point in the grade level for 4th graders). No other students in these grades exceeded the national mean. (See Figure 12 and Table A-15.)

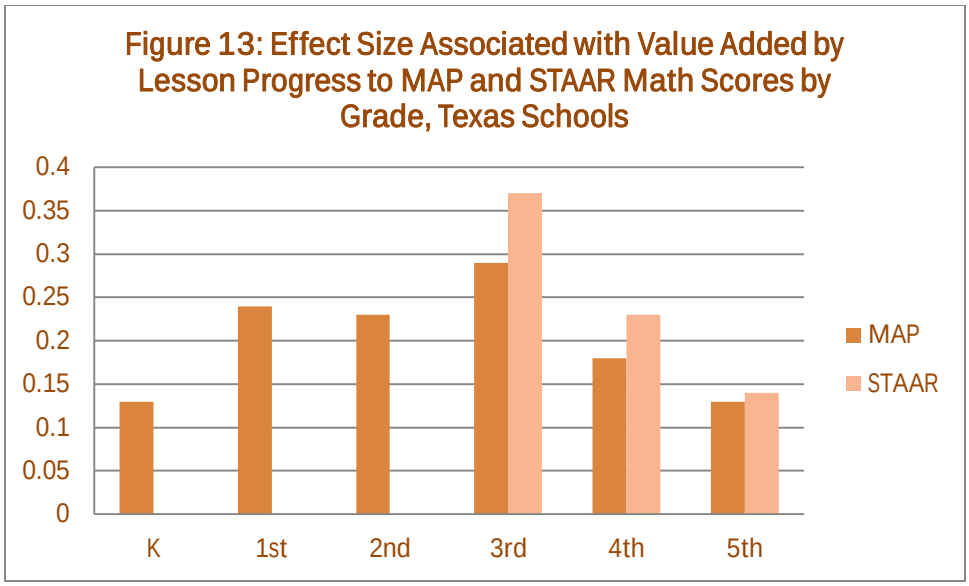


Note: The "At or Near Target" group varied by grade level. For first grade it included students who were at the end of their grade level curriculum. For second and third grade it included students who were beyond the mid-point of the year's material. For fourth graders it included any students who were in the fourth grade material.

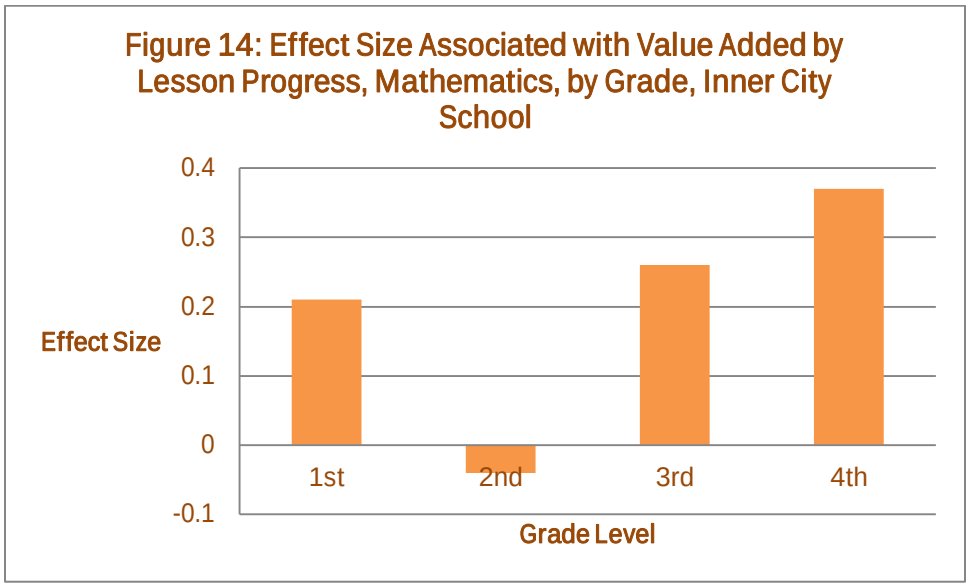
Value Added by Lesson Progress in *CMCCE*

Figures 13 and 14 summarize the results of the regression analyses for the mathematics measures with controls for previous levels of achievement.[7] As explained above, the reported values can be interpreted as effect sizes, for they indicate the association between lesson progress and achievement in standard deviation units. Importantly, because they control for prior levels of achievement, these coefficients indicate the "value added," or the extra boost, that students' cumulative mastery of the material in *CMCCE* gave to their test scores beyond what would be expected by their previous achievement. Figure 13 gives the results associated with MAP and STAAR scores for the Texas students. All of the effects were positive and highly significant. Figure 14 gives the results associated with MAP scores for the inner city students. Two of the results were statistically insignificant (for grades one and two), but those for the upper grades were significant and surpassed the level generally used to denote educational importance.

[7] As with reading, no information was available for prior test results for the Colorado school. Starting lesson in *CMCCE* was the only indicator of prior learning. However, this measure was very highly correlated (r=.95) with the final lesson in *CMCCE* (reflecting the very low level of lesson progress in the school described in the text). Given this very high correlation, a value added analysis was not statistically appropriate for this site. More information is included in the appendix.



Note: For the MAP analyses the effects are net of students' fall MAP scores. For the STAAR analyses the results for grades 4 and 5 are net of students' STAAR score in the previous year and for grade 3 they are net of the fall MAP score. The STAAR was only administered to students in grades 3 to 5.



Note: Prior achievement was measured by Fall MAP mathematics scores.

Summary

The results regarding mathematics are similar to those regarding reading. Students who were at or approaching mastery of grade level material in *CMCCE* were far more likely than those who were at lower levels in the program to score at the proficient level or higher on their state assessments. On average, the probability that students in the Texas schools

would score at the proficient level on the mathematics portion of the STAAR was .87 for those who were at or approaching mastery of grade level material, but only .55 for those who were at lower levels. Students in the Colorado school were far less likely to be progressing in their program, but half of those who were at any point within their grade level curriculum by the spring of the year scored at the proficient level on their state assessment compared to only two students (7%) who were more than a year behind. Similar results appeared with the nationally normed MAP achievement test, with those who were at or approaching mastery of grade level material in *CMCCE* being much more likely to score at or above the national mean.

As with reading, these results were confirmed by the value added analysis, which included strong controls for students' prior achievement. All but 2 of the 13 value added effects were statistically significant. The average effect size was .21 of a standard deviation, identical to that found for reading.

Summary and Discussion

All Direct Instruction programs incorporate mastery learning, sometimes described as analogous to a stair-stepped progression (Engelmann, 2014b). Each step or lesson in the program builds on previous learning, thus making it easier to learn material thoroughly and to progress more quickly through a subject. DI programs include explicit instructions on how teachers can assess their students' mastery and make sure that they are appropriately placed for building on previously learned material. This report has examined the extent to which students' cumulative mastery of the programs is related to scores on external measures of achievement: state assessments for Texas (STAAR) and Colorado (TCAP) and the Measures of Academic Progress (MAP) derived from a nationally normed test developed by the Northwest Evaluation Association.

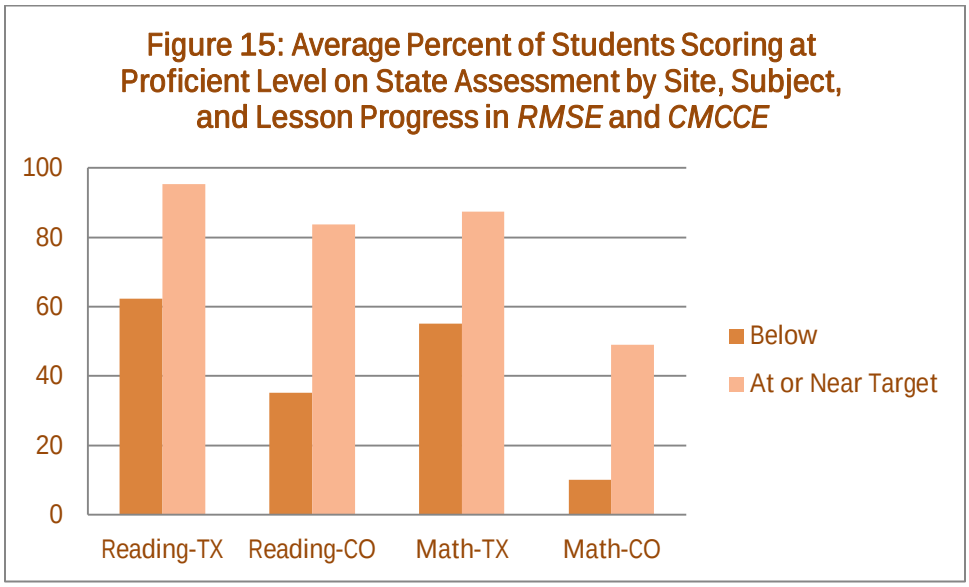
Data were obtained from schools in three different sites: Texas, Colorado, and an east coast inner city and from a range of grades (K to 5 in Texas, grade 3 in Colorado, and grades 1-4 for the inner city school). There was substantial variation both within and between schools on students' progress through the curriculum. There was also substantial variation in the demographic characteristics of students, including both race-ethnicity and levels of poverty. This diversity is especially helpful for providing numerous replications of the analysis.

All of the schools received technical support from the National Institute for Direct Instruction and had strong procedures to help ensure that students were placed at points within the DI programs where they were at mastery. Thus the students' placement in the curriculum

indicated the point at which they had learned, or mastered, the material presented in the curriculum to that point.[8]

The Findings

The results indicate that students who were at or approaching mastery at their assigned grade level in *RMSE* and *CMCCE* had a strong probability of scoring at the proficient level on state assessments, while far fewer students who were at lower levels in the program were found to be proficient. Figure 15 summarizes these differences by displaying, for both of these groups, the average percentage of students who scored at the proficient level at each site and for each subject. In most cases, on average, over four-fifths of the students who were at or near mastery at their grade level placement scored at the proficient level. The only exception involves mathematics for the students in Colorado, where the definition of "near target" included any student who was in the third grade level of the program. Half of those students scored at the proficient level, compared to only 10 percent of those below that level. All of the comparisons were statistically significant. The associated effect sizes were large and generally several times the usual level used to denote educationally important effects (averaging 1.12 for reading and .95 for mathematics).

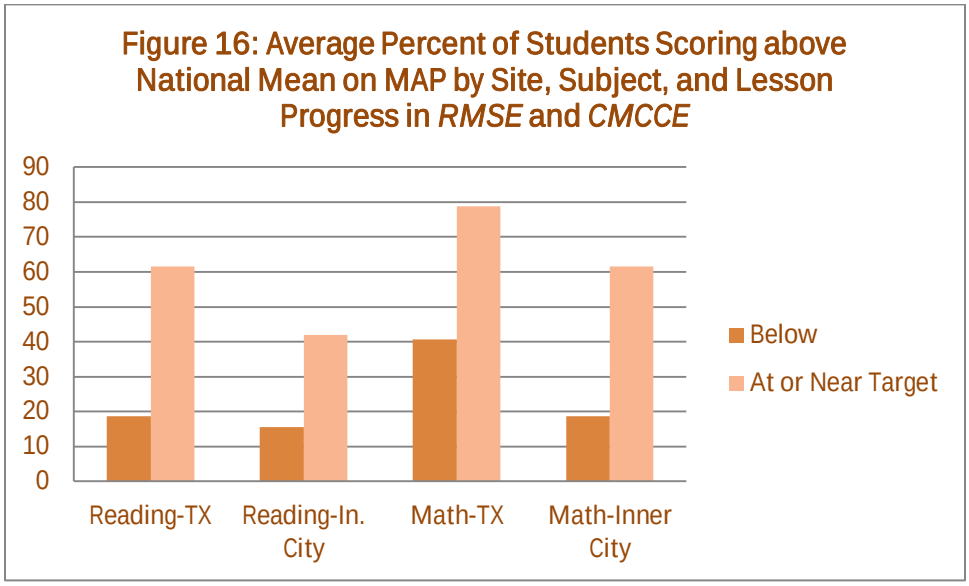


Note: The values in the figure represent the average percentage across the grades within each of the sites. As explained in the body of the text above, to have an adequate sample size for statistical comparisons and given varying rates of lesson progress, the definition of "at or near target" varied across sites and grades. For Texas it was defined as the last lesson for a given grade for grades K-4 for reading and for grades K-3 for mathematics. For grade 5 in reading and grade 4 in mathematics, it was defined as being 90% through the grade level material. For grade 5 in mathematics it was defined as being 60% through the grade level material. For the Colorado school "at or near target" was defined as being at any point in grade 3 material for both reading and mathematics.

[8] It is, of course, possible that some students were not truly at mastery for their designated point in the curriculum. However, to the extent that this occurred, the estimates of the effects of *RMSE* and *CMCCE* on students' test scores are conservative in nature.

Similar results appeared with the MAP data, which were available for the Texas and inner city sites. As shown in Figure 16 students who were at or near mastery of grade level material in their DI programs were much more likely to score above the national mean. Almost all of the differences were statistically significant, even though some of the sample sizes were relatively small. Again, the effect sizes associated with these comparisons were substantial, averaging .93 for reading and .83 for math.

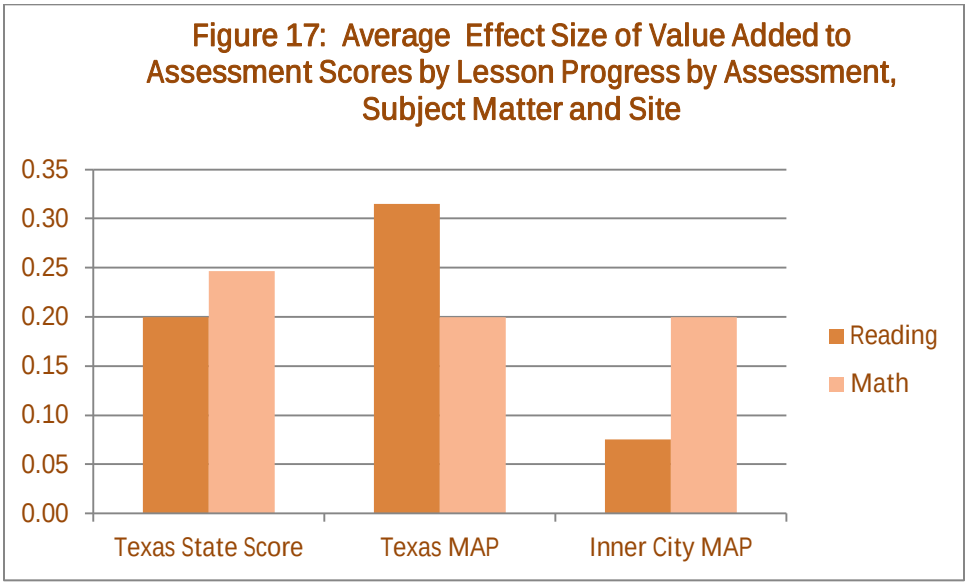
Most importantly, the strong relationship of lesson progress to achievement scores persisted when students' prior achievement was controlled. Average effect sizes associated with the value added for each site, assessment, and subject are shown in Figure 18.[9] All of the values are positive and all but one is greater than .20. Most of the individual estimates of value added were statistically significant, even with relatively small samples and strong



Note: The values in the figure represent the average percentage across the grades within each of the sites. As explained in the body of the text above, to have an adequate sample size for statistical comparisons and given varying rates of lesson progress, the definition of "at or near target" varied across sites and grades. For Texas it was defined as the last lesson for a given grade for grades K-4 for reading and for grades K-3 for mathematics. For grade 5 in reading and grade 4 in mathematics, it was defined as being 90% through the grade level material. For grade 5 in mathematics it was defined as being 60% through the grade level material. For the inner city school "at or near target was defined as being at the last lesson of the grade level for all grades for reading. For mathematics it was defined as being at the last lesson of the grade level or beyond for first grade, at the mid-point of the lessons for the grade level for second and third graders and at any place within the grade level material for fourth grade.

[9] As explained in earlier sections and in the Appendix the measure of earlier achievement for the Colorado site was the starting lesson in the program, rather than earlier achievement scores. The resulting estimates of value added were, as with the other analyses, statistically significant. However, the effects were almost twice as results for the analysis of mathematics for the Colorado site are not included here because of statistical issues involving high rates of intercorrelations of the variables in the analysis.

controls. In other words, when students' previous test scores were used to predict their achievement scores, their current placement in the DI programs provided significant "value added" to their state assessment and MAP scores. *Progress at mastery in their DI programs added significantly to students' achievement scores beyond what they would be expected to score given their earlier performance on the assessment*.



Taken together, these results provide strong evidence that *progress, at mastery, through the DI programs, RMSE and CMCCE, has a strong influence on students' tests scores and that this influence is independent of students' prior levels of achievement.* Students who have mastered material at, or near, their assigned grade level in the programs are very likely to score at the proficient level on state assessments and much more likely than other students to score above the national average on a normed achievement test. The value added to their scores by progress through the program is independent of their prior learning or achievement.

The findings reported here are noteworthy given the replication of results across three different sites and several grade levels. Such replication across settings is a key element required of the accumulation and verification of findings. To have such consistent results provides additional confidence in support of the hypothesis that lesson progress at mastery in the DI programs is related to higher scores on both state assessments and nationally normed tests and that this impact is independent of students' prior levels of achievement.

Additional replications of the analysis would, of course, be interesting and potentially informative. Possible areas of further research could include examining data from other

sites, with other assessments, with other subjects and for time periods longer than an academic year.

Implications for Policy and Practice

Most curricular programs use a "spiral approach," briefly visiting a topic and then returning again later with a slightly different approach, In contrast, Direct Instruction programs embody a mastery curriculum with careful design and extensive field testing that ensure that students who have progressed through the program will, with appropriate instruction, have learned all that has been taught to that point. Each new element that is taught builds on previously learned material. The analysis presented in this paper suggests that this mastery approach may be crucial to the association of students' progress through the programs and their scores on assessments. When students have mastered the curriculum for their grade (or, in several analyses, especially in the higher grades, only the material partly through their grade level), they have a very high probability of passing their state assessment and are much more likely than other students to score above the national mean on standardized tests. The author knows of no documentation in the literature of another curricular program having such a strong association of progress through the program and assessment scores.

The author has not been able to locate studies that document the effect size associated with test preparation for students in the elementary grades, perhaps because this industry is relatively new. In contrast, there is literature, dating from the 1950s through the current century, on the effect of preparation for the SAT and PSAT. This is a relatively large and established industry devoted to helping high school students prepare for college entrance examinations. Summaries of this work indicate that such preparation has only a small impact on students' scores, with effect sizes ranging from .06 to .07 for the verbal tests and .11 to .20 for mathematics (Briggs, 2001; Dominguez and Briggs, 2009). It could be reasonable to assume that the fledgling elementary school test preparation programs would, at maximum, have effect sizes similar to these values. The average effect size for the value added analysis in this study was .21, several times the average reported for the verbal training for the SAT and PSAT and slightly higher than the maximum estimate for the mathematics training.

Given this analysis, it is appropriate to ask if there are situations in which test preparation strategies might be appropriate as adjuncts to study of the curriculum. One could argue that there are some elements of test preparation that would not be seen as "cheating" or "teaching to the test," such as helping students understand the format of exams and the ways in which they should fill out the forms. Moreover, there may be some students who are in need of special help in this regard, for whom performance on standardized assessments does not conform to their placement in their curriculum or their performance on day-to-day

classroom assessments. These students can, undoubtedly, be identified and helped to learn appropriate test taking skills. Given the strong association of lesson placement and assessment scores described in this document, one would, however, presume that these students are, undoubtedly, a small minority.

It is no doubt tempting for administrators to employ elaborate test preparation strategies as a way to raise test scores of students who are substantially behind their peers. Yet, such an approach would, given the effect sizes noted above, probably be relatively ineffective. Moreover, because they detract from the curriculum, they could result in the students being even further behind their peers at the end of the intervention. The DI curriculum includes specific, highly successful, strategies to help students catch up to grade level.[10] In the long run, devoting time to helping students who are the furthest behind their peers begin to catch up would appear to be both more humane and more effective than spending time in the relatively less effective test preparation.

Given the high association of lesson placement and students' test scores, it would probably also be tempting for administrators to place students at higher points within the curriculum, hoping that simple exposure to the material at the higher levels would result in higher test scores. In fact, however, it is very likely that such a tactic would result in these students falling even further behind their peers and, perhaps even worse, coming to think of themselves as poor students. (See Engelmann, 2014b, pp.14-15 for a convincing discussion of the problems with such an approach.)

In short, the results described in this document suggest that a rational alternative to extensive test preparation is to ensure that students make adequate progress, at mastery, through *RMSE* and *CMCCE*. The data suggest that when students have attained mastery at their assigned grade level in these programs the probability that they will pass the state assessments is very high. The influence of lesson progress on achievement is independent of students' prior achievement, providing hope for students who have struggled with low assessment scores in the past.

In contrast to the learning that occurs with test preparation, students engaged in the subject matter through *RMSE* and *CMCCE* are building a broad array of content knowledge that prepares them for continued academic success. Enhancing student progress through the

[10] For students in the primary grades the strategies often involve alterations to the schedule to ensure that they have adequate time for moving through the programs.. With older students special programs, such as *Corrective Reading* and *Corrective Mathematics*, are designed to go through the curriculum at about twice the pace of the regular programs (*RMSE* and *CMCCE*), thus helping these students also begin to catch up with their peers.

curriculum, rather than devoting excessive time to test preparation, could be a "win-win" situation. Students would be more likely to pass state assessments and have higher scores on national exams. They would also be learning academic content and the models provided to them would involve mastery of learning rather than simply learning to "take a test" or "game the system." A direct, effective, and morally appropriate way for school administrators to address the pressure to have higher test scores would be to facilitate their progress at mastery through the DI curriculum.

Appendix: Statistical Details

This appendix provides statistical details on the data reported in the body of this report. The first major section provides data on the reading analysis and the second section provides data on the analysis of mathematics. The tables follow the text.

Reading Analyses

The discussion below is divided into three subsections. The first gives descriptive information on the schools in the analysis. The second provides the results regarding the association of lesson progress with achieving at the proficient level on state exams and scoring above the national mean on the MAP. The third discusses details on the value added analysis for measures of reading achievement.

Descriptive Data

Descriptive data on the Texas Schools are in Table A-1. The analysis was limited to cases with valid data for a given achievement measure and data on lesson progress. Descriptive statistics are given on the scale scores for the STAAR and MAP, the percentage of students meeting benchmark on the STAAR and exceeding the national norm on the MAP, the average cumulative reading lesson that students had reached at the start and end of the year (as defined in Table 1 in the text), and the percentage of students at grade level at both the beginning and end of the school year.[11] There were more missing data for students in the lower grades than in the upper grades and more missing data for MAP measures than for the STAAR. Details on sample size for each measure are also in Table A-1.

The data in Table A-1 show substantial variability on all measures. Some students had much higher achievement scores than others; some had much more lesson progress than others. Those in the higher grades were, on average, far less likely than those in the lower grades to have reached the end of the year target for their grade. In fact, only eight percent of the fifth grade students were on a trajectory to have reached that point.

Descriptive data on lesson progress for the Colorado school, including placement at both the beginning and end of the school year, are in Table A-2. Data were available for 63 students on both TCAP scores and progress in *Reading Mastery*. The average student completed more than 200 lessons over the year, substantially more lessons than comprise one year's study. Thus the students were advancing at a highly accelerated pace. However, given their

[11] Beginning of the year data on lesson progress were available only for those in grades 3 through 5. There were a few cases (less than 5 for each grade) where the beginning of the year placement was higher than the end of the year placement. These cases were omitted, assuming that the initial starting point involved an error in placement.

very low initial placements, even this accelerated pace was not sufficient to bring the majority to grade level. Slightly more than half of the Colorado students scored at the proficient level or higher on the TCAP.

Data on progress in *RMSE* for students in the inner city school are given in Table A-3. School staff provided information on lesson placement in reading for the start of the 2013-14, and these data were used to calculate students' cumulative lesson placement in the program in the spring of 2013. (Note that these data are limited to the students who returned to the school for the 2013-14 year and who were in grades one to four in 2012-13.) For instance, for first grade, the average number of lessons completed was 274 (the mean in column 2), although half of the students completed 295 or more lessons (the median in column 3). Data indicate that, at each grade, the average student was nearing the end of their grade material by the spring. (Compare Table A-3 with Table 1 in the text.) However, there was, at each grade, a fair amount of variability in the final lesson placement. In addition, there was variability across grades. While less than a third of the first grade students were at the end of the year goal of 320 lessons, over half of the students in third and fourth grade met their grade level goals. As noted in the text, this may indicate that students at the school, which has the highest rates of poverty of those examined, were gradually catching up.

Lesson Progress and Meeting Normative Standards in Reading

Table A-4 reports data related to two state assessments, the STAAR and the TCAP, comparing results for students who were at or nearing their assigned grade level in *RMSE* and those who were below that level. The top panel reports results for the STAAR and the second panel reports results for the TCAP. The first column of data gives the percentage of students who were not at the targeted lesson and met the state defined proficiency level; the second column gives the percentage of those who were at or near (for the Colorado students) the targeted lesson and met the proficiency standard. The third and fourth columns of data give the results of a chi-square test, assessing the extent to which the results would occur by chance, and the final column reports an effect size (Cohen's d) associated with the difference. In all of the comparisons the students who were at or nearing grade level were far more likely to reach the proficient level on the state assessments. All of the comparisons were statistically significant and the effect sizes were large.

Table A-5 parallels the data reported in Table A-4, but focuses on the percentage of students who met or surpassed the national mean on the MAP. Results are given in the first panel for the Texas schools and in the second panel for the inner city school. In all comparisons students who were approaching or at grade level were more likely to meet or exceed the national mean score on the MAP. All of the comparisons for the Texas schools were

statistically significant ($p < .001$ in all cases), and all of the associated effect sizes were large, ranging from .71 to 1.29. While all of the effect sizes for the inner city site exceeded the criterion used to establish educational importance (ranging from .29 to 1.39), those for grades 2 and 3 failed to reach standard levels of statistical significance, no doubt because of the relatively small sample size.

Value Added Analyses – Reading

Tables A-6 through A-9 give detailed results for the regressions of scale scores on lesson progress. Table A-6 gives results for the analysis of the STAAR exam, Table A-7 gives results for the MAP data for the Texas schools, Table A-8 gives results for the MAP data for the inner city school, and Table A-9 gives results for the TCAP data from the Colorado school. These analyses were designed to examine the relationship of students' assessment scores to their progress through Reading Mastery (defined as their ending lesson for the school year), while controlling for their achievement at the start of the year. Results were calculated separately for each grade and assessment resulting in 17 different analyses.

The measures used as controls differed slightly from one analysis to another, but all provide a strong indication of students' learning at the start of the school year. For instance, for the analysis of the STAAR data (Table A-6) students' STAAR scores in the previous year were used for the analysis for students in grades 4 and 5. Because the third graders had no previous STAAR exposure, their fall MAP scores were used as a control. For the analyses of spring MAP scores (Tables A-7 and A-8), fall MAP scores were used as the control variable. For the analyses of TCAP scores (Table A-9), no prior test data were available. However, data were available on students' placement in *RMSE* at the beginning of the year, indicating their reading skills at that point, and this placement was used as the control variable.[12] The analyses of the Texas data included free and reduced lunch status and English proficiency as control variables, but these had little influence on the results and did not alter the conclusions regarding the impact of placement in *RMSE*.

The fourth column of data in each table reports the zero-order (without controls) correlations between the scale score (the raw STAAR, TCAP, and MAP scores) and the variables in the analysis. All of these correlations are positive and, in most cases, quite large, indicating a strong positive association between students' scale scores and their placement within the *RMSE* program. The first, second, and third columns of data report the results of the

[12] The fall and spring RM placements were correlated at $r = .9157$, indicating that about 84% of the variance was held in common. This collinearity increases the standard error associated with the coefficient estimates. (Note that the larger standard error is considered in the calculation of the tests of significance.) Because the results so strongly replicate those at the other sites, they were included. However, because they were much larger than those with the other sites, they have not been emphasized in the text or included in calculations of the average value added effect.

regression analyses, including the measure of students' prior achievement and their placement in *RMSE* in the spring of the year. The first column reports the unstandardized regression coefficients, the second reports the standardized regression coefficients (beta weights), and the third reports the t-ratios associated with each coefficient and the associated probability.

Because both the correlation coefficient and the standardized regression coefficient are in standard deviation units, they can be directly compared. The correlation coefficient can be interpreted as the "gross" or "total" effect of a variable on the achievement score, while the beta weight can be interpreted as the "net" effect once the other variables in the equation have been equalized or controlled. Thus, the beta weight associated with the cumulative lesson at spring can be seen as the influence of progress through *RMSE* net or independent of students' prior achievement. These are the values reported in the figures shown in the main body of the text.

Of the 17 standardized coefficients, only one was negative, All but four were statistically significant. The four non-significant results were in the smaller inner city and Colorado settings. The average effect size across the analyses was .37, when the Colorado results were included, and .21 when they were not included. This provides strong support for the conclusion that progress in *RMSE* is related to higher scale scores and this relationship is independent of prior learning and achievement.

Mathematics

As with the analysis of reading, this section is divided into three parts: The first gives descriptive information on the schools in the analysis. The second provides the results regarding the association of lesson progress in *CMCCE* to scoring at the proficient level on state exams or above the national mean on the MAP. The third discusses details on the value added analysis for measures of mathematics achievement.

Descriptive Data

For the Colorado school data were available for 64 students on both TCAP scores and progress in *Connecting Math Concepts: Comprehensive Edition (CMCCE)*. Table A-10 reports data on lesson placement in *CMCCE* of third grade students at the Colorado school in the fall of 2012 and the spring of 2013 using the cumulative measure of progress described in Table 2 in the text. The top rows report the average lesson number, minimum, maximum, and standard deviation. It can be seen that the average student was in Grade 2 material at both the start and the end of the year, well below grade level. There was relatively little variability, and few of the students appeared to have made substantial progress. The average student completed only 57 lessons (range from 48 to 180 with a standard deviation

of only 16). Thus, in contrast to the results with reading, the lesson progress in mathematics was far less than optimal.

Table A-11 reports data on the cumulative lessons completed by students in spring 2013 for each grade in the inner city school. While students' progress was superior to that in the Colorado school, in all grades the average values were far below the end of year lesson for their grade. While 59 percent of the first graders completed the math program, only two second graders and no third or fourth graders did so. As with reading, there was a fair amount of variability within each grade in the amount of lesson progress.

Table A-12 gives descriptive data for the measures of lesson progress in *CMCCE* and mathematics assessment scores for the students in the Texas schools. Data are given on scale scores on the STAAR (grades 3 to 5) and MAP (grades K to 5), the percentage of students meeting the STAAR level of proficiency or meeting or surpassing the national mean for the MAP, and lesson placement in CMCCE. Note that because so few fourth and fifth grade students were at grade level in CMCCE at the end of the year the percentage at grade level reported represents, for fifth grade students, the percentage who were at least 60% of the way through the program in the spring and, for fourth grade students, the percentage who were at least 90% of the way through the grade level program at that time.

Lesson Progress in CMC and Meeting Normative Standards in Mathematics

Tables A-13, A-14, and A-15 compare the percentage of students scoring at the proficient level or higher in the STAAR (Table A-13) and TCAP (Table (A-14) or above the national norms on the MAP (Tables A-13 and A-15) for those who were at or near grade level in *CMCCE* and those below that point. In all cases the percentages were substantially higher for those placed higher in *CMCCE*. All but one comparison (third graders in the inner city on the MAP) were statistically significant, and all of the comparisons had associated effect sizes that were large and far beyond the level considered educationally important, ranging from .48 to 1.52.

Value Added Analysis - Mathematics

The results for the value added analysis for mathematics are in the remaining tables. Results for the STAAR are in Table A-16, those for the MAP in the Texas Schools are in Table A-17, those for the MAP for the inner city school are in Table A-18, and those for the TCAP are in Table A-19. The beginning and ending lessons in *CMCCE*, which were used as predictors for the TCAP analysis, were very highly correlated (r = .97). Thus these results were not discussed in the text or used in the computation of average effect sizes. They are included in this appendix only for informational purposes.

As discussed in the text, the results of the value added analysis replicate the results with reading. All but one of the 13 coefficients associated with the spring lesson placement were positive. Coefficients ranged from -.04 to .37, with an average of .21. All but two (both in the inner city school) were statistically significant.

Table A-1

Descriptive Statistics Reading Measures, Texas Charter Schools, by Grade

Kindergarten Students					
	Mean	SD	Min.	Max	N
Fall Reading Scores					
MAP Reading Fall	138.9	9.9	111	183	825
MAP Greater than Norm, Fall	0.38	0.48	0.00	1.00	825
Spring Reading Scores					
MAP Reading Spring	151.8	11.1	121	193	825
MAP Greater than Norm, Spring	0.33	0.47	0.00	1.00	825
Lesson Progress					
Cum. Reading Lesson, EOY	154.6	38.3	34	372	999
> = Target Lesson, EOY	0.46	0.50	0.00	1.00	999
First Grade Students					
	Mean	SD	Min.	Max	N
Fall Reading Scores					
MAP Reading Fall	153	12.1	117	193	1076
MAP Greater than Norm, Fall	0.29	0.46	0.00	1.00	1076
Spring Reading Scores					
MAP Reading Spring	166	12.4	123	202	1076
MAP Greater than Norm, Spring	0.2	0.4	0.00	1.00	1076
Lesson Progress					
Cum. Reading Lesson, EOY	280.5	78.1	45	602	1043
> = Target Lesson, EOY	0.40	0.49	0.00	1.00	1043
Second Grade Students					
	Mean	SD	Min.	Max	N
Fall Reading Scores					
MAP Reading Fall	167.7	15.1	125	220	1032
MAP Greater than Norm, Fall	0.34	0.47	0.00	1.00	1032
Spring Reading Scores					
MAP Reading Spring	180.8	13.8	138	220	1032
MAP Greater than Norm, Spring	0.28	0.45	0.00	1.00	1032
Lesson Progress					
Cum. Reading Lesson, EOY	420.1	91.8	28	644	1019
> = Target Lesson, EOY	0.49	0.5	0.00	1.00	1019
Third Grade Students					
	Mean	SD	Min.	Max	N

Fall Reading Scores					
MAP Reading Fall	182.2	15.9	141	230	713
MAP Greater than Norm, Fall	0.36	0.48	0.00	1.00	713
Spring Reading Scores					
MAP Reading Spring	192.9	16.2	147	241	713
STAAR Scale Score	1412.8	132	1079	1909	724
MAP Greater than Norm, Spring	0.40	0.49	0.00	1.00	713
STAAR at Proficient Level	0.72	0.45	0.00	1.00	724
Lesson Progress					
Cum. Reading Lesson, BOY	415	108.6	97	665	724
= > Target Lesson, BOY	0.5	0.5	0.00	1.00	724
Cum. Reading Lesson, EOY	567.5	95.3	213.3	765.9	724
> = Target Lesson, EOY	0.40	0.49	0.00	1.00	724

Fourth Grade Students

	Mean	SD	Min.	Max	N
Fall Reading Scores					
MAP Reading Fall	190.3	16.3	143	228	563
STAAR Scale Score Spring 2012	1398	131.4	736	1771	552
MAP Greater than Norm, Fall	0.32	0.47	0.00	1.00	563
STAAR at Proficient Level, Spring 2012	0.71	0.45	0.00	1.00	552
Spring Reading Scores					
MAP Reading Spring	200	15.503	153	233	563
STAAR Scale Score	1461	128.6	811	1855	552
MAP Greater than Norm, Spring	0.41	0.49	0.00	1.00	563
STAAR at Proficient Level	0.61	0.49	0.00	1.00	552
Lesson Progress					
Cum. Reading Lesson, BOY	503.4	138.5	2	726	552
= > Target Lesson, BOY	0.43	0.49	0.00	1.00	552
Cum. Reading Lesson, EOY	662	105.5	262.6	847.7	552
> = Target Lesson, EOY	0.27	0.44	0.00	1.00	552

Fifth Grade Students

	Mean	SD	Min.	Max	N
Fall Reading Scores					
MAP Reading Fall	200.3	15.8	149	231	343
STAAR Scale Score Spring 2012	1489.9	119.2	1181	1780	346
MAP Greater than Norm, Fall	0.39	0.49	0.00	1.00	343
STAAR at Prof. Level, Spring 2012	0.69	0.46	0.00	1.00	346

Spring Reading Scores					
MAP Reading Spring	207.5	15.3	153	240	343
STAAR Scale Score	1535.7	109.7	1216	2021	346
MAP Greater than Norm, Spring	0.43	0.5	0.00	1.00	343
STAAR at Proficient Level	0.83	0.38	0.00	1.00	346
Lesson Progress					
Cum. Reading Lesson, BOY	598.4	162.5	2	726	338
= > Target Lesson, BOY	0.3	0.46	0.00	1.00	338
Cum. Reading Lesson, EOY (90%)	729	107.9	230	844	339
> = Target Lesson, EOY (90%)	0.17	0.37	0.00	1.00	339

Note: For students in third and fourth grade the DI placements for fall, 2013, were available for only about three-fifths of the students. To increase the sample size, the fall 2013 lesson placements were regressed on the spring scores and the resulting equation was used to get a predicted spring value for all students. The R squared value for these predictions was .93 for both the third grade and fourth grade. For fifth graders the only final lesson placement data was for spring of 2013. Only 8% of the students in fifth grade were on a trajectory to be at the target lesson at the end of the year, so the percentage on target at spring represents a projection of being at 90% of the end of year target (lesson 820 in April and higher) .

Table A-2

Lesson Placement, Reading Mastery Signature Edition, Third Graders, Colorado School, Fall 2012 and Spring 2013

	Fall, 2012	Spring, 2013
Mean	205	419
Median	161	451.5
Minimum	1	50
Maximum	466	654
S.D.	155	156
More than One Year Behind Grade Level (n)	36	31
Within One Year of Grade Level (n)	19	24
At Grade Level (n)	8	8

Data were available for 64 students. The category of within one year of grade level includes students who were in 2nd grade material at the start of the year and in 3rd grade material at the end of the year (and when they would be starting fourth grade). More than one year behind grade lavel includes students who were in kindergarten or first grade material at the start of the year and in material in 2nd grade or lower at the end of the year. Those who were at grade level began the year in third grade material and ended the year ready to start fourth grade material.

Table A-3

Ending Lesson in Reading, Inner City School, Spring 2013, by Grade

Grade	Mean	Median	Min	Max	Goal	Met Goal (%)	N
1	274	295	191	411	320	30.8	65
2	412	411	68	585	465	39.3	56
3	579	611	342	696	605	54.9	51
4	688	726	342	726	725	55.1	49

Note: The "Goal" refers to the ending lesson at each grade level. Differences by grade in the percentage of students meeting the goal were statistically significant (chi-square = 10.04, p = .02)

Table A-4

Relationship of On-Target Lesson Progress in Reading Mastery Signature Edition (RMSE) to Scoring at Proficient Level, STAAR, by Grade, and TCAP

Percent At Proficient Level or Higher, STAAR, by Grade)

	Lesson Progress				
Grade	Not at target	At or beyond Target	Chi-square	Prob.	Effect Size
Third Grade	55.9	97.2	148.64	<.001	1.25
Fourth Grade	51.1	88.6	64.35	<.001	0.92
Fifth Grade	79.8	100	13.85	<.001	1.01

Percent at Proficient Level or Higher, TCAP

	One year or more Behind Grade Level	Less Than One Year Behind Grade Level	Chi Square	Prob.	Effect Size
Total Score	29	88	22.21	<.001	1.52
Fiction/Poetry	42	81	10.32	0.001	0.88
Non-Fiction	35	78	11.69	0.001	0.96
Vocabulary	35	88	18.07	<.001	1.32
N	31	32			

For Texas the “target” level was the last lesson for the grade for grades K-4 and 90% of the year’s curriculum for Grade 5. For the Colorado school “at or near target” was defined as being at any point in grade 3 material.

Table A-5

Relationship of On-Target Lesson Progress in Reading Mastery Signature Edition to MAP Scores at or above National Average, Texas Charter Schools and Inner City School, by Grade

Texas Schools

	Lesson Progress				
Grade	Below Goal	At or Above Goal	Chi-square	Prob.	Effect Size
Kindergarten	15.4	52.7	130.06	<.001	0.87
First Grade	2.9	45.9	300.45	<.001	1.29
Second Grade	6.9	50.2	241.83	<.001	1.15
Third Grade	19	71.8	198.49	<.001	1.25
Fourth Grade	32.1	65.5	50.65	<.001	0.71
Fifth Grade	35.1	82.8	44.65	<.001	1.12

Inner City School

	Lesson Progress				
Grade	Below Goal	At or Above Goal	Chi-square	Prob.	Effect Size
First Grade	2.3	47.4	19.76	<.001	1.39
Second Grade	32.4	50.0	1.74	0.19	0.36
Third Grade	22.7	35.7	0.99	0.32	0.29
Fourth Grade	4.6	34.6	6.53	0.01	0.88

Note: Numbers under "Lesson Progress" tell the percentage of students who were at or above the MAP National Average for those who were below the reading goal for their grade (first data column) or at or above this goal (second data column). The “on target” criterion was the last lesson for the grade level for all instances, except for fifth graders in Texas. For that group the criterion was 90% of the lessons within the grade level curriculum.

Table A-6

Regressions of Spring STAAR Reading Scale Scores on Prior Scores, Free and Reduced Lunch, Limited English Proficiency, and Lesson Progress, Texas Charter Schools, by Grade

Third Grade				
	b	beta	t-ratio	R
Fall MAP	3.67	0.44	13.49***	0.74
FARM	-25.41	-0.07	-3.08**	-0.27
LEP	-4.46	-0.02	-0.65	-0.24
Cumulative Lesson	0.54	0.39	12.09***	0.72
Constant	459.17		11.55***	
R squared	.64***			
Fourth Grade				
	B	Beta	t-ratio	R
STAAR 3rd grade	0.66	0.68	21.70***	0.8
FARM	-2.45	-0.01	-0.28	-0.17
LEP	-7.58	-0.03	-1.05	-0.32
Cumulative Lesson	0.24	0.20	6.61***	0.58
Constant	381.58		9.86***	
R squared	.67***			
Fifth Grade				
	B	Beta	t-ratio	r
STAAR 4th grade	0.63	0.69	14.23***	0.76
FARM	-33.78	-0.11	-2.91**	-0.23
LEP	-0.73	0.00	-0.07	-0.35
Cumulative Lesson	0.08	0.08	1.65	0.53
Constant	567.26		9.89***	
R squared	.58***			

Table A-7

Regressions of Spring MAP Reading Scores on Fall Score, Demographics, and Lesson Progress, Texas Charter Schools

Kindergarten Students				
	b	beta	t-ratio	r
Fall MAP	0.54	0.49	18.74***	0.68
FARM	-1.63	-0.05	-2.31*	-0.18
LEP	-0.47	-0.02	-0.86	-0.26
Cumulative Lesson	0.11	0.38	14.55***	0.62
Constant	60.65		15.79***	
R squared	.56***			
First Grade Students				
	b	beta	t-ratio	r
Fall MAP	0.56	0.55	23.62***	0.78
FARM	-1.23	-0.03	-1.88	-0.18
LEP	0.1	0	0.22	-0.23
Cumulative Lesson	0.06	0.36	15.79***	0.71
Constant	65.34		19.92***	
R squared	.68***			
Second Grade Students				
	b	beta	t-ratio	r
Fall MAP	0.43	0.48	16.33***	0.75
FARM	-2.39	-0.06	-2.95**	-0.21
LEP	-0.26	-0.01	-0.45	-0.24
Cumulative Lesson	0.05	0.34	11.91***	0.71
Constant	88.34		24.55***	
R squared	.62***			
Third Grade Students				
	b	beta	t-ratio	r
Fall MAP	0.56	0.55	16.77***	0.77
FARM	-0.22	-0.01	-0.22	-0.21
LEP	1.08	0.03	1.29	-0.19
Cumulative Lesson	0.05	0.32	9.84***	0.7
Constant	59.99	.	12.28***	
R squared	.64***			
Fourth Grade Students				
	b	beta	t-ratio	r

Fall MAP	0.6	0.63	17.26***	0.74
FARM	-0.14	0	-0.12	-0.13
LEP	-0.05	0	-0.05	-0.26
Cumulative Lesson	0.03	0.18	5.23***	0.57
Constant	69.06		12.17***	
R squared	.57***			

Fifth Grade Students

	b	beta	t-ratio	r
Fall MAP	0.57	0.59	16.26***	0.78
FARM	-2.76	-0.06	-1.9	-0.2
LEP	-1.98	-0.06	-1.7	-0.39
Cumulative Lesson	0.04	0.31	8.39***	0.64
Constant	63.11		9.39***	
R squared	.69***			

Table A-8

Regressions of Spring MAP Reading Scores on Fall Score and Ending RMSE Lesson, by Grade, Inner City

		First Grade		
	b	beta	t-ratio	r
Fall MAP	0.74	0.62	5.87***	0.77
Cum. Lesson	0.06	0.24	2.26*	0.64
Constant	35.26		2.17*	
R squared	.63***			
		Second Grade		
	b	beta	t-ratio	r
Fall MAP	0.26	0.25	1.67	0.30
Cum. Lesson	0.02	0.11	0.75	0.23
Constant	132.51		5.72***	
R squared	0.1			
		Third Grade		
	b	beta	t-ratio	r
Fall MAP	0.46	0.49	3.30**	0.54
Cum. Lesson	0.01	0.08	0.56	0.35
Constant	100.91		4.67***	
R squared	.29***			
		Fourth Grade		
	b	beta	t-ratio	r
Fall MAP	0.79	0.66	4.73***	0.53
Cum. Lesson	-0.02	-0.13	-0.93	0.21
Constant	61.36		2.27*	
R squared	.36***			

Table A-9

Regressions of Spring TCAP Reading Score on RMSE Lesson Placement, Beginning and End of Year, Third Grade, Colorado School

Total Scale Score				
	b	Beta	t-ratio	r
Constant	435.82		18.84***	
Beginning Cum. Lesson	-0.07	-0.2	-0.78	0.52
Ending Cum. Lesson	0.29	0.79	3.10**	0.61
R Squared	.37***			
Fiction and Poetry				
	b	Beta	t-ratio	r
Constant	424.49		15.31***	
Beginning Cum. Lesson	-0.15	-0.35	-1.3	0.43
Ending Cum. Lesson	0.35	0.86	3.20**	0.54
R Squared	.31***			
Non-Fiction				
	b	Beta	t-ratio	r
Constant	415.11		6.02***	
Beginning Cum. Lesson	0.26	0.28	0.94	0.37
Ending Cum. Lesson	0.09	0.10	0.32	0.35
R Squared	.14**			
Vocabulary				
	b	Beta	t-ratio	r
Constant	398.35		9.10***	
Beginning Cum. Lesson	-0.09	-0.14	-0.51	0.45
Ending Cum. Lesson	0.40	0.7	2.32*	0.51
R Squared	.26***			

Table A-10

Lesson Placement, CMCCE, Third Graders, Colorado School, Fall 2012 and Spring 2013 (n=64)

	Fall, 2012	Spring, 2013
Mean	317	375
Median	376	434
Minimum	246	295
Maximum	386	439
S.D.	65	70
More than One Year Behind Grade Level (n)	0	28
Within One Year of Grade Level (n)	29	36
At Grade Level (n)	35	0

Data were available for 64 students. For Fall, 2012 the category of "within one year of grade level" includes students who were in 2nd grade material at the start of the year and the category of "at grade level" includes those who began the year in 3rd grade material. For Spring, 2013, there were no students at grade level. Those within one year of grade level were in third grade material but not at the end of the program. Those who were one year behind grade level were in 2nd grade material.

Table A-11

Ending Lesson in CMCCE, Spring 2013, by Grade, Inner City School

Grade	Mean	Median	Min	Max	Goal	Met Goal (%)	N
1	225	245	120	281	245	59.1	66
2	312	326	181	375	375	3.6	56
3	417	480	203	480	505	0.0	51
4	520	560	203	560	635	0.0	48

Note: The "Goal" refers to the ending lesson at each grade level. Differences by grade in the percentage of students meeting the goal were statistically significant (chi-square = 102.65, $p < .001$).

Table A-12

Descriptive Statistics on All Measures, Math Analysis, by Grade, Texas Schools

Kindergarten Students					
Fall Math Scores	Mean	SD	Min.	Max.	N
MAP Math Fall	139.8	11.0	104.0	177.0	865
MAP Greater than Norm, Fall	0.37	0.48	0.00	1.00	865
Spring Math Scores					
MAP Math Spring	156.2	12.1	116.0	192.0	865
MAP Greater than Norm, Spring	0.43	0.50	0	1	865
Lesson Progress					
Cumulative Math Lesson	107.6	33.3	20.0	265.0	865
> = Target Lesson, EOY	0.44	0.50	0	1	865
First Grade Students					
Fall Math Scores	Mean	SD	Min.	Max.	N
MAP Math Fall	158.8	12.0	117.0	216.0	1131
MAP Greater than Norm, Fall	0.42	0.49	0.00	1.00	1131
Spring Math Scores					
MAP Math Spring	173.6	12.3	124.0	223.0	1131
MAP Greater than Norm, Spring	0.45	0.50	0.00	1.00	1131
Lesson Progress					
Cumulative Math Lesson	238.5	66.5	20.0	742.0	1131
> = Target Lesson, EOY	0.45	0.50	0.00	1.00	1131
Second Grade Students					
Fall Math Scores	Mean	SD	Min.	Max.	N
MAP Math Fall	170.9	12.5	135.0	206.0	1048
MAP Greater than Norm, Fall	0.40	0.49	0.00	1.00	1048
Spring Math Scores					
MAP Math Spring	184.3	11.5	144.0	220.0	1048
MAP Greater than Norm, Spring	0.30	0.46	0.00	1.00	1048
Lesson Progress					
Cumulative Math Lesson	343.7	78.7	31.0	636.0	1048
> = Target Lesson, EOY	0.41	0.49	0.00	1.00	1048
Third Grade Students					
Fall Math Scores	Mean	SD	Min.	Max.	N
MAP Math Fall	185.4	13.0	138.0	225.0	752
MAP Greater than Norm, Fall	0.33	0.47	0.00	1.00	752

Spring Math Scores					
MAP Math Spring	201.1	12.0	143.0	235.0	752
STAAR Scale Score	1428.4	136.7	1123.0	1927.0	764
MAP Greater than Norm, Spring	0.50	0.50	0.00	1.00	752
STAAR at Proficient Level	0.60	0.49	0.00	1.00	764
Lesson Progress					
Cumulative Math Lesson, BOY	319.9	89.8	1.00	536.0	752
>= Target Lesson, BOY	0.35	0.48	0.00	1.00	752
Lesson Progress (completed in year)	147.6	44.4	14.4	466.9	752
> = Target Lesson, EOY	0.37	0.48	0.00	1.00	752
Fourth Grade Students					
Fall Math Scores	Mean	SD	Min.	Max.	N
MAP Math Fall	198.3	11.6	155.0	227.0	558
STAAR Scale Score Spring 2012	1427.0	136.9	770.0	1942.0	496
MAP Greater than Norm, Fall	0.40	0.49	0.00	1.00	558
STAAR at Proficient Level, Spring 2012	0.59	0.49	0.00	1.00	496
Spring Math Scores					
MAP Math Spring	207.3	12.4	158.0	241.0	558
STAAR Scale Score	1477.5	124.9	825.0	1876.0	496
MAP Greater than Norm, Spring	0.37	0.48	0.00	1.00	558
STAAR at Proficient Level	0.56	0.50	0.00	1.00	496
Lesson Progress					
Cumulative Math Lesson, BOY	424.9	80.8	1.0	636.0	496
= > Target Lesson, BOY	0.24	0.42	0.00	1.00	496
Cumulative Math Lesson, EOY	541.8	75.5	200.4	742.5	496
> = 90% of Target Lesson, EOY	0.41	0.49	0.00	1.00	496
Fifth Grade Students					
Fall Math Scores	Mean	SD	Min.	Max.	N
MAP Math Fall	209.0	10.6	173.0	240.0	331
STAAR Scale Score Spring 2012	1506.5	120.1	974.0	2016.0	316
MAP Greater than Norm, Fall	0.42	0.49	0.00	1.00	331
STAAR at Proficient Level, Spring 2012	0.64	0.48	0.00	1.00	316
Spring Math Scores					
MAP Math Spring	219.4	12.7	153.0	250.0	331
STAAR Scale Score	1569.5	113.6	1290.0	2064.0	316
MAP Greater than Norm, Spring	0.47	0.50	0.00	1.00	331
STAAR at Proficient Level	0.85	0.36	0.00	1.00	316

Lesson Progress					
Cumulative Math Lesson, BOY	526.2	86.6	121.0	636.0	331
= > Target Lesson, BOY	0.24	0.43	0.00	1.00	331
Cumulative Math Lesson, April	617.2	83.0	220.0	718.0	331
> = Target Lesson, EOY (>=60%)	0.31	0.46	0.00	1.00	331

Note: For fifth grade no data were available for the end of the year so data on the lesson at April was used, calculation projections, based on progress to date, on the end of year lesson. No students in fifth grade were on a trajectory to be at the target lesson at the end of the year (the maximum lesson of 718 represents only 69% of the target), so the data given represent the proportion of students who were at least 60% of the way through the program in April (lesson 708 and higher). For fourth graders the percentage given represents an estimate of students who would be at 90% of the target at the end of the school year.

Table A-13

Relationship of On-Target Lesson Progress in CMCCE to MAP Scores at or above National Mean and Scoring Proficient on the STAAR, by Grade, Texas Schools

Comparisons to MAP Norms (% At or Above Norms)					
	Lesson Progress				
Grade	Not at target (%)	At or beyond Target (%)	Chi-square	Prob.	Effect Size
Kindergarten (NWF)	26.6	55.3	85.13	<.001	0.61
First Grade	25.4	68.2	207.21	<.001	0.95
Second Grade	19.1	45.6	84.78	<.001	0.59
Third Grade	35	75.3	113.11	<.001	0.89
Fourth Grade	24.4	56.3	58.3	<.001	0.69
Fifth Grade*	36	71.8	36.66	<.001	0.77
Comparisons to STAAR Proficiency Levels					
	Lesson Progress				
Grade	Not at target (%)	At or beyond Target (%)	Chi-square	Prob.	
Third Grade	45	87.4	132.62	<.001	1.02
Fourth Grade	41.6	76.1	58.11	<.001	0.75
Fifth Grade*	78.8	98.9	20.73	<.001	0.78

Table A-14

Percentage of Students Passing Mathematics TCAP by Lesson Placement in CMCCE, End of Year, Third Graders, Colorado, Spring 2013

	Grade 2	Grade 3	Chi Square	Prob.	Effect Size
Total Score	7	50	13.46	<.001	1.14
Standard 1	7	53	14.88	<.001	1.22
Standards 2 & 3	11	56	13.76	<.001	1.11
Standards 4 & 5	18	39	3.34	0.07	0.48
Standard 6	7	47	12.12	<.001	1.06
N	28	36			

Note: None of the students were at grade level in CMC at the end of the year. Thus the comparison is between those who were somewhere in grade 3 material and those who were still in grade 2 material. The standards are defined as "Standard One: Students develop number sense and use numbers and number relationships in problem-solving situations and communicate the reasoning used in solving these problems. Standard Two: Students use algebraic methods to explore, model, and describe patterns and functions involving numbers, shapes, data, and graphs in problem-solving situations and communicate the reasoning used in solving these problems. Standard Three: Students use data collection and analysis, statistics, and probability in problem-solving situations and communicate the reasoning used in solving these problems. Standard Four: Students use geometric concepts, properties, and relationships in problem-solving situations and communicate the reasoning used in solving these problems. Standard Five: Students use a variety of tools and techniques to measure, apply the results in problem-solving situations, and communicate the reasoning used in solving these problems. Standard 6: Students link concepts and procedures as they develop and use computational techniques, including estimation, mental arithmetic, paper-and-pencil, calculators, and computers, in problem-solving situations and communicate the reasoning used in solving these problems." (TCAP, 2013).

Table A-15

Relationship of On-Target Lesson Progress in CMCCE to MAP Scores at or above Norm, by Grade, Inner City School

Comparisons to MAP Norms (% At or Above Norms)					
	Lesson Progress				
Grade	Well Below target (%)	At or Approaching Target (%)	Chi-square	Prob.	Effect Size
First Grade	11	45	8.38	0.004	0.84
Second Grade	0	26	4.56	0.03	1.19
Third Grade	0	11	2.60	0.11	0.70
Fourth Grade	0	22	5.09	0.02	1.06

Note: For second and third grade the lesson progress target is only at the mid point of the given grade level; for fourth grade it is at any point within fourth grade. Only 2 students in second grade and none in third and fourth grade finished their grade level programs in CMC.

Table A-16
Regressions of 2013 STAAR Mathematics Scale Scores on 2012 Scores, Demographic Characteristics, and Lesson Progress in CMCCE, by Grade, Texas Schools

Third Grade				
	b	beta	t	r
Fall MAP	5.25	0.5	18.12***	0.7
FARM	-13.51	-0.04	-1.55	-0.22
LEP	-4.82	-0.02	-0.69	-0.18
Cum Lesson	0.53	0.37	13.75***	0.63
Constant	221.31		4.43***	
R squared	.60***			
Fourth Grade				
	b	beta	t	r
STAAR 3rd grade	0.61	0.67	21.84***	0.77
FARM	-9.54	-0.03	-1.1	-0.1
LEP	-7.08	-0.03	-0.93	-0.18
Lessons Completed	0.38	0.23	7.69***	0.52
Constant	414.6		10.74***	
R squared	.64***			
Fifth Grade				
	b	beta	t	r
STAAR 4th grade	0.6	0.64	14.05***	0.73
FARM	-23.17	-0.07	-1.85	-0.13
LEP	-15.87	-0.06	-1.51	-0.23
Cum Lesson	0.19	0.14	3.19**	0.48
Constant	565.69		9.89***	
R squared	.56***			

Note: For third grade the MAP scores for the fall were used as a control because there were no STAAR scores for the prior year.

Table A-17

Regressions of Spring MAP Mathematics Scores on Fall Scores, Demographic Characteristics, and Progress in CMCCE, by Grade, Texas Schools

Kindergarten				
	b	beta	t	r
Fall MAP	0.77	0.7	29.48***	0.75
FARM	-0.68	-0.02	-0.88	-0.14
LEP	-0.42	-0.02	-0.74	-0.18
Cum. Lesson	0.05	0.13	5.65***	0.32
Constant	44.08		11.86***	
R squared	.57***			
First Grade				
	b	beta	t	r
Fall MAP	0.65	0.63	29.45***	0.77
FARM	-1.21	-0.03	-1.77	-0.18
LEP	-0.3	-0.01	-0.65	-0.16
Cum. Lesson	0.04	0.24	11.08***	0.6
Constant	60.44		18.47***	
R squared	.64***			
Second Grade				
	b	beta	t	r
Fall Scores	0.54	0.59	24.49***	0.7
FARM	-1.14	-0.03	-1.55	-0.16
LEP	-0.66	-0.03	-1.32	-0.14
Cum. Lesson	0.03	0.23	9.80***	0.51
Constant	82.15		23.03***	
R sq.	.54***			
Third Grade				
	b	beta	t	r
Fall Scores	0.56	0.6	22.91***	0.76
FARM	-0.5	-0.02	-0.69	-0.21
LEP	-0.6	-0.02	-1.03	-0.2
Cum. Lesson	0.04	0.29	11.09***	0.6
Constant	82.04		19.64***	

R sq.	.63***			
		Fourth Grade		
	b	beta	t	r
Fall Scores	0.68	0.64	19.95***	0.74
FARM	-0.83	-0.03	-0.89	-0.11
LEP	-1.8	-0.07	-2.41*	-0.25
Cum. Lesson	0.03	0.18	5.66***	0.49
Constant	59.08		9.60***	
R sq.	.58***			
		Fifth Grade		
	b	beta	t	r
Fall Scores	0.79	0.67	16.20***	0.76
FARM	-2.16	-0.06	-1.6	-0.15
LEP	-2.61	-0.09	-2.42*	-0.27
Cum. Lesson	0.02	0.13	3.35***	0.46
Constant	43.35			
R sq.	.60***			

Table A-18

Regression of Spring MAP Mathematics Scores on Fall Scores and Progress in CMCCE, by Grade, Inner City School

First Grade				
	b	beta	t	r
Fall MAP	0.74	0.49	4.22***	0.57
Cum. Lesson	0.04	0.21	1.82	0.41
Constant	34.02		1.32	
R squared	.37***			
Second Grade				
	b	beta	t	r
Fall Scores	0.67	0.75	7.06***	0.73
Cum. Lesson	-0.01	-0.04	-0.43	0.31
Constant	71.21		5.03***	
R sq.	.54***			
Third Grade				
	b	beta	t	r
Fall Scores	0.75	0.59	5.10***	0.74
Cum. Lesson	0.04	0.26	2.20*	0.60
Constant	40.97		1.74	
R sq.	.60***			
Fourth Grade				
	b	beta	t	r
Fall Scores	0.55	0.59	6.09***	0.87
Cum. Lesson	0.06	0.37	3.75***	0.81
Constant	60.72		5.00***	
R sq.	.81***			

Table A-19

Regression of TCAP Mathematics Scores on Fall and Spring Lesson Placement in CMCCE, Total Scale Score and Subtests

Total Score				
	b	beta	t	r
Fall Cum Lesson	-0.08	-0.07	-0.14	0.54
Spring Com. Lesson	0.68	0.62	1.36	0.56
Constant	154.19		3.37***	
R squared	.31***			
Standard One				
	b	beta	t	r
Fall Cum Lesson	0.73	0.52	1.1	0.52
Spring Com. Lesson	0.007	0.005	0.01	0.51
Constant	156.67		2.74**	
R sq.	.27***			
Standards Two and Three				
	b	beta	t	r
Fall Cum Lesson	-0.04	-0.03	-0.06	0.49
Spring Com. Lesson	0.68	0.53	1.12	0.50
Constant	143.44		2.57**	
R sq.	.25***			
Standards Four and Five				
	b	beta	t	r
Fall Cum Lesson	-0.26	-0.18	-0.35	0.37
Cum. Lesson	0.77	0.56	1.11	0.39
Constant	163.6		2.57**	
R sq.	.15***			
Standard Six				
	b	beta	t	r
Fall Cum Lesson	-0.83	-0.51	-1.09	0.46
Cum. Lesson	1.51	1.00	2.13*	0.50
Constant	78.73		1.21	
R sq.	.27***			

Note: The fall and spring lesson placement were very highly correlated (r = .97), resulting in highly inflated standard errors. The table is provided only for informational purposes.

References

Adams G., & Engelmann, S. (1996). *Research on Direct Instruction: 25 years beyond DISTAR.* Seattle, WA: Educational Achievement Systems.

Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, *73*(2), 125–230.

Briggs, D.C. (2001). The effect of admissions test preparation: Evidence from NELS:88. Chance, 14 (1), 10-18.

Colorado Department of Education (2011). Transitional Colorado Assessment Program (TCAP) Assessment Framework, Mathematics – Grade 3. http://www.cde.state.co.us/sites/default/files/documents/assessment/documents/frameworks/tcap_math_3rd%20grade.pdf

Coughlin, C. (2014). Outcomes of Engelmann's Direct Instruction: Research Syntheses, pp. 25-54 in J. Stockard (Ed.). The Science and Success of Engelmann's Direct Instruction. Eugene, OR: NIFDI Press.

Domingue, B. & Briggs, D.C. (2009). Using linear regression and propensity score matching to estimate the effect of coaching on the SAT. University of Colorado.

Engelmann, S.E. (2014a). Research from the Inside: The Development and Testing of DI Programs, pp. 3-24 in J. Stockard (Ed.). The Science and Success of Engelmann's Direct Instruction. Eugene, OR: NIFDI Press.

Engelmann, S.E. (2014b). *Successful and Confident Students with Direct Instruction*. Eugene, OR: NIFDI Press.

Engelmann, S.E. & Engelmann, K.E. (2004). Impediments to scaling up effective comprehensive school reform models. Pp. 107-133 in T.K. Glennan, Jr., S.J. Bodilly, J.R. Galegher, & K.A. Kerr (Eds.) Expanding the reach of education reforms: Perspectives from leaders in the scale-up of educational interventions). Santa Monica, CA: Rand Corporation.

Hattie, John A.C. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. London and New York: Routledge.

Liem, A., & Martin, A. (2013). Direct Instruction. In John Hattie and Eric M. Anderman (Eds.), *International guide to student achievement* (pp. 366-368). New York: Routledge.

Northwest Evaluation Association. (2011a). RIT Scale Norms: For Use with Measures of Academic Progress (MAP®) and MAP® for Primary Grades. Portland, OR: Author.

Northwest Evaluation Association. (2011b). Technical Manual for Measures of Academic Progress (MAP) and Measures of Academic Progress for Primary Grades (MPG). Portland, OR: Author.

Przychodzin, A. M., Marchand-Martella, N. E., Martella, R. C., & Azim, D. (2004). Direct Instruction mathematics programs: An overview and research summary. *Journal of Direct Instruction, 4*(1), 53–84.

Schieffer, C., Marchand-Martella, N. E., Martella, R. C., Simonsen, F. L., & Waldron-Soler, K. M. (2002). An analysis of the *Reading Mastery* program: Effective components and research review. *Journal of Direct Instruction, 2*(2), 87–119.

Stockard, J. (2011a). Direct Instruction and first grade reading achievement: The role of technical support and time of implementation. *Journal of Direct Instruction*, *11*(1), 31-50.

Stockard, J. (2011b). Increasing reading skills in rural areas: An analysis of three school districts. *Journal of Research in Rural Education*, *26*(8), 1-19.

Tallmadge, G. K. (1977). *The Joint Dissemination Review Panel ideabook.* Washington, DC: National Institute of Education and U.S. Office of Education.

Vitale, M. & Joseph, B. (2008). Broadening the institutional value of Direct Instruction Implemented in a low-SES elementary school: Implications for scale-up and school reform. *Journal of Direct Instruction, 8*(1), 1-18