Spring 2014

Measuring Student Gains on the Connecticut Mastery Test

Taylor K. Godfrey
Trinity College, taylor.godfrey@trincoll.edu

Measuring Student Gains on the Connecticut Mastery Test

Taylor Godfrey

Trinity College

Senior Research Seminar, Ed400

Fall 2013

Final Paper

Acknowledgements

I would like to acknowledge several people for their help throughout my research project. First, I would like to thank Dr. George Michna for granting me permission to work with the dataset I used to conduct my research. Second, I would like to thank Professor Jack Dougherty for providing me with this dataset, as well as his guidance and support throughout my entire project.

I would also like to thank Rob Walsh for his input in regards to my research articles for my literature review, and his participation in my final presentation. Additionally, I would like to thank Rachael Barlow and Professor Reuman for their advice in regards to working through various SPSS calculations. I would especially like to thank Rachael for her consistent availably to meet and speak with me in regards to my calculations.

Finally, I would like to thank Professor Rachel Leventhal-Weiner for her endless support for my project throughout the entire semester. Whether it was meeting with her three times in one week to perfect my final presentation, or determining how to put my passion for this research into words, she was always available for help and advice and I could not have completed this project without her.

Abstract

The implementation of No Child Left Behind (NCLB) in 2002 led to an accountability system in schools based on high-stakes standardized tests (Orlich, 2004). NCLB measures schools' performance on tests based on the percentage of students performing at proficiency level. Due to the need to reach proficiency, Connecticut has placed an emphasis on reporting static yearly test scores on the Connecticut Mastery Test (CMT), without indicating gains made by students as they transition from one grade level to the next. This research study used independent student-level data to calculate gains on the CMT across three continuous school years for students in grades three to eight. It was found that notable gains were made on particular subtests of the CMT, which otherwise would have gone unnoticed based on the public data currently being released by the state of Connecticut. Additionally, in Hartford Public Schools, residential mobility resulted in losses on test scores and English Language Learners exhibited higher gains than non-English Language Learner students on the CMT. In conclusion, as opposed to static yearly scores, gain indicators provide more accurate information in regards to the gains and losses being made on standardized tests. This better information from the use of gain indicators would lead to more effective educational reforms contributing to improvements in student and school performance.

Introduction

Accountability Based on Testing

Historically, testing began with the development of psychological intelligence tests. These tests were used as standard indicators of intelligence across various disciplines and were developed based on the skillsets teachers implemented in the classroom (Siegler, DeLoache, & Eisenberg, 2010). In the United States, intelligence tests were used within different organizations for job placements, as well as in the school systems. The movement of testing into education was initially a strategy to track students into the appropriate grade levels. Additionally, tests were implemented to provide students with the opportunity to hold themselves accountable for their own success in order to graduate (Siegler et al., 2010).

The shift from using testing as a tracking method to using it as a means of accountability in schools, originated from the publication of the report *A Nation at Risk*. Schools were supposed to be held accountable based on their students' performance on new academic goals, curricula, and tests. However, state policy makers failed to carry out rewards and sanctions based on the quality of performance and the impact of the report did not fully influence the stakes of standardized tests (Walberg, 2003).

The implementation of the No Child Left Behind Act (NCLB) in 2002 created accountability systems in all schools based on high-stakes state mandated standardized tests (Orlich, 2004). Under NCLB, one hundred percent of students are expected to reach proficiency on state-level standardized tests by 2014. Public schools are required to achieve adequate yearly progress (AYP) every school year towards the goal of attaining one hundred percent proficiency (Orlich, 2004). High-stakes testing, in regards to NCLB, refers to the fact that if schools do not meet their AYP requirements on state-level standardized tests, the consequences can be as

serious as school closures (Harris, 2011), a consequence *A Nation at Risk* failed to instill in the minds of policy makers. Additionally, the need to reach proficiency and meet AYP has led certain states, such as Connecticut, to release public data showing solely static yearly scores.

Research Question and Thesis

My research question seeks to answer, what does independent student-level data show about average grade-level gains on the CMT, compared to the public data that is released by the Connecticut State Department of Education (as shown in Appendix A)? I argue that information on grade-level gains, based on student-level data, provides valuable insight in regards to the gains and losses schools experience on standardized tests from one year to the next. Gains represent positive numbers and losses represent negative numbers based on my gain indicator calculations. I also argue that residential mobility, the students who moved between schools, results in greater losses on CMT scores compared to students who remain in the same school. Finally, I argue English Language Learner (ELL) students experience higher gains than non-ELL students on their CMT scores.

Implications of Research

As stated above, the contents of NCLB have resulted in states reporting scores at a single point in time, instead of observing gains made over time. Low achieving schools are being held most accountable for improving their test scores in order to meet the requirements of NCLB. However, these low achieving schools are left at a disadvantage due to the initial achievement level of their students when entering the school system, compared to high performing schools (Harris, 2011). For example, a number of students in low performing schools do not have the opportunity to attend pre-school, leaving them at a disadvantage from the first day they enter

kindergarten. The initial inequalities created by these circumstances follow students throughout their entire academic career (Harris, 2011).

The use of static scores to measure the performance of schools and students leads to numerous repercussions for low performing schools. First, poor performing students can be excluded from testing by schools in order to raise scores. Second, the sanctions created by high-stakes testing can lead to the pushing out of quality teachers because they seek opportunities at higher performing schools (Darling-Hammond, 2003). Finally, high-scoring schools are allotted the ability to cease attempts for reform and improvements if they are consistently meeting proficiency requirements. Low-scoring schools are not provided with the opportunity to resist change and experience turnover in their policies and systems based on the results of static scores. Due to the fact that static scores are not accurate measures of school performance, turnovers in policies are then being made when they may not be necessary or could be allocated elsewhere to make improvements (Harris, 2011).

Literature Review

Value-Added Assessment

The consequences created by holding schools accountable based on static scores indicates a need for change in the assessment method of various states in regards to their standardized tests. A value-added assessment for standardized test scores would be the ideal strategy for measuring the performance of students and schools. Value-added assessment is based on individual student growth measures; however, this assessment method accounts for external factors outside the control of schools, such as student demographic factors and class size (Harris, 2011). Depending on the extent to which the external factors disadvantage the achievement of the students within the school, a value-added system would account for these disadvantages

when assessing the performance of schools. School experiencing disadvantages would be compensated for these uncontrollable factors, in comparison to the schools that do not experience these disadvantages (Harris, 2011).

Growth Models

Accounting for inequalities when measuring the performance of students and schools is the ideal strategy; however, when information on external factors cannot be provided, the best assessment option entails determining the amount of progress students make from one year to the next. Currently, the static test score data is valuable for AYP reports. As mentioned previously, this data does not provide information on individual student growth rates and it also fails to account for initial differences in achievement between cohorts of students. Performance in education should be a measure of the impact schools and teachers have on student outcomes and it cannot be measured accurately if schools and teachers are being compared based on static data (Harris, 2011).

Growth measures can be calculated in three different ways: cohort-to-cohort, growth-to-proficiency, and individual student growth. Cohort-to-cohort measures have two significant consequences when used to measure school performance. First, solely calculating growth based on the percentage of students performing at or above the proficiency level is a poor indicator of performance and neglects to reveal important information about student gains. Additionally, basing growth calculations off of the students performing at the proficiency level neglects to account for initial differences in achievement across schools. Second, cohort-to-cohort growth is limited to measuring only the growth of different groups of students in the same grade over time. This limitation means cohort-to-cohort measures fail to account for the growth of individual continuous students within each cohort (Harris, 2011).

Growth-to-proficiency measures, like student growth measures, account for the initial differences in achievement for each school and calculate student growth. However, the first problem with growth-to-proficiency measures is that proficiency standards were developed with no systematic basis. Additionally, disadvantaged schools with low initial achievement levels are expected to make learning gains at a faster rate than other schools in order to reach proficiency at the same time (Harris, 2011).

Student growth measures are the closest assessment method to value-added measures in regards to providing accurate test score information to contribute to school performance standards. These measures account for the growth made on test scores from one year to the next for each individual student (Harris, 2011). For example, a student's score on the CMT from third grade would be subtracted from their score in fourth grade in order to determine their test score improvement from third to fourth grade.

Student growth measures could potentially eliminate incentives to exclude certain students if teachers were given the opportunity to determine how much each student was expected to learn throughout the year. Additionally, teachers will not be punished for the starting inequalities of their students. They will be evaluated based on the gains made by students from their starting point, instead of in comparison to other cohorts of students. Low-scoring schools would be able to stop the endless cycle of implementing new programs and curriculums and certain high-scoring schools may see they are in need of improvements and reforms to their systems. Finally, Harris (2011) claims student-growth measures could eliminate large amounts of frustration felt by teachers and schools due to the inaccuracy of "snapshot" measures.

Gain Indicators

School gains can be calculated based on gain indicators, which observe the average growth in achievement from one time period to another for the same cohort of students. Gain indicators do not control for the external factors, such as student and family demographic information; however, gain indicators are the most advantageous assessment measure when information on these external factors cannot be provided.

Gain indicators calculate individual student growth from one year to the next, and then average this growth for each year to determine overall gains. Additionally, gain indicators can only be used when the test administered to the students is scored based on a common scale for every grade (Meyer, 1997). The CMT is a standardized test with a common scale, allowing for gain calculations for individual students from one school year to the next.

English Language Learners and Residential Mobility

Research indicates that ELL students perform markedly lower on standardized tests than native English speakers. The achievement gap is as large as 20-40% points on standardized tests between these two groups. No Child Left Behind (NCLB) requires a 95% participation rate on all state assessments, indicating very few English Language Learner (ELL) students are exempt from standardizing testing. As shown in Appendix A, only 16 third grade students in Hartford in 2013 were exempt from the CMT mathematics subtest due to ELL status. Additionally, every state must develop objects to measure the achievement of ELL students to determine whether or not they are making adequate yearly progress in regards to their acquisition of the English language, in addition to meeting the same AYP requirements of native English speakers (Menken, 2006).

State mandated standardized tests were created in order to measure the achievement of native English speakers and not ELL students. However, ELL students are expected to achieve the same level of proficiency as native English speakers, which results in standardized tests becoming measures of ELL students' language proficiency, instead of their academic abilities (Menken, 2006). The progress ELL students are making over time is substantially more important to their academic development compared to their language proficiency. ELL students are not performing well on high-stakes standardized tests resulting from the implementation of NCLB and a new evaluation system needs to be created for these students. They need to be assessed based on their improvements on standardized test scores, opposed to their yearly percentage scores.

Additionally, research has shown residential mobility results in lower performance on test scores compared to students who do not move. Only five percent of the discrepancy between students who remain at the same school and students who move between schools is due to the stress associated with moving (Pribesh, 1999). Poor performance of students with high levels of residential mobility is typically due to factors present prior to the moves. Moving has also been shown to result in a loss of social ties. Moving from one school to another leads to a lack of identification with a set social group. Without social acceptance, students typically fall into place with less motivated students who perform at lower levels than students with social security (South, 2005).

Methodology

For my study, I used an exclusively quantitative approach. The dataset I used was a pre-existing dataset, which contained independent student-level data on standardized test scores for all Hartford Public School students. Hartford Public Schools provided this dataset to Trinity

College and I received permission to gain access to this data under the agreement that I would have no contact with any of the subjects, or any ability to track their identities. A professor on Trinity's campus ensured the security of the student identities by replacing all student identification codes with new codes for the purpose of use by Trinity College student researchers. Additionally, home addresses and zip codes were deleted for every student from my personal dataset.

The variables provided in my dataset included: students' masked Hartford public school ID, City, Facility Code, Grade Code, Resident Town, Gender, Race, Dominant Language Code, ELL, Special Education, CMT Mathematics-Level Score, CMT Mathematics-Scale Score, CMT Mathematics-Vertical Scale Score, CMT Reading-Level Score, CMT Reading-Scale Score, CMT Reading-Vertical Scale Score, CMT Writing-Level, and CMT Writing-Scale Score. The data for each of these variables was provided for the 2010-11, 2011-12, and 2012-13 academic school years.

The student-level data was provided on three separate excel spreadsheets, one for each school year. In order to conduct my analysis, I merged all three spreadsheets into one SPSS file. The merged file originally contained information on 30,659 students. I decided to work solely with scale scores, opposed to level scores. Scale scores account for the difficulty level of questions on standardized tests, while level scores do not. For example, when two students get the same number of questions correct on one of the subtests, they would receive the same raw score. However, if one student got all questions labeled difficult correct and the other only got questions labeled easy correct, the student who got the difficult questions correct would receive a higher scale score and the student who got the easy questions correct would receive a lower scale score (Harris, 2011).

Additionally, I decided to work with students who had continuous data in regards to their

CMT scores; meaning these students had no missing information for any of the three CMT

subtests across all three school years. I created a new variable to identify the students with

missing test score information and put a filter on my dataset to eliminate these students from

further analysis. As a result of this filtering, I ended up working with a sample of 3,585 students.

Once I filtered my data, I calculated the gains for each individual student on the three

CMT subtests for the 2010-11 school year to the 2011-12 school year, for the 2011-12 school

year to the 2012-13 school year, and I calculated an overall gain for the 2010-11 school year to

the 2012-13 school year. I will use the CMT mathematics scale scores to provide an example of

how I completed these calculations (as shown in the equation below).

$$\text{MathGain\_1012} = (\text{CMTMathScaleScore\_1112}) - (\text{CMTMathScaleScore\_1011})$$

I first created a new variable in my dataset that would represent the gain on the CMT

mathematics scale score from 2010-11 to 2011-12. In the equation, this would be the variable

labeled MathGain_1012. In order to calculate this gain, I subtracted the 2010-11 mathematics

scale score from the 2011-12 mathematics scale score, as shown above in the equation. I created

two new gain variables for the mathematics gain from 2011-12 to 2012-13 and for the

mathematics gain from 2010-11 to 2012-13. If I were to show these two equations, the new

variables would be labeled MathGain_1113 and MathGain_1013, respectively.

After calculating nine new variables in total, three for each of the subtests, I determined

the average gains for each grade level based on my new variables. I first looked at the cohort of

third grade students from the 2010-11 school year by selecting only the students in my dataset

with the grade code for third grade. After applying this filter, I did a comparison of means to

determine the average gains this cohort of students made from third to fourth grade, fourth to fifth grade, and then the overall gains from third to fifth grade for all three subtests. I followed this same process for the cohort of fourth grade students from the 2010-11 school year, the fifth grade students from the 2010-11 school year, and the sixth grade students from the 2010-11 school year.

I then calculated the same gains for these four cohorts of students based on their residential mobility, meaning whether or not they moved between schools, in addition to whether or not they were labeled as English Language Learners. In order to calculate gains based on residential mobility, I created a new variable to identify the students with the same facility code across all three years and the students with a change in facility code at any point in time over the three school years. I did a comparison of means for each cohort to determine the average gains made for students who remained in the same school in comparison to the average gains made for the students who moved between schools. In order to calculate gains based on ELL status, I also did a comparison of means for each cohort to determine the average gains made by students who were ELL in comparison to non-ELL students for all three subtests.

Findings

CMT Gains Based on Grade-level

My research found that the 2010-11 third grade cohort ($N = 923$) exhibited substantial gains on the mathematics and reading subtests and losses on the writing subtest from third to fourth grade. This cohort exhibited losses on the mathematics and reading subtests and gains on the writing subtests from fourth to fifth grade. Overall, there were gains exhibited for all three subtests from third grade to fifth grade (as shown in Table 1.1).

The 2010-11 fourth grade cohort ($N = 793$) exhibited slight gains on the mathematics subtest, losses on the reading subtest, and notable gains on the writing subtest from third to fourth grade. This cohort exhibited losses on the mathematics and writing subtests and notable gains on the reading subtest from fifth to sixth grade. Overall, there was a slight loss for the mathematics subtest and notable gains for the reading and writing subtests from fourth grade to sixth grade (as shown in Table 1.2).

The 2010-11 fifth grade cohort ($N = 799$) exhibited losses on the mathematics and writing subtests and the highest gain of all four cohorts on the reading subtest from fifth to sixth grade. The next year, this cohort exhibited losses on the reading subtest, in addition to losses again on the mathematics and writing subtests, from sixth to seventh grade. Overall, losses were exhibited on the mathematics and writing subtests and the highest overall gain was exhibited on the reading subtest from fifth to seventh grade (as shown in Table 1.3).

The 2010-11 sixth grade cohort ($N = 1053$) exhibited losses on all three subtests from sixth to seventh grade. This cohort exhibited losses on the mathematics subtest and gains on the reading and writing subtests from seventh to eight grade. Overall, losses were exhibited on all three subtests from sixth to eighth grade (as shown in Table 1.4).

CMT Gains Based on Residential Mobility

Students who moved between schools ($N = 235$) sometime across the three school years in the 2010-11 third grade cohort exhibited lower gains on all three subtests from third to fourth grade and fourth to fifth grade compared to students who remained in the same school. Thus, indicating overall lower gains on all three subtests for the students who moved between schools from third to fifth grade (as shown in Table 2.1).

Students who moved between schools ($N = 320$) in the 2010-11 fourth grade cohort exhibited lower gains on the reading and writing subtests and higher gains on the mathematics subtest from fourth to fifth grade compared to students who remained in the same school. Students who moved exhibited lower gains in the mathematics and reading subtests and higher gains on the writing subtest from fifth to sixth grade compared to students who remained in the same school. Overall, students who moved between schools exhibited lower gains on the mathematics and writing subtests and higher gains on the reading subtest from fourth to sixth grade (as shown in Table 2.2).

Students who moved between schools ($N = 330$) in the 2010-11 fifth grade cohort exhibited overall lower gains on all three subtests from fifth to sixth grade compared to students who remained in the same school. All students exhibited losses from sixth to seventh grade; however, students who moved exhibited lower gains for the math and reading subtests and higher gains for the writing subtest compared to the students who remained in the same school. Overall, students who moved exhibited lower gains on all three subtests from fifth to seventh grade compared to students who remained in the same school (as shown in Table 2.3).

All students exhibited losses on all three subtests from sixth to seventh grade for the 2010-11 sixth grade cohort; however, students who moved between schools ($N = 207$) exhibited lower gains on the mathematics and reading subtests and higher gains in the writing subtest than students who remained in the same school. Students who moved exhibited lower gains on the writing subtest and higher gains on the mathematics and reading subtests from seventh to eighth grade. Overall, students who moved exhibited lower gains on all three subtests from sixth to eighth grade compared to students who remained in the same school (as shown in Table 2.4).

CMT Gains Based on English Language Learner (ELL) Status

ELL students ($N = 136$) in the 2010-11 third grade cohort exhibited higher gains than non-ELL students on the reading subtest and lower gains on the mathematics and writing subtests from third to fourth grade. ELL students exhibited higher gains than non-ELL students on the reading and writing subtests and lower gains on the mathematics subtest from fourth to fifth grade. ELL students exhibited overall higher gains than non-ELL students on the reading and writing subtests and lower gains on the mathematics subtest from third to fifth grade (as shown in Table 3.1).

ELL students ($N = 122$) in the 2010-11 fourth grade cohort exhibited higher gains than non-ELL students on the reading and writing subtests and lower gains on the mathematics subtest from fourth to fifth grade. ELL students exhibited higher gains than non-ELL students on the mathematics subtest and lower gains on the reading and writing subtests from fifth to sixth grade. Overall, ELL students made higher gains than non-ELL students on the reading and writing subtests and lower gains on the mathematics subtest from fourth to sixth grade (as shown in Table 3.2).

ELL students ($N = 129$) in the 2010-11 fifth grade cohort exhibited higher gains than non-ELL students on the mathematics and reading subtests and lower gains on the writing subtest from fifth to sixth grade. All three subtests showed losses from sixth to seventh grade for both ELL and non-ELL students; however, ELL students exhibited higher gains on all three subtests. Overall, ELL students exhibited higher gains than non-ELL students on all three subtests from fifth to seventh grade (as shown in Table 3.3).

Both ELL ($N = 141$) and non-ELL students in the 2010-11 sixth grade cohort from sixth to seventh grade exhibited losses on all three subtests; however, ELL students exhibited higher

gains on the reading and writing subtests and lower gains on the mathematics subtest. ELL

students exhibited higher gains than non-ELL students on the mathematics and reading subtests

and lower gains on the writing subtest from seventh to eighth grade. Overall, from sixth to eighth

grade both ELL and non-ELL students exhibited losses on all three subtests; however, ELL

students made higher gains on all three subtests (as shown in Table 3.4).

Discussion and Conclusions

According to the findings, student-level data can be used to calculate the growth of

individual continuous students within the Hartford Public School system. These growth scores

can then be averaged together in order to determine the gains and losses made by the same group

of students as they progress from one grade-level to the next. Notable gains, such as the 25 point

gain on the reading subtest for the 2010-11 fifth grade cohort from fifth to sixth grade, could be

used to help identify changes that occurred in school policies for reading or changes to items on

the reading subtest between those two years.

Overall, the implementation of policies, such as the Common Core State Standards,

should be leading to gains on standardized tests in every district. The Common Core State

Standards provide a framework for what students are expected to learn. This framework is meant

to help guide teachers and schools towards developing curricula relevant to the real world in

order to promote success among students after high school. Connecticut adopted these standards

in 2010 and is predicted to have them fully implemented by the end of the 2013-14 school year

(National Governors Association Center for Best Practices, Council of Chief State School

Officers, 2010). The four cohorts of students observed in this research study did show gains on

certain subtests, however, there were not overall notable gains in every subtest from 2010-2013.

This information could indicate the Common Core Standards had a lack of impact on

performance on standardized test scores or that school curricula geared towards success in the real world in not applicable to performance on standardized testing.

Additionally, observing notable losses could be beneficial to identifying mistaken policy changes or other root causes of declines in scores. For example, CMT testing in the 2012-13 school year occurred at the beginning of March, less than a month after the devastating nor'easter Nemo, which sent Connecticut into a state of emergency and led to the closing of Hartford Public Schools for almost an entire week. The data shows notable losses on many of the CMT subtests from the 2011-12 school year to the 2012-13 school year. It is possible these losses could be due to declines in student performance; however, it is also critical to note that the chaos created by Nemo could have had a substantial impact on test scores that year, resulting in the large losses from the 2011-12 school year to 2012-13 school year.

Compared to data provided by the Connecticut State Department of Education (as shown in Appendix A), data provided on the gains and losses of various cohorts of students are essential to identifying when improvements are occurring within schools and when changes need to be made. Without gain indicators from one school year to the next, schools are continuing to make decisions based on whether or not they have made AYP. Hartford Public Schools continue to perform below proficiency level, leading to changes in policy and education reforms that are being made without knowledge of where and when gains are being made.

Residential mobility, meaning the students who moved between schools, had a negative impact on test score gains, compared to students who remained in the same school. The losses experienced by students who moved between schools could be a result of various factors. First, the most substantial losses occurred during the shift from elementary school to middle school, from sixth to seventh grade for most Hartford Public Schools. For the 2010-11 sixth grade

cohort, the overall losses from sixth to eighth grade for the students who moved between schools were the lowest out of the four cohorts of students. Second, the curriculums between schools may vary, and transitions from one school to the next may result in new students falling behind and performing lower on standardized tests. Finally, students who change schools have been shown to have weaker social ties than friends who remain in the same school, thus leading them to develop friendships with lower performing students (South, 2005).

Overall, the findings indicated that English Language Learner (ELL) students exhibited higher gains than non-ELL students. The overall gains made on the mathematics subtest for the 2010-11 third and fourth grade cohorts were the only indicators of ELL students exhibiting lower gains than non-ELL students. The lower gains on the mathematics subtest for ELL students was likely due to the fact that mathematics does not require proficiency in the English language. Without the language barrier, the discrepancy between the two groups is eliminated.

As stated in the literature, ELL students perform at a lower level on standardized test scores than non-ELL students (Menken, 2006). However, the substantial gains ELL students have made on standardized tests indicate that static scores may not be sufficient measures of their performance. Static scores represent ELL students' ability to take a test meant for English speakers, in comparison to students who speak English as their primary language. Calculating the gains made by ELL students provides indicators of the progress they are making in the English language.

Future Research

Based on my findings, I suggest future research take into consideration the grade-level gains made by interdistrict schools compared to district schools. The comparison of interdistrict and district schools has the potential to identify whether charter schools are making a positive or

negative impact on test scores in Hartford. Additionally, future research should consider further

connecting the notable gains and losses found in my research to statewide or school-wide policy

changes. The linkage of my findings to policy changes has the potential to provide schools with

information on the policy changes that are effective and the policy changes that are leading to

losses on test scores. Finally, future research should consider conducting gain calculations for all

school districts in Connecticut. The comparison of all districts would allow for the control of

major events, such as weather conditions and economic recessions, which could be causing

overall gains and losses on tests scores for the entire state. Controlling for these external factors

affecting school performance would allow for each school district to be assessed more accurately

based on their internal performance.

References

Connecticut State Department of Education. *Connecticut Mastery Test Public Summary Performance Reports*. Retrieved from http://www.ctreports.com/

Darling-Hammond, L. (2003). Standards and Assessments: Where We Are and What We Need. *Teachers College Record.*

Harris, D. N. (2011). *Value-Added Measures in Education*. Cambridge, MA: Harvard Education Press.

Menken, K. (2006). Teaching to the Test: How No Child Left Behind Impacts Language Policy, Curriculum, and Instruction for English Language Learners. *Bilingual Research Journal, 30 (2)*. 521-546.

Meyer, R. H. (1997). Value-Added Indicators of School Performance: A Primer. *Economics of Education Review, 16 (3)*. 283-301.

National Governors Association Center for Best Practices, Council of Chief State School Officers (2010). *Common Core State Standards.* Washington D.C.: National Governors Association Center for Best Practices, Council of Chief State School Officers.

Orlich, D. C. (2004). No Child Left Behind: An Illogical Accountability Model. *The Clearing House, 78 (1).* 6-11.

Pribesh, S., & Downey, D.B. (1999). Why Are Residential and School Moves Associated with Poor School Performance? *Demography, 36 (4)*. 521-534.

Siegler, R., DeLoache, J., & Eisenberg, N. (2010). *How Children Develop.* United States of America: Worth Publishers.

South, S. (2005). Adolescent Residential Mobility and Premature Life-Course Transitions: The Role of Peer Networks. *Sociological studies of children and youth, 11*. 23-52.

Walberg, H. J. (2003). Accountability Unplugged: Time to Actually Try Standards-Based Reform. *Education Next, 3(2)*.

Table 1.1: 2010-11 3rd Grade Cohort Average CMT Gains

Year	Mathematics	Reading	Writing
2010-2012 (3rd-4th Grade)	7.86	10.50	-3.98
2011-2013 (4th-5th Grade)	-4.16	-8.96	6.26
2010-2013 (3rd-5th Grade)	3.70	1.54	2.28

Note: *N* = 923

Table 1.2: 2010-11 4th Grade Cohort Average CMT Gains

Year	Mathematics	Reading	Writing
2010-2012 (4th-5th Grade)	.88	-3.06	6.11
2011-2013 (5th-6th Grade)	-1.57	12.67	-.58
2010-2013 (4th-6th Grade)	-.69	9.62	5.54

Note: *N* = 793

Table 1.3: 2010-11 5th Grade Cohort Average CMT Gains

Year	Mathematics	Reading	Writing
2010-2012 (5th-6th Grade)	-3.24	25.13	-2.71
2011-2013 (6th-7th Grade)	-5.55	-8.56	-3.57
2010-2013 (5th-7th Grade)	-8.78	16.57	-6.28

Note: *N* = 799

Table 1.4: 2010-11 6th Grade Cohort Average CMT Gains

Year	Mathematics	Reading	Writing
2010-2012 (6th-7th Grade)	-2.17	-8.17	-7.10
2011-2013 (7th-8th Grade)	-6.23	4.50	3.02
2010-2013 (6th-8th Grade)	-8.40	-3.68	-4.08

Note: *N* = 1053

Table 2.1: 2010-11 3rd Grade Cohort: Average Gains for Residential Mobility

Year	Mathematics	Reading	Writing
2010-2012 (3rd-4th Grade)			
Moved	4.82	7.73	-3.93
Remained	8.89	11.45	-4.00
2011-2013 (4th-5th Grade)			
Moved	-5.41	-8.26	5.75
Remained	-3.73	-9.20	6.43
2010-2013 (3rd-5th Grade)			
Moved	-.59	-.54	1.82
Remained	5.16	2.25	2.43

Note: Moved between schools $N = 235$; Remained in same school $N = 688$

Table 2.2: 2010-11 4th Grade Cohort: Average Gains for Residential Mobility

Year	Mathematics	Reading	Writing
2010-2012 (4th-5th Grade)			
Moved	1.19	-2.61	3.68
Remained	.67	21.15	7.76
2011-2013 (5th-6th Grade)			
Moved	-3.49	12.32	1.42
Remained	-.26	12.90	-1.93
2010-2013 (4th-6th Grade)			
Moved	-2.31	9.72	5.09
Remained	-.41	9.57	5.83

Note: Moved between schools $N = 320$; Remained in same school $N = 473$

Table 2.3: 2010-11 5th Grade Cohort: Average Gains for Residential Mobility

Year	Mathematics	Reading	Writing
2010-2012 (5th-6th Grade)			
Moved	-6.20	23.00	-6.26
Remained	-1.16	26.62	-.20
2011-2013 (6th-7th Grade)			
Moved	-6.30	-9.39	-1.39
Remained	-5.03	-7.97	-5.10
2010-2013 (5th-7th Grade)			
Moved	-12.50	13.62	-7.65
Remained	-6.19	18.65	-5.31

Note: Moved between schools $N = 330$; Remained in same school $N = 469$

Table 2.4: 2010-11 6th Grade Cohort: Average Gains for Residential Mobility

Year	Mathematics	Reading	Writing
2010-2012 (6th-7th Grade)			
Moved	-5.94	-10.75	-6.02
Remained	-1.24	-7.54	-7.37
2011-2013 (7th-8th Grade)			
Moved	-3.59	6.58	-.16
Remained	-6.88	3.98	3.80
2010-2013 (6th-8th Grade)			
Moved	-9.53	-4.16	-6.18
Remained	-8.12	-3.56	-3.57

Note: Moved between schools $N = 207$; Remained in same school $N = 846$

Table 3.1: 2010-11 3rd Grade Cohort: Average Gains for ELL vs. Non-ELL Students

Year	Mathematics	Reading	Writing
2010-2012 (3rd-4th Grade)			
ELL	7.64	13.90	-3.74
Non-ELL	7.89	9.91	-4.03
2011-2013 (4th-5th Grade)			
ELL	-6.63	-3.72	10.9
Non-ELL	-3.73	-9.87	5.45
2010-2013 (3rd-5th Grade)			
ELL	1.01	10.17	7.20
Non-ELL	4.16	.05	1.42

Note: ELL N = 136; Non-ELL N = 787

Table 3.2: 2010-11 4th Grade Cohort: Average Gains for ELL vs. Non-ELL Students

Year	Mathematics	Reading	Writing
2010-2012 (4th-5th Grade)			
ELL	-7.43	4.87	9.41
Non-ELL	2.39	-4.49	5.51
2011-2013 (5th-6th Grade)			
ELL	3.21	12.48	-.93
Non-ELL	-2.43	12.70	-.51
2010-2013 (4th-6th Grade)			
ELL	-4.21	17.35	8.48
Non-ELL	-.05	8.21	5.00

Note: ELL N = 122; Non-ELL N = 671

Table 3.3: 2010-11 5th Grade Cohort: Average Gains for ELL vs. Non-ELL Students

Year	Mathematics	Reading	Writing
2010-2012 (5th-6th Grade)			
ELL	-2.05	29.00	-3.16
Non-ELL	-3.47	24.38	-2.62
2011-2013 (6th-7th Grade)			
ELL	-4.52	-6.31	-1.20
Non-ELL	-5.75	-8.99	-4.03
2010-2013 (5th-7th Grade)			
ELL	-6.57	22.69	-4.36
Non-ELL	-9.22	15.39	-6.64

Note: ELL $N = 129$; Non-ELL $N = 670$

Table 3.4: 2010-11 6th Grade Cohort: Average Gains for ELL vs. Non-ELL Students

Year	Mathematics	Reading	Writing
2010-2012 (6th-7th Grade)			
ELL	-2.21	-6.16	-1.11
Non-ELL	-2.16	-8.49	-8.03
2011-2013 (7th-8th Grade)			
ELL	-2.81	5.18	.08
Non-ELL	-6.76	4.39	3.47
2010-2013 (6th-8th Grade)			
ELL	-5.02	-.98	-1.04
Non-ELL	-8.92	-4.10	-4.55

Note: ELL $N = 141$; Non-ELL $N = 912$

Appendix A

2013 Connecticut Mastery Test Summary Report Grade 3

Mathematics

Standard CMT Score Summary		**Standard CMT Results by Level**		**Assessment Participation**	
Average Math scale score	221.7	*Advanced*	7.9%	*Standard CMT*	1293/87.5%
Average raw score	76.8			*Skills Checklist*	39/2.6%
		Goal	22.3%	*Modified Assessment*	124/8.4%
Average # of content strands Mastered	12.5			*ELL Exempt*	16/1.1%
		Proficient	29.3%	*No Valid Score*	3/0.2%
Percent at/above goal level	30.2				
		Basic	18.6%	Total Participation	1475/99.8%
				Absent	3/0.2%
Percent at/above proficient level	59.5	*Below Basic*	22.0%		
				Total Enrollment	1478/100.0%

Source: Connecticut State Department of Education, CMT Public Summary Performance Reports