

January 01, 2011

Effects of standards-based report cards on student learning

Theresa A. Craig
Northeastern University

Recommended Citation

Craig, Theresa A., "Effects of standards-based report cards on student learning" (2011). *Education Doctoral Theses*. Paper 16.
<http://hdl.handle.net/2047/d20002355>

This work is available open access, hosted by Northeastern University.

EFFECTS OF STANDARDS-BASED REPORT CARDS ON STUDENT LEARNING

A thesis presented
by

Theresa A. Craig

to
The School of Education

In partial fulfillment of the requirements for the degree of
Doctor of Education

in the field of

Educational Leadership

College of Professional Studies
Northeastern University
Boston, Massachusetts
November 2011

EFFECTS OF STANDARDS-BASED REPORT CARDS ON STUDENT LEARNING

by

Theresa A. Craig

ABSTRACT OF DISSERTATION

The purpose of this causal-comparative study is to examine the effect of standards-based report cards on the growth and performance of Grade 4 students on the Massachusetts measure of accountability. This study examines the effect of the report card format on the grade level student growth percentiles (SGP) and composite performance index (CPI) on the 2010 Grade 4 Massachusetts Comprehensive Assessment System (MCAS) Mathematics Test. The sample population for this study is 103 elementary schools educating fourth grade students in southeastern Massachusetts. The study uses historical performance data from the Massachusetts Department of Elementary and Secondary Education, a brief questionnaire collected from elementary school leaders, and an ordinal product rating scale. The study conducts a document analysis of standards-based report cards collected from schools by applying a product rating scale created for this study. A multivariate analysis of covariance is used to build the ex post facto relationship between schools using standards-based report cards and results for the Grade 4 growth percentile and performance index on the MCAS for Mathematics. Key terms:

- *Standards-based report cards*
- *Composite Performance Index (CPI)*
- *Student Growth Percentile (SGP)*
- *MCAS – Massachusetts Comprehensive Assessment System*

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Education in the field of Educational Leadership in the College of Professional Studies
of Northeastern University, November, 2011

Acknowledgements

I wish to acknowledge the guidance and assistance of my advisor Dr. David Szabla. Through his persistent emphasis on deeper understanding, this research project has evolved into a comprehensive study. I want to thank Dr. Claire Jackson for her support during the initial stages of my project development as well as for her professional and personal guidance throughout the doctoral program. I also owe a debt of gratitude to the members of my cohort whose insights and contributions to class discussions greatly contributed to my learning at Northeastern University.

This project would not have been possible without the cooperation of over one hundred elementary principals throughout southeastern Massachusetts. I want to thank each of them for completing my questionnaire and sending copies of their report cards. I was deeply touched by the many personal notes that were attached to the questionnaires wishing me good luck in my pursuits.

I cannot thank my family enough for their support during my many years of post-graduate studies. To my parents Jack and Linda Harrington: Thank you for never wavering in your belief that I can accomplish anything I set my mind to. To my children, Jared, Ashley and Lacey and daughter-in-law, Katie: Thank you for understanding when I had to work intermittently on my laptop during conversations and for recognizing that while I was reluctant to miss family functions it was sometimes a necessity. To my loving husband Glenn: I could not have done any of this without your support, love, understanding and willingness to take on almost every other responsibility in our lives so that I could complete my studies. Thank you for the sacrifices that you have made to help me accomplish this goal.

TABLE OF CONTENTS

Abstract	2
Acknowledgments	3
CHAPTER 1: INTRODUCTION	7
Purpose of the Study	7
Statement of the Problem and its Significance	9
Research Questions	12
Definitions	13
Conceptual Framework	15
Theoretical Framework	21
Research Design	28
Limitations of the Study	30
Delimitations of the Study	32
CHAPTER 2: LITERATURE REVIEW	33
Historical Perspective	33
Grading Practices	33
Standards-Based Teaching	35
Accountability Movement	37
Impact of Grading Practices on Self-Efficacy and Motivation	39
Current Trends in Grading Practices	40
A Continuum of Learning	40
Failing Grades as Punishment	43
Separation of Learner Behaviors	45
Standards and Indicators	47
Resistance: Transforming the Culture	51
Literature Review Conclusion	52
CHAPTER 3: RESEARCH DESIGN	53
Variables	54
Independent variable	55
Dependent variables	55
<i>Student Growth Percentiles</i>	55
<i>Composite Performance Index</i>	56
Covariates	56
Threats to Internal Validity	56
Research Questions	57
Methodology	59
Participants and Population Sample	60
Measures	60
<i>Rating Scale</i>	60
<i>Grouping Rubric</i>	62
Procedures	65
Data Collection	65

<i>Questionnaires</i>	66
<i>Report Card Coding</i>	66
<i>Growth and Performance</i>	67
<i>Covariates</i>	67
<i>Data Fields</i>	68
Data Analysis: Multivariate Analysis of Covariance	68
Protection of Human Subjects	69
Timeline	70
CHAPTER 4: REPORT OF RESEARCH FINDINGS	71
Independent Variable	71
Questionnaires	71
Report card coding	71
<i>Continuum of progress</i>	73
<i>Removal of failing grades</i>	77
<i>Separating learner and social behaviors</i>	78
<i>Standards and indicators</i>	78
Independent Variable Statistics	80
Population and Sample data	81
Data analysis	82
Covariate statistics	83
Multivariate Analysis of Covariance	85
<i>Controlling for Assumptions of MANCOVA</i>	86
<i>MANCOVA Test of Significance</i>	87
<i>Mean and Standard Deviation</i>	89
Rejection or Acceptance of Hypotheses	90
Hypothesis 1	90
Hypothesis 2	91
Hypothesis 3	92
Hypothesis 4	93
Summary of MANCOVA results	94
Additional Analysis of Results	94
Tests of linear regression	94
Additional Analysis of Hypothesis 2	98
Summary of Research Findings	98
Continuum of Progress	98
Failing Grades	100
Social and Learner Behaviors	101
Standards and Indicators	101
CHAPTER 5: DISCUSSION OF RESEARCH FINDINGS; IMPLICATIONS FOR EDUCATIONAL PRACTICE	101
Research Questions	102
Interpretation of Findings	103
Shared Understanding	104
Performance Levels	105

Multiple indicators of proficiency	106
Failing Grades	108
Effect of Student Feedback on Performance	109
Implications	110
Implications for Practice	110
Implications for Future Research	113
Conclusion	115
References	118
Bibliography	124
Appendix A: Questionnaire and Survey	128
Appendix B: District and School Assistance Center Regions	131
Appendix C: Standards-Based Report Card Product Rating Scale	132
Appendix D: Standards-Based Report Card Grouping Rubric	133
Appendix E: CPI Performance Rating and Range	134
Appendix F: Descriptive Statistics for Covariates Separated by Independent Variable Grouping	135
Appendix G: Tests of Between-Subjects Effects	136
Appendix H: MANCOVA Results for the Independent Variable	137
Appendix I: Univariate Test for the SGP Dependent Variable	138
Appendix J: Univariate Test for the CPI Dependent Variable	139

Chapter 1: Introduction

Purpose of the Study

Many leaders of elementary schools are adopting standards-based report cards in order to align local forms of reporting achievement with the format that is used to report results on high-stakes testing. Standards-based report cards provide information on student mastery of content and skills within the strands, standards, and indicators of the established curriculum and inform students of their progress along a continuum of proficiency. Research clearly identifies the positive impact of the best practices that are engrained in standards-based grading such as: administering formative and summative assessments; providing students with timely feedback; posting standards and expected outcomes for lessons and activities; and continually providing opportunities to master concepts and skills (Aidman, Gates & Deterra Sims, 2001; Guskey, 2004a; Hargis, 1990; Marzano, 2003; Marzano & Kendall, 1996; McTighe & Bailey, 1996; Salend, 2005; Stiggins, 2005, Ward, 2004). Researchers have written extensively on the powerful impact of standards-based practices on student self-efficacy and motivation (Friedman, 2005; Hirsch, 2007; Stiggins, 1997; Stiggins, 2005; Stiggins & Chappuis, 2008, Stiggins, 2010). The impetus for introducing a standards-based report card to stakeholders is the need to improve student achievement. However, a search of the literature provides scarce supporting quantitative evidence that the utilization of standards-based report cards to report progress to students and families has a positive impact on the students' achievement on the high stakes assessments of proficiency. Families, students, and teachers are entitled to additional quantitative evidence that the effort expended to develop and implement standards-based report cards is producing positive results for student performance and growth.

Achievement depicted as an amalgamated ‘A in Math’ is a mismatch to the current assessment systems and levels of mastery that students are expected to attain in numerous discrete, well-defined areas of curriculum content. School leaders, families, and educators recognize that the traditional reporting tools that have been used for the past century to report student progress provide minimal representation of student achievement. The adoption of standards-based report cards has gained momentum in recent years. However, in many towns and cities, parents, community members, teachers, and students object to plans to alter what is viewed as a sacrosanct institution - the easily recognized, but often inaccurate portrayal of student achievement as As and Bs on traditional report cards.

The federal No Child Left Behind (NCLB) Act’s accountability designations and the resulting consequences have created a sense of urgency for school leaders to develop reporting tools that can match required accountability results. There is a clear disconnection between the state of Massachusetts’ measure of performance on the curriculum standards and the grades that represent learning on traditional report cards. It is possible for a student to receive all ‘As’ in mathematics on a traditional report card throughout the elementary grades, yet receive a performance level of ‘Needs Improvement’ or ‘Warning’ on the Massachusetts Comprehensive Assessment System (MCAS) in mathematics. In contrast, some students who have traditionally received poor grades due to a general disinterest, lack of participation, incomplete homework assignments, and low test grades have demonstrated conflicting evidence of mastery by scoring in the ‘Proficient’ range on the MCAS.

This research project examines the connection between the use of standards-based report cards and corresponding student performance and growth on the MCAS, a criterion-referenced assessment. The practical goal of this research project was to build a quantitative foundation for

the expansion of the use of standards-based report cards to the secondary level in public schools. While many elementary schools in the state of Massachusetts have adopted standards-based report cards, few secondary schools have embarked on the process. As district leaders prepare to facilitate the cultural shift from traditional grades to performance levels at the middle school, confirmation that the change will build stronger student performance over time will assist them in convincing stakeholders that this paradigm shift is worth the time, effort, and anguish.

The intellectual goal of the project is to examine the potential effect of standards-based report cards on the growth and performance of students who are most at risk for achievement. The low income subgroup of students as designated by the Department of Elementary and Secondary Education (DESE) and No Child Left Behind (NCLB) has a lower percentage of students scoring proficient on the MCAS than the aggregate. This project will help to build understanding of the best practices of reporting student progress and the elements that are essential to promote student learning.

Statement of the Problem and its Significance

The current NCLB legislation resulted in the designation of a status of *Needs Improvement, Corrective Action* or *Restructuring* for 409 elementary schools in Massachusetts in 2010 due to the underperformance of designated subgroups (MA DESE, 2011). The subgroups identified by NCLB are: low income, special education, limited English proficient, and ethnic groups of white, African American/Black, Hispanic/Latino, and Asian/Pacific Islander. Schools with a status are required to take significant corrective action according to NCLB such as: ensuring that all students have the option to transfer; ensuring that supplemental educational services are available to eligible students in the school; and providing for technical assistance to the school. In addition, the districts must take at least one of the following corrective actions:

institute a new curriculum grounded in scientifically based research and provide appropriate professional development to support its implementation; extend the length of the school year or school day; replace the school staff who are deemed relevant to the school not making adequate progress; significantly decrease management authority at the school; restructure the internal organization of the school; or appoint one or more outside experts to advise the school. The designation of almost 40% of Massachusetts schools as not making adequate yearly progress in 2010 creates a growing sense of urgency to promote more accelerated growth for students in subgroups who are most at risk.

Students who are from families experiencing poverty, students who are native speakers of another language, or those who have a disability may be especially victimized by traditional forms of grading because traditional grades are often more representative of work habits and conformity than mastery of key concepts and skills. It is an accepted practice in traditional grading to give a student the equivalent of a quiz or test grade for homework completion or preparation for class. This grading of homework completion and classroom compliance as an assessment punishes students who have fewer supports at home and may result in class failure for students even if they have acquired much of the content knowledge and skills. Standards-based grading practices separate those learner behaviors and social behaviors such as homework completion and class participation from mastery of content knowledge and skills. The traditional grading procedures that assigned students letter grades or percentages were designed to reward the highest achievers and typically permitted only a few students to feel truly successfully. These antiquated methods of assessing and reporting are now being actively revised to “permit the possibility that all students could succeed at some appropriate level” (Stiggins, 2005, “A New Mission,” para. 3).

Educational leaders have responded to the research by articulating clear standards for each grade level and assessing students' mastery toward those standards. With revised curriculum standards clearly articulated and assessment procedures to measure those standards well established, educational leaders implemented the next step and developed standards-based report cards. Standards-based report cards represent performance along a continuum of mastery levels as well as a continuum of standards and/or benchmarks. The effective feedback of standards-based grading informs the child of his/her progress, puts him/her in charge of his/her own learning and facilitates ownership of that progress. These essential elements of standards-based report cards build a foundation of self-efficacy.

Massachusetts school leaders will soon engage stakeholders in revisions of standards-based report cards in response to the March 2011 adoption of the *Massachusetts Curriculum Frameworks Incorporating the Common Core State Standards*. Educational leaders will begin navigating the new lexicon associated with the adoption of the Common Core and considering the most effective components to include on the redesigned report cards: strands, critical areas, domains, clusters, college and career readiness anchor standards, foundational skills, indicators, or individual grade-specific learning expectations (NGA Center & CCSSO, 2010). This study may provide support for leaders as they engage stakeholders in the arduous task of designing or re-designing a report card that will inform students and parents of the level of mastery of essential content-specific grade level knowledge and skills. In order to inform these efforts this study examines the relationship of standards-based report cards to student learning growth and achievement.

Research Questions

The primary research question for this research study is: What effect do standards-based report cards have on the growth percentiles and performance indices of fourth grade students on the MCAS Mathematics Test? The primary hypothesis of this study is: Fourth grade results for students in the schools with standards-based report cards will have higher percentages of growth and higher indices of performance as measured by the MCAS Mathematics Test than schools with non-standards-based report cards. The additional research questions and hypotheses presented in this research study are the following:

- Research Question 1: Do standards-based report cards that report student performance on a continuum of progress in mathematics rather than a cumulative grade average result in higher levels of growth and performance on the MCAS? Hypothesis 1: Standards-based reporting methods that provide feedback to families and students of progress along a continuum sustain motivation toward attainment of proficiency more than traditional grades, therefore producing a higher median SGP and CPI.
- Research Question 2: Do standards-based report cards that remove indications of failure such as Fs and zeros in mathematics result in higher levels of growth on the MCAS for students at risk for learning? Hypothesis 2: Grading systems that eliminate failing grades result in higher levels of motivation, self-efficacy, and growth for students who are at risk, therefore producing a higher median SGP for schools using them.
- Research Question 3: Do standards-based report cards that separate academic progress in mathematics from social and learner behaviors result in higher levels of growth and performance on the MCAS? Hypothesis 3: Standards-based report cards that separate mastery of content knowledge and skills from learner behaviors lead to higher levels of

growth and proficiency than amalgamated grades on traditional report cards, therefore producing a higher median SGP and higher CPI for schools using them.

- Research Question 4: Do standards-based report cards that grade in power standards result in higher levels of growth and performance than report cards with a cumulative grade or standards-based report cards that do not provide power standards? Hypothesis 4: Standards-based report cards that clearly identify power standards in math, specific to Grade 4, result in higher levels of growth and performance than report cards with one overall content area grade, therefore producing a higher median SGP and CPI.

This research project explores evidence that providing students with feedback on their performance on the mathematics standards through standards-based report cards improves their performance on the Mathematics MCAS. The study draws inferences regarding the impact of standards-based report cards on improving outcomes for students by gathering data on Student Growth Percentile (SGP) and the Composite Performance Index (CPI) on the Massachusetts Comprehensive Assessment System for Mathematics. This study examines the components of standards-based report cards and makes inferences regarding the effect of standards-based grading practices on self-efficacy, motivation, and achievement.

Definitions

The operational definitions developed for this study are based on commonalities among researchers Robert Marzano, Richard Stiggins, Thomas Guskey, and Larry Ainsworth. Empirical evidence is provided within the study where it assists with developing a shared understanding of the terminology. Some constitutive definitions are taken from the Massachusetts Department of Elementary and Secondary Education website and the text *How to Design and Evaluate Research in Education* by Fraenkel and Wallen (2009).

- *Causal-comparative research* – Research to explore the cause for, or consequences of, existing differences in groups of individuals; also referred to as ex post facto research (Fraenkel & Wallen, 2009).
- *Composite performance index* - A measure of the extent to which students are progressing toward proficiency. The CPI is a 100-point index that combines the scores of students who take standard MCAS tests (the Proficiency Index) with the scores of those who take the MCAS-Alternate Assessment (MCAS-Alt) (the MCAS-Alt Index). CPI scores range from 0 to 100, and correspond to one of the six Performance Rating categories (MA DESE, 2011).
- *Criterion-referenced instrument* – An instrument that specifies a particular goal or criterion for students to achieve (Fraenkel & Wallen, 2009).
- *Massachusetts Comprehension Assessment System*- The Massachusetts Comprehensive Assessment System (MCAS) is designed to meet the requirements of the Education Reform Law of 1993. This law specifies that the testing program must test all public school students in Massachusetts, including students with disabilities and English Language Learner students; measure performance based on the Massachusetts Curriculum Framework learning standards; and report on the performance of individual students, schools, and districts. As required by the Education Reform Law, students must pass the Grade 10 tests in English Language Arts (ELA), Mathematics, and one of the four high school Science and Technology Engineering tests as one condition of eligibility for a high school diploma (in addition to fulfilling local requirements). In addition, the MCAS program is used to hold schools and districts accountable, on a yearly basis, for

the progress they have made toward the objective of the No Child Left Behind Law that all students be proficient in Reading and Mathematics by 2014. (MA DESE, 2011)

- *Standards* - Statements which indicate what students are expected to know and be able to do in each curriculum area.
- *Standards-Based Report Card* - An alternate way of reporting student progress to parents which involves assessing student proficiency in alignment with the state/local standards and benchmarks. For the purposes of this study the operational definition of a standards-based report card includes meeting the criteria developed and reported in the Standards-Based Report Card Grouping Rubric included in Appendix D.
- *Student Growth Percentile (SGP)* - A measure of student progress that compares changes in a student's MCAS scores to changes in MCAS scores of other students with similar achievement profiles. The model establishes cohorts of students with "similar performance profiles" by identifying all students with the same (or very similar) MCAS scores in prior years.

Conceptual Framework

The conceptual framework for this study is established on the assumption that providing students with effective feedback through standards-based report cards results in improved student learning. Providing feedback to students on their learning helps them to see and understand the path they will need to follow to achieve proficiency in knowledge and skills. Richard Stiggins (2005) describes how students become partners in the learning process because they are equipped to monitor where they are and how they can work toward the next level. Students take an active role in recognizing their level of learning and communicating their

understanding to teachers and families. This study examines the effect of providing student feedback in the form of standards-based report cards.

Standards-based grading establishes a framework for providing feedback to students on their achievement levels within a set of standards. Student learning standards depicted on report cards are intended to indicate what students should know and be able to do at each of the grade levels or grade spans across their educational experience. The use of standards-based report cards enables teachers, students, and parents to track success continually (Marzano, 2003). Students receive a grade or performance level within the standard or strand that is being taught on periodic assessments throughout the school year. Guskey (2004) reports on the evidence that standards-based report cards account for what students know and are able to do. Aidman, Gates, and Deterra Sims (2001) further support that parents who receive a standards-based report on their child indicate that the report card shows them what their students know and can do. The documentation of mastery of specific skills in reading and math particularly has been a focal point of standards-based report cards (American School Board Journal, 2006; Silberglitt, Burns, Madyun & Lail, 2006). As reported by each of these studies, accurate reporting of mastery is essential information for parents and students given the high-stakes accountability measures mandated by NCLB in reading, math, and science.

Standards-based reports provide information about student progress but may take many different forms. The report itself can take the form of age- or grade-appropriate expectations expressed in benchmarks, complete standards, open-ended statements, or abbreviated focal points (Salend, 2005). The sample report card in Figure 1 shows the report card indicators for three trimesters for the Number Sense and Operations strand on one school's report card for Grade 4.

Trimester 1	Trimester 2	Trimester 3	Proficiency Level		
			1	2	3
Mastery of multiplication through 10 x 10 and related division facts	Mastery of multiplication through 10 x 10 and related division facts	Mastery of multiplication through 12 x 12 and related division facts			
Add and subtract (up to 5 digit numbers) and multiply (up to 3 digit numbers by 2 digit numbers) accurately and efficiently	Solve multiplication and division number stories	Select, use, and explain models to relate common fractions and mixed numbers ($\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{5}$, $\frac{1}{6}$, $\frac{1}{8}$, $\frac{1}{10}$, and $\frac{1}{12}$), find equivalent fractions, mixed numbers and decimals, and order fractions			
Exhibit an understanding of the base ten number system by reading, writing, modeling, and interpreting whole numbers to at least 100,000; demonstrating an understanding of the values of the digits and comparing and ordering numbers	Demonstrate an understanding of and the ability to use the conventional algorithm for division of up to a three-digit whole number with a single-digit divisor (with or without remainders)	Identify and generate equivalent forms of common decimals and fractions less than one whole (halves, quarters, fifths, and tenths)			
Demonstrates the ability to explain mathematical thinking both orally and in writing	Select and use a variety of strategies to estimate quantities, measures, and the results of whole number computations up to three-digit whole numbers and amounts of money up to \$1,000	Demonstrates the ability to explain mathematical thinking both orally and in writing			
	Demonstrates the ability to explain mathematical thinking both orally and in writing				
	Round whole numbers through 100,000 to the nearest 10, 100, 1,000, 10,000 and 100,000				

Figure 1. A sample section of a Grade 4 standards-based report card for the Number Sense strand of mathematics.

Report cards can vary greatly from school to school. The conceptual framework of this study establishes a continuum of report card types. Report cards were placed along a continuum from more traditional, non-standards-based report cards to report cards that mix features of both traditional and standards-based to the farthest end of the continuum which includes report cards that fit this study's definition of standards-based.

The Massachusetts Comprehensive Assessment System (MCAS) is designed to meet the requirements of both the Education Reform Law of 1993 and NCLB. Annual assessments are required for students in grades 3 to 8 in Mathematics and Reading. High school students must attain a level of proficiency in Mathematics, Reading, and Science in order to be eligible to receive a high school diploma. School districts receive annual MCAS reports on the performance of their district, individual schools, and specific grade levels. The data that districts receive provide two measures of student performance. The first is the Composite Performance Index. The second is the more recently developed Student Growth Percentile.

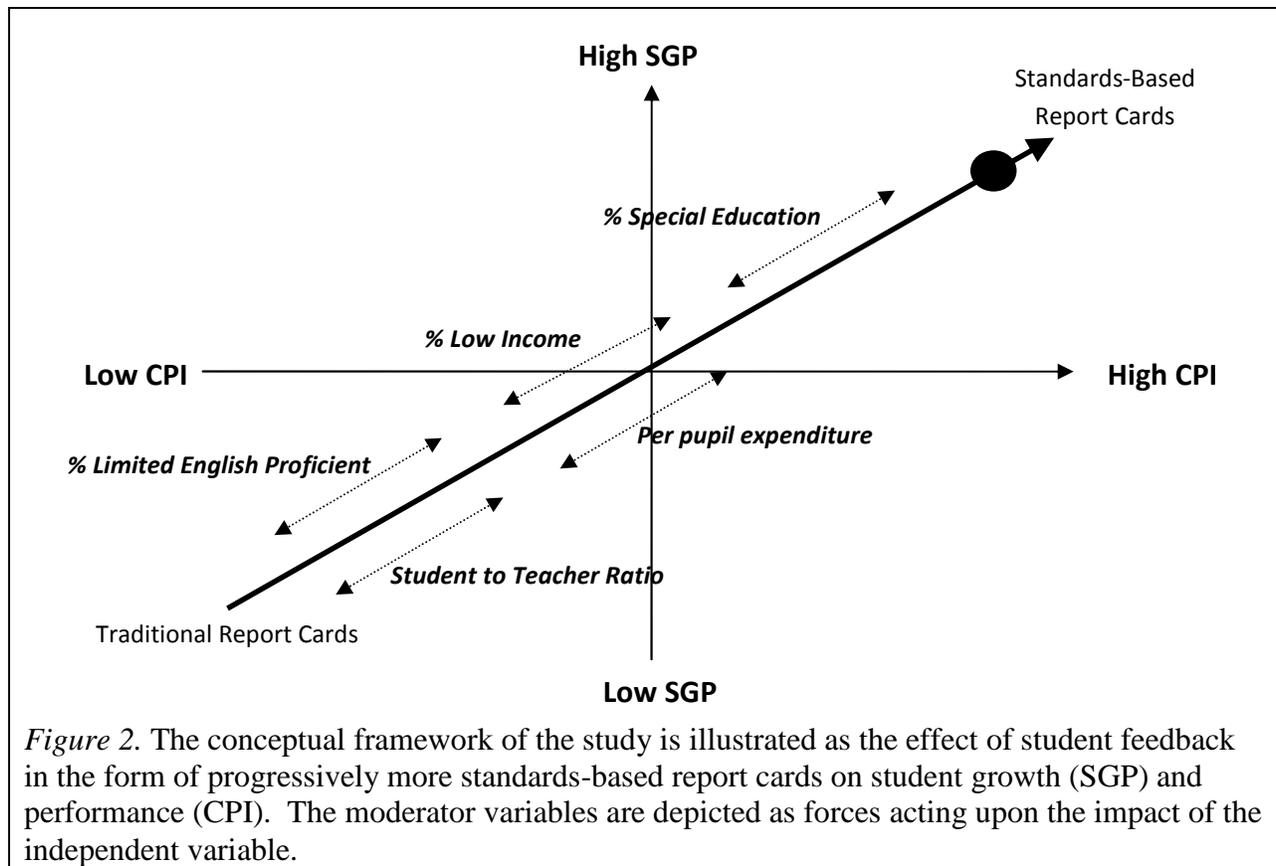
The Composite Performance Index (CPI) is defined by the DESE as a measure of the extent to which students are progressing toward proficiency in English Language Arts (ELA) and Mathematics, respectively. The CPI is a 100-point index that combines the scores of students who take standard MCAS tests (the Proficiency Index) with the scores of those who take the MCAS-Alternate Assessment (MCAS-Alt). CPI scores range from 0 to 100 and correspond to one of the six Performance Rating categories listed in Appendix E.

The Massachusetts Department of Elementary and Secondary Education (2009) describes the more recently developed Student Growth Percentile (SGP) as:

“...a measure of student progress that compares changes in a student's MCAS scores to changes in MCAS scores of other students with similar achievement

profiles. The model establishes cohorts of students with “similar performance profiles” by identifying all students with the same (or very similar) MCAS scores in prior years; all MCAS data for a student since 2006 are used (where available) to establish academic peers” (p.1).

Student growth percentiles can also be aggregated to understand growth at the subgroup, school, or district level. The most appropriate measure for reporting growth for a group is the median student growth percentile (the middle score if one ranks the individual student growth percentiles from highest to lowest). The use of the median SGP informs this study as to the changes in growth of students in each school in relation to the independent categorical variable of standards-based report cards. This study uses the grade level SGP and CPI reported on the MCAS as measures of growth and performance. This study’s framework represents the continuum as an ascending line of progression toward a report card that more closely represents the best practices of standards-based report cards and results in higher levels of growth and performance as shown in Figure 2. The measures of student performance that are impacted by report card type are represented as high or low SGP and high or low CPI.

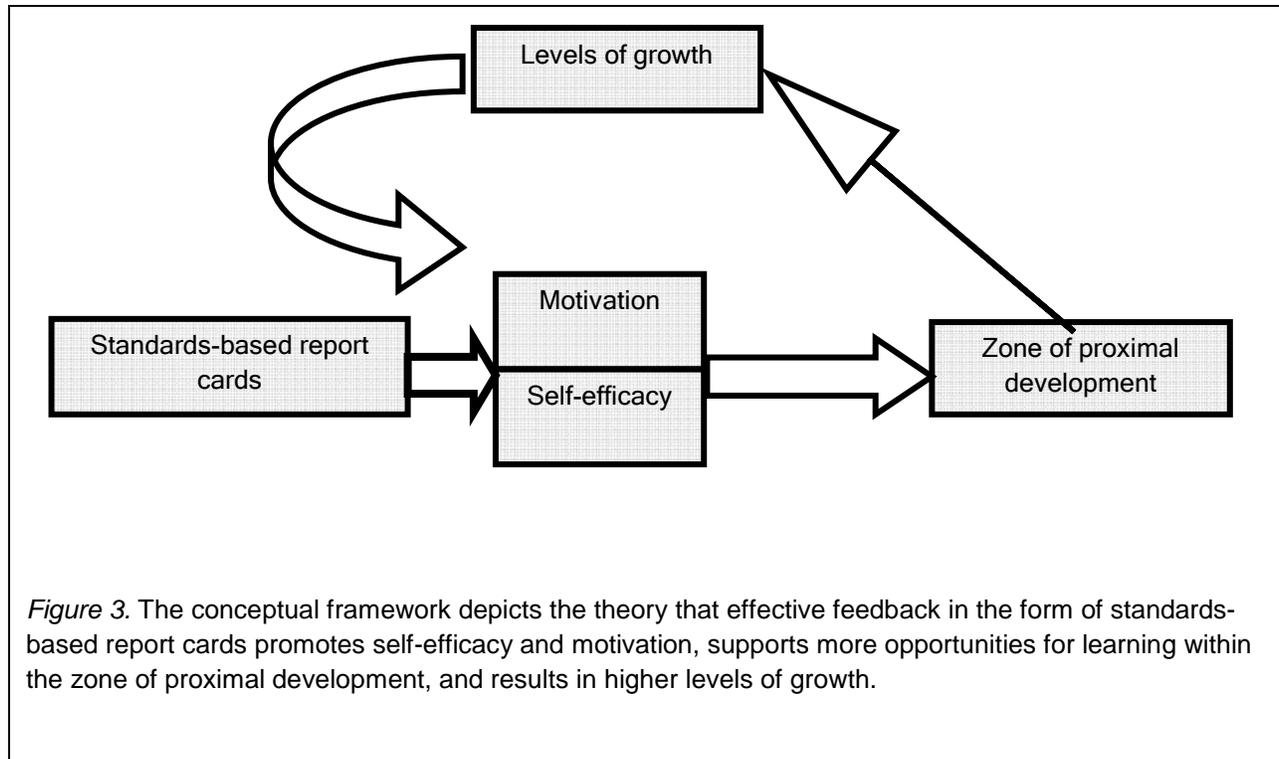


This study includes three covariates that represent subgroups identified by the Department of Elementary and Secondary Education (DESE). They are presented as the percentage of students in each school that qualify as low income, the percentage of students who have been identified with a disability, and the percentage of students who have limited English proficiency. Trends in the results of the MCAS show that students from low income homes, students in special education, and students who have limited English proficiency over the past 10 years have shown slower growth and lower levels of performance as subgroups than the general population (MA DESE, 2011). Two additional covariates are controlled in this study. The variable of per pupil expenditure is included because students from schools with higher per pupil expenditure typically perform better on high-stakes testing. The final variable of student-teacher ratio is

included because results show higher levels of performance for students in schools with lower class size and more capacity for individualized instruction. These five covariates are depicted in Figure 2 as having a positive or negative influence on the progress that is measured and attributed to the use of standards-based report cards.

Theoretical Framework

The theoretical framework of this study is developed from theories of self-efficacy, the zone of proximal development, and standards-based instruction. The underlying theory is that feedback provided by standards-based report cards will improve self-efficacy and increase motivation, thereby resulting in increased opportunities for students to be in a zone of proximal development. The increased opportunities to learn within the zone of proximal development are evidenced by expanded growth and improved performance. This research project is grounded in the theory that providing an achievable goal (i.e. the grade level content standards and indicators), removing designations of failure and informing students of their progress along a continuum promotes self efficacy and motivation, provides opportunities for students to learn in a zone of proximal development, and results in improved performance and higher levels of growth as illustrated in Figure 3.



The theoretical framework that provides the foundation for this project encompasses a review of the work of Albert Bandura, Lev Vygotsky, and Richard Stiggins. Albert Bandura's work is the primary theoretical framework that informs this study of standards-based report cards. Bandura (1993) provided a foundation for the current movement to empower students to monitor their own progress through the curriculum. The Social Cognitive Theory shaped by Bandura outlined how children and adults operate cognitively on their social experiences and how these cognitive operations then come to influence their behavior and development. Children begin to represent their environment and themselves in terms of response-outcome expectancies, perceptions of self-efficacy, and evaluative self-reactions (Grusec, 1992). These cognitive processes set the stage for a student's view of self and the opportunities, challenges and experiences that they seek out for their learning and eventual living environments. Beginning in the elementary grades students begin to shape their own view of self and their perception of their

ability to succeed in their environment as determined by the grades they receive on report cards. Strobino, Gravitz, and Liddle (2002) note that students experience grades as rewards or punishments that shape their perceptions and behaviors. Students become conditioned to the “extrinsic rewards that grades convey. And so, it seems that the experiences one has throughout the years may serve as an impetus for the efforts one expends and also for the motivation to engage in the learning process” (Strobino, et. al., 2002, p. 146). In order to become more confident, proficient learners, students need grading procedures that explicitly provide opportunities for improvement in order to strengthen images of self-efficacy and motivation.

Students must have purposeful intentions and self-motivation in order to take the actions that are necessary to improve achievement. Bandura’s Social Cognitive Theory defines agency as acts done intentionally, which involve “not only the deliberative ability to make choices and action plans, but the ability to give shape to appropriate courses of action and to motivate and regulate their execution” (Bandura, 2001, p. 7). Schools are increasingly utilizing educational practices that not only inform student self-monitoring in the development of action steps to improve performance, but also promote motivation to engage in those action steps and encourage the corrective self-regulation that is essential to moving along a continuum of growth. The use of standards-based report cards is one such research-based practice that informs students of their level of proficiency and promotes the development of future-directed plans. The self-regulative sub-functions of motivation, affect, and action are dependent upon constructive feedback on progress such as that depicted on standards-based report cards.

Bandura’s theory of self-efficacy proposes that people develop “domain-specific beliefs about their own abilities and characteristics that guide their behavior by determining what they try to achieve and how much effort they put into their performance in that particular situation or

domain” (Grusec, 1992, p.782). When considering traditional grading practices, Bandura’s theory is applicable to students who struggle with learning and are rewarded for their efforts with failing grades. Their sense of self-efficacy is impacted by the recognition that others readily receive good grades for similar effort, or less, than they expend to receive failing grades.

Children who are disadvantaged must work the hardest to progress in learning but have the least incentive to do so in grading systems that identify them as failures. “These students find the relationship between high effort and low grades frustrating and often express their frustration with indifference, deception, or disruption” (Guskey, 2006, p. 673). This recognition may lead students who are disadvantaged to make a decision to expend less effort, become less motivated, and begin a vicious cycle of failure. In a review of the literature on self-efficacy beliefs and achievement in reading and writing Frank Pajares (2003) emphasizes the theory of Bandura and subsequent researchers stating that the “focus on students’ self-beliefs as a principal component of academic motivation is grounded on the assumption that the beliefs that students create, develop, and hold to be true about themselves are vital forces in their success or failure in school” (p. 140). There cannot be a more profound impact on the self-belief of a student than to receive a grade report that depicts them as a failure.

There are four primary sources of self-efficacy as identified by Bandura. The sources are: mastery experience; vicarious experience as the performance of others is observed (modeling); verbal messages and social persuasion; and physiological states such as anxiety and stress. Positive experiences help to build self-efficacy which strengthens motivation and empowers the student. Negative experiences defeat self-efficacy and weaken motivation, resulting in stress and anxiety, and thereby reducing the student’s learning effectiveness. Pajares (2003) restates the development of self-efficacy simply as “outcomes interpreted as successful raise self-efficacy;

those interpreted as failures lower it” (p. 140). Bandura highlights the necessity of informing struggling students of progress as he notes that “unless people believe they can produce desired results and forestall detrimental ones by their actions, they have little incentive to act or to persevere in the face of difficulties” (Bandura, 2001, p. 7). A child’s beliefs about his or her ability to affect a difference through his or her own efforts influence sense of self-efficacy and can hinder or enhance future efforts. Motivation to achieve goals and expect positive outcomes are centrally influenced by efficacy beliefs (Bandura, 2001). “It is partly on the basis of efficacy beliefs that people choose what challenges to undertake, how much effort to expend in the endeavor, how long to persevere in the face of obstacles and failures, and whether failures are motivating or demoralizing” (Bandura, 2001, p.7). Enhancement of perceived self-efficacy could improve children’s cognitive skill development and their intrinsic interest in academic subjects as noted by Grusec (1992). This alone should be enough reason to move from traditional grading systems, which ‘fail’ students starting in the early grades, to a more motivating system of grading that places students along a continuum of acquisition of knowledge and skills.

The connection between increased sense of self-efficacy and motivation, and higher achievement seems almost intuitive when considering the process of student learning. Grusec (1992) postulates that Bandura’s Social Cognitive Theory has become less evident in research over time because its basic concepts of observational learning have become an “accepted part of our knowledge base” (p. 776). Lev Vygotsky provides the foundation for another accepted part of the knowledge base; students have optimal potential for learning when they are in their zone of proximal development. The students’ opportunities for optimal learning are most effective when they are motivated and challenged to move along a continual process of learning.

Vygotsky established that students operating within their zone of proximal development learn concepts and solve problems in a particularly efficient manner. Vygotsky's zone of proximal development is defined as the distance between the collective group problem solving possible for a child today and the independent problem-solving, emerging from that group context, which will be possible tomorrow. Vygotsky theorized that development of mastery involved transforming thinking from an absence of self-regulation to self-regulation and egocentric speech (Vygotsky, 1929 cited in Gredler, 2009). Students must have information about their own achievement in order to regulate their learning, build self-efficacy, maintain motivation, and continue to grow as learners. Standards-based teaching and assessment practices facilitate a zone of proximal development in classroom settings. As Vygotsky noted in his later research, "students rely on the information and feedback provided in learning situations in order to develop the inner regulation of purposeful activity" (Vygotsky & Luria, 1944 cited in Gredler, 2009, p. 6). Standards-based report cards are guiding factors for an informed learner to practice egocentric speech regarding the progress made and the amount of mastery still needed in relation to the curriculum standards.

The theoretical framework of both Bandura and Vygotsky generates additional questions about how informing learners of their progress through standards-based report cards may also be empowering students to promote growth and improve achievement. If standards-based report cards improve self-efficacy and increase motivation, thereby resulting in increased opportunities for students to be in their zone of proximal development, then the result should be improved achievement and increased growth for those most at risk for learning.

Dr. Richard Stiggins developed a modern theory of educational practice encompassing self-efficacy and motivation which informs this study of the effect of standards-based report

cards. His theory of Assessment for Learning has been the catalyst for the development of meaningful assessment and reporting of progress to students and families. Stiggins (1997, 2005) identifies the essential element of learning as bringing all students to believe in themselves. Stiggins argues that assessing and reporting systems that serve to punish and demoralize students must be replaced by others that “promote hope and continuous effort” (Stiggins, 2005, “A New Mission,” para. 3). Stiggins’ promotion of instilling hope and continuous effort in learners mirrors almost explicitly Bandura’s cognitive learning theory of building student self-efficacy and motivation and Vygotsky’s egocentric speech and internal regulation that form optimal learning opportunities in the zone of proximal development. Stiggins emphasizes that it is not the teachers, school leaders, or policy makers who decide whether students will succeed; it is the students. “Students are deciding whether success is within or beyond reach, whether the learning is worth the required effort, and so whether to try or not” (Stiggins, 2005, “The Old Mission,” para. 5). The students experience anxiety, fear of failure, uncertainty, and unwillingness to take risks based on their perceptions of their own capabilities. Stiggins (1997; Stiggins & Chappuis, 2008) has refocused educational assessment on two elements – formative and summative assessments. Formative assessments are periodic interim assessments that provide both the teacher and the student with an understanding of progress toward proficiency in grade level standards. Summative assessments give a picture of mastery at a particular point in time. Stiggins implores us to re-examine grading practices in order to provide opportunities for all students to demonstrate success at some level. “Procedures that permitted (perhaps even encouraged) some students to give up in hopelessness and to stop trying must now be replaced by others that promote hope and continuous effort. In short, the entire emotional environment surrounding the prospect of being evaluated must change, especially for perennial low achievers”

(Stiggins, 2005, “A New Mission,” para. 3). As district leaders engage in revisions to local curriculum to align with the Common Core State Standards Initiative this is the opportune time to revisit grading practices and adopt the tools that will best facilitate student learning (NGA Center & CCSS, 2010). This research study will examine whether standards-based report cards are a viable tool for promoting hope and continuous effort in order to improve growth and performance.

Research Design

This research study employs an unobtrusive collection of data. Data collected includes: Grade 4 Student Growth Percentile (SGP) and Composite Performance Index (CPI) data available from the DESE *District Profiles of Assessment and Accountability*; information gathered from questionnaires returned from elementary schools; empirical data collected from report cards; and product rating scale data recorded for each report card collected.

The hypothesis that drives this research study is: Fourth grade results for students in the schools with standards-based report cards will have higher percentages of growth and higher indices of performance as measured by the MCAS Mathematics Test than schools with non-standards-based report cards. Research indicates that standards-based report cards should feature (a) the possibility to progress along a continuum of performance, (b) the removal of failing grades, (c) the separation of assessment of learner and social behaviors from grades for content knowledge and skills, and (d) a focus on grade level standards and indicators. The Standards-Based Report Card Product Rating Scale used in this study was developed in accordance with the research of Robert Marzano, Richard Stiggins, Thomas Guskey, and Larry Ainsworth with a primary focus on the guidelines provided by Thomas Guskey and Jane Bailey (2010) in *Developing Standards-Based Report Cards* and Douglas Reeves and Larry Ainsworth in

Ainsworth's 2003 text *Power Standards, Identifying the Standards that Matter Most*. Guskey and Bailey (2010) identify how reporting standards must be differentiated from curriculum standards and provide 5 criteria which inform the rating scale developed for this study:

1. Designed for reporting on student learning not planning instruction and assessments;
2. Relatively few in number (usually 4-6 per subject) while curriculum standards are many in number (10-50 per subject), (*Doug Reeves recommends 7-12 in Ainsworth, 2003*);
3. Broad and more general than the highly specific curriculum standards;
4. Clear and understandable as opposed to complicated and detailed;
5. Expressed in parent-friendly language instead of the complex educator language. (Guskey & Bailey, 2010, p. 37)

The document analysis conducted in this study utilizes a criterion group design as an extension of the causal-comparative study to determine if the dependent variables are affected by the categorical variable of the standards-based report cards that are collected for this study. The Standards-Based Report Card Product Rating Scale developed for the document analysis in this study is included as Appendix C. This rating scale is a significant limitation of the study as it has not been utilized prior to this study. The researcher developed the rating scale for the purpose of setting criteria for sorting the report cards used in the study into independent variable groups.

The study controls for several covariates among the categorical independent variables: percentage of students in the low income, special education and limited English proficient categories; student-teacher ratio, and per pupil expenditure. This demographic data is gathered

from the *School and District Profiles* directory within the DESE website by selecting the town and then selecting *Enrollment/Indicators*.

This is a causal-comparative study of the effect of standards-based report cards on student growth percentiles and composite performance index scores on the Massachusetts Comprehensive Assessment System. There is one categorical independent variable: standards-based report card. A product rating scale developed for the study is used to delineate the categorical variable of standards-based report card according to quantitative variables in a criterion-group design. The study is conducted at the organizational level and examines the effects of standards-based report cards on growth and performance of students in Grade 4 in mathematics in elementary schools in southeastern Massachusetts.

Limitations of the Study

The major threat to internal validity of a causal-comparative design is the possibility of a subject characteristics threat (Fraenkel & Wallen, 2009). This threat can be reduced by matching subjects from the comparison groups on the extraneous variable that may be influencing the results, finding or creating homogeneous subgroups, or statistical matching. The most serious limitation of a causal-comparative design is the lack of control over threats to internal validity. Because the manipulation of the independent variable has already occurred, many controls cannot be applied. Relationships can be identified but causation cannot be fully established. Additional threats include “loss of subjects, location, instrumentation, and sometimes history and maturation” (Fraenkel & Wallen, 2009, p. 368).

To control for subject characteristics threats this study created three groups of relative homogeneity. All members of the independent variable for this study have the following characteristics: participants are from southeastern Massachusetts and have MCAS that results are

reported for Grade 4 mathematics in 2010. In order to reduce the loss of subjects threat, a gift card was included in the mailing to all elementary principals along with the questionnaire to solicit participation in the study. An internal threat to the validity of the study was the possibility of a small comparison group of schools in southeastern Massachusetts that continue to use traditional report cards. The actual return for non-standards-based report cards ($N = 30$) is considered a valid number for comparison to the number of standards-based report cards ($N = 52$) and the number of mixed report cards ($N = 20$).

This study includes several covariates as an additional method to control for causal relationships that may be impacting results and to increase internal validity. The impact of socio-economic status, special education needs, limited English proficiency, student-teacher ratio, and per pupil expenditure is evident in the results of the MCAS over the past decade. Each of these factors is included as a covariate in the data analysis to reduce the likelihood that the causal relationship of each factor may be influencing the results for the independent variable. The limitations of this study are the following:

1. This research study is limited to Grade 4 mathematics results on the Massachusetts Comprehensive Assessment System. Grade 4 mathematics was selected for study because the assessment has been administered annually since 2002 and has the most data available. Educators in Grade 4 have the most experience in administering the MCAS test and responding to the results. This study does not include results for other grades nor does it include any other content area assessments.
2. The study is limited to one year of data (2010) and includes schools that may have engaged standards-based report cards for varied periods. The initial questionnaire was designed to gather information on length of implementation as a moderating variable but

few respondents provided this information so the data was not considered in the final analysis.

3. The sample size is a limitation of the study as it represents 48% of the elementary schools in the targeted area of southeastern Massachusetts and approximately 9% of the elementary schools in the state.
4. The study does not account for standards-based practices that may or may not be in place in the classrooms of participants.
5. The rating of the report cards according to the ordinal rating scale and the coding of report cards was completed by the researcher only.
6. The Standards-Based Report Card Product Rating Scale developed for this study has not been utilized before and is a significant limitation of the study.

Delimitations of the Study

This study analyzes data from participating public elementary schools within southeastern Massachusetts that are educating students in grade 4. The study does not encompass any charter schools due to the difference in funding structure that would have affected use of the covariate of per pupil expenditure. While ethnicity is a reported demographic statistic that is available through the DESE, it is not utilized as a covariate in this study. The public elementary schools were not disaggregated by type of district – urban, suburban or rural. The study does not examine factors such as grading policies, professional development around standards-based instruction and grading. This study did not examine information on student performance in years prior to 2010.

Chapter 2: Literature Review

This review of scholarly literature outlines the primary features of standards-based report cards and the supporting research that standards-based practices influence the success of students. The literature review is introduced with a brief review of the historical perspective on grading, standards-based teaching and the accountability movement. The historical perspective is followed by an examination of the current trends in standards-based grading and reporting that relate to this study, specifically: grading along a continuum of progress; elimination of failing grades; separation of learner behaviors from academic achievement; and the determination of standards and indicators. The impact of trends in standards-based grading practices on student self-efficacy, motivation, and achievement is interwoven throughout the literature review. The relative weakness of traditional grading practices as compared to trends in standards-based grading is explored within the context of each subheading. The literature review concludes with a brief summary of the resistance to standards-based grading.

Historical Perspective

Grading practices. “A single letter grade or a percentage score is not a good way to report student achievement in any subject area because it simply cannot present the level of detailed feedback necessary for effective learning” (Marzano, 2000, p. 106). As surprising as it may seem to many parents, teachers, administrators, and educators of higher learning, grades did not always exist for students. Prior to the mid-nineteenth century students were judged by their progress within content areas and skills (Hargis, 1990). Student progress was documented through narratives or statements of skills and knowledge in order to determine areas for further instruction, to determine competency for apprenticeship, or to determine eligibility for entrance

into higher education. Parents and students received an accurate picture of their progress, knowledge, and skills. There were no grades issued in the form of A, B, C.

As the number of government-supported elementary schools and high schools began to increase during the 40 years between 1870 and 1910, the number of public high schools increased from 500 to 10,000 (Hargis, 1990; Marzano, 1996). This sudden influx of student attendees necessitated a move toward the more streamlined, industrial model of percentage scales and letter grades for recording progress in high school. This industrialized model of reporting progress was the beginning of the grading system that students, parents, and teachers find comfort in today:

Today's adults grew up in schools designed to sort us into the various segments of our social and economic system. The amount of time available to learn was fixed: one year per grade. The amount learned by the end of that time was free to vary: some of us learned a great deal; some, very little. As we advanced through the grades, those who had learned a great deal in previous grades continued to build on those foundations. Those who had failed to master the early prerequisites within the allotted time failed to learn that which followed. (Stiggins, 2005, para.1)

It appears that the educational pendulum is swinging away from what has been recognized as *traditional* reporting in the form of letter grades or percentages back toward the narrative descriptions of mastery that actually originated in the early years of public education in the United States (Hargis, 1990; Marzano, 1996). Thomas Guskey (in Guskey & Bailey, 2010), a leader in the movement toward standards-based report cards, notes that questionable grading practices in traditional grading include: averaging to obtain a student's grade or mark; assigning a

'zero' to work that is late, missed, or neglected; and taking credit away from students for behavioral infractions. The current movement in reporting and grading is a return to the more descriptive documentation of progress through narratives, statements of curriculum expectations, open-ended statements, and portfolios (Guskey, 2004a; Allen 1998; Salend, 2005). The term that has been given to this more comprehensive form of documenting progress as it now surfaces is *standards-based report cards*. While educational leaders in Massachusetts have recognized the need for better reporting systems and responded to them by developing a range of reporting tools that vary district to district, there is limited evidence to support the supposition that the new reporting systems impact student performance on the high-stakes criterion-referenced assessments used by the State to measure performance.

The use of letter grades is firmly entrenched in today's society and will most likely remain so at the secondary level for the distant future (Marzano & Kendall, 1996). The primary factor in maintaining traditional grades has been their simplicity. Years of utilization of the letter grading systems has built familiarity that has made the system quick to record and easy to understand (Hargis, 1990). The A-B-C system of grading in content areas has attained a cult status. Parents, teachers, community members, and college admissions officers refer to students in terms of an A-student or C-student. The shift away from this familiar paradigm to a new system at the secondary level will require conclusive evidence that the change will positively affect student achievement.

Standards-based teaching. In order to discuss standards-based report cards it is necessary to first review standards-based teaching. Researchers share a common definition of standards-based teaching: the teaching toward mastery of skills and content knowledge that all students should be able to demonstrate as they progress through the grades (Aidman, 2001;

Guskey, 2004a; Marzano, 1996). Standards-based teaching has its roots in the call for uniform education policies that began in the early 1980s with the publication of *A Nation at Risk*. This Commission report indicated a need to focus on raising standards, implementing standardized tests, and holding students and teachers accountable (United States National Commission on Excellence in Education, 1983). The movement toward standards-based teaching was accelerated by education reform in the 1990s (Marzano, 1996). The Education Reform Act of 1993 in Massachusetts directed the development of a common core of learning for all students which then developed into the Massachusetts Curriculum Frameworks [Massachusetts Department of Elementary and Secondary Education (MA DESE), 2010]. The Frameworks provided the grade span indicators of learning for all students in the Commonwealth of Massachusetts from 1993 to 2010. As of July 2010, Massachusetts adopted the national Common Core State Standards, which with the addition of several Massachusetts standards, have become the new Massachusetts Curriculum Frameworks Incorporating the Common Core State Standards for English Language Arts & Literacy, and Mathematics.

The body of research supporting the implementation of standards-based grading is based on three primary elements that impact classroom practice: improved communication of student progress, the standardization of grading benchmarks, and the centralization of assessment by the teacher (Aidman, Gates & Deterra Sims, 2001; Guskey, 2004a; Hargis, 1990; Marzano, 2003; Marzano & Kendall, 1996; McTighe & Bailey, 1996; Salend, 2005; Stiggins, 2005, Ward, 2004). The use of standards-based report cards enables teachers, students, and families to track success continually (Marzano, 2003). All students within a grade level should have the same opportunity to show what they know and can do. The lack of a clear standards-based curriculum disproportionately hurts children of poverty, children of color, and children with disabilities

(Carter, 2007). The old methods of averaging grades, punishing students for poor learner behaviors with poor grades and rewarding good behavior with good grades, must be abandoned for a new paradigm. Traditional report cards provide grades that represent student achievement in relation to other students in the classroom and allow for high levels of achievement to be parceled to only a few along a bell curve. Standards-based report cards are tools for reporting a progression toward competency or proficiency. Each student's performance should be interpreted relative to established instructional goals and standards, independent of other students' performances (Glaser & Nitko, 1971, Lalley & Gentile, 2009, Popham, 1978). Performance must be consistently reported across grades and schools through standards-based report cards to inform students and parents of mastery that students have demonstrated within the curriculum and to promote progress toward higher levels of achievement.

Accountability movement. Early in public education Horace Mann traveled the state of Massachusetts assessing the status of schools using criteria that emphasized the educational inputs: how many books were available, the number of programs employed, the availability of library resources, the physical status of the building, etc. The Coleman Report released in 1966 became the impetus for a shift from evaluating educational inputs to evaluating learning outcomes (Marzano, 1996). Assessment of learning outcomes has gained momentum since the publication of *A Nation at Risk* and the reauthorization of the Elementary and Secondary Education Act now known as the No Child Left Behind Act of 2002. The NCLB legislation has called upon educators to raise all students to a specified level of competence in academic achievement standards and will hold states, districts, and schools accountable for making sure that all students meet them (Stiggins, 2005). Over the past decade 49 of the 50 states have established testing requirements in accordance with NCLB (Hursh, 2005). Testing is now a

federal mandate for students in grades 3 to 8 in Mathematics and English Language Arts.

Science and Technology Engineering tests are administered in grades 5 and 8, and then again in Grade 9 or 10. Attaining a level of Proficiency on the English Language Arts; Mathematics; and Science, Technology and Engineering MCAS tests is required for Massachusetts high school students in order to meet the competency determination for a high school diploma.

For students in the lower grades, performance below proficiency means that an Individual Student Success Plan (ISSP) is developed for them which may indicate remedial services, additional support in the classroom, or specific interventions that will be available to improve their MCAS performance. These consequences qualify the MCAS as a *high-stakes assessment* and have created a sense of urgency to develop assessment, grading, and reporting systems that will mark student progress in the standards more frequently than the annual test and will support improved performance.

The impact of the students' performance on these high-stakes assessments extends far beyond the students themselves. Although the MCAS tool was designed to inform schools of areas of curriculum strengths and weakness and was intended to be a formative assessment, the test has actually become a tool for rating schools and districts and imparting punitive measures on those that rank below the accountability formula established by the state. The implications of poor performance on MCAS are devastating for individual students, their families, their schools, and their districts as more and more schools find themselves in the accountability status of *Needs Improvement, Corrective Action, or Restructuring*. As a result of the performance of schools on the 2010 MCAS 40% of schools in Massachusetts have an accountability status. NCLB mandates that students have the option to transfer from schools with low test scores to those with higher test scores. Resources are directed to under-performing student subgroups and specific

content areas. Grant funding is outsourced to private agencies with the goal of improving student performance. At the same time that sanctions are being imposed school districts are continuing efforts to make improvements and offer evidence that student learning is progressing. Policy makers argue that teacher bias and subjectivity are responsible for the “mismatch between grades and scores on accountability assessments” (Guskey, 2006). Therefore, it makes sense to align district-wide grading practices with state-wide assessment practices in order to clearly identify students’ strengths and weakness in the mastery of standards. This study examines the MCAS results of schools that have retained traditional grading practices in comparison to those that have adopted standards-based report cards to provide evidence that the efforts to match local grades with scores on accountability assessments are impacting student achievement.

Impact of Grading Practices on Self-Efficacy and Motivation

Beliefs about self-efficacy arise from an individual’s history of achievement, from observation of what others are able to accomplish, from attempts of others to mold feelings of self-efficacy through persuasion, and from consideration of one’s own physiological state during a task (Grusec, 1992). Researchers have noted the impact of feedback on student learning:

From the very earliest grades, some students learned a great deal very quickly and consistently scored high on assessments. The emotional effect of this was to help them to see themselves as capable learners, and so these students became increasingly confident in school...They believed that success was within their reach. (Stiggins, 2005, “The Old Mission,” para. 2)

Students who score low on tests and receive poor grades beginning in early elementary grades will question their own capabilities as learners and enter into a vicious cycle of reduced motivation, compromised views of self-efficacy, and correspondingly lowered achievement.

Stiggins emphasizes that for those students with compromised views of self-efficacy performance plummets as their motivation declines (2005). Lalley and Gentile (2009) summarize the foundation for grading along a continuum toward mastery with their observation that “not all students get it the first time and need additional attempts, and perhaps other methods or examples, before they try and try again” (p. 33).

It is no longer acceptable or tolerable that educators allow some students to give up in hopeless failure. The performance levels of ‘approaching grade level standard’ or ‘below grade level standard’ provide for a more hopeful perspective on learning for students and leave opportunities for continued self-regulation toward mastery. Providing opportunities for all students to achieve at higher levels is an essential component of standards-based grading and the primary justification for adopting the practice of using performance levels instead of grades to measure progress in achieving proficiency in standards and indicators.

Current Trends in Grading Practices

A continuum of learning. Hypothesis 1: Standards-based reporting methods that provide feedback to families and students of progress along a continuum sustain motivation toward attainment of proficiency more than traditional grades, therefore producing a median SGP and mean CPI that are higher than traditional report cards.

The subjectivity of letter grading and the cumulative assumption approach to grading (averaging) have led to discrepancies between report card grades and accountability measures. Some students who have received A’s and B’s on their report cards throughout the grades have received performance levels of ‘Warning’ or ‘Needs Improvement’ on the MCAS. This discrepancy has led some parents to file legal action alleging grade inflation. Parents assert that they have been misled into thinking that their student was achieving at levels commensurate with

their peers (Freedman, 2005). Many teachers allege that schools are merely passing students through the system without real mastery of the skills necessary to receive a diploma. These allegations are supported by teacher perception as reported in Johnson, Arumi, and Ott (2006) that 40% of teachers, in schools of students who are mainly minority, believe that students are being passed through the system without learning. Traditional grading is so subjective that an “A, for example, may mean that the student knew what the teacher expected before instruction began (product), didn't learn as well as expected but tried very hard (process), or simply made significant improvement (progress)” (Guskey, 2001, p. 21). Standards-based reporting alleviates this discrepancy in grading because students receive a grade or performance level within the standard or strand that is being taught. Measuring a student against progress toward a standard should inherently include the assumption that the student will attain that standard, or at least continue progress toward it, given the right amount of time and supports.

Guskey (2004a) reports on the evidence that standards-based report cards account for what students know and can do. Aidman, Gates, and Deterra Sims (2001) further support that parents who receive a standards-based report on their child indicate that the report card shows them what their students know and can do. The documentation of mastery of specific skills in reading and math particularly has been a focal point of standards-based report cards (American School Board Journal, 2006; Silbergliitt, Burns, Madyun & Lail, 2006). Accurate reporting of mastery is essential information for parents and students given the high-stakes accountability measures mandated by NCLB in reading, math, and science. Rick Stiggins (2005) emphasizes that educators must “use many different assessment methods to provide students, teachers, and parents with a continuing stream of evidence of student progress in mastering the knowledge and skills that underpin or lead up to state standards” (“Assessment FOR Learning”, para. 1).

The focus of standards-based grading is the assessment of student mastery at the point of assessment on designated standards within a content area. Marzano and Kendall (1996) emphasize that the key to standards-based grading is the use of the most recent scores, not a compilation of work over time. A student who struggles at the beginning of the term but masters the skill toward the end of the term can attain a performance level of 3 for 'At the grade level standard' or perhaps a 4 for 'Advanced at the grade level standard' despite the difficult start. The information conveyed to both students and families gives a more accurate picture of the student's learning progress. Lalley and Gentile (2009) liken standards-based grading to the process of earning a driver's license:

A good model for this is the driving test on which, in most states, it is necessary to pass a written test at 80% correct on rules of the road and other essential facts, as well as to demonstrate competence in authentic skills of driving. The material to be tested and the passing standards for the driving test are published for all applicants, and passing both parts of the exam is necessary to receive a license to drive. Note also that a person who scores 80, one who scores 100, and one who needs three tries to attain at least an 80 are all treated the same: All are considered sufficiently competent to receive a license that allows them to begin their careers as drivers (with some restrictions in enlightened states). Although the analogy to driving tests probably ends there, we teachers need to impress upon students as well as the general public that initial mastery is only that: the beginning. (Lalley & Gentile, 2009, p. 32)

The impetus for standards-based grading has been the urgency to identify student strengths and weaknesses in terms of content knowledge and skills. The more defined standards-

based reports of learning give much more information than the overall 'C in Math'. A more informative report that provides performance information on specific standards in terms such as 'approaching grade level benchmarks' in 'understanding place value' is essential. One overall letter grade in Mathematics could not possibly represent the student's knowledge and skills across five strands and dozens of standards. This content-driven standards-based reporting format is designed to inform students and families of their progress along a continuum of skills and builds self-regulation and internal dialogues as necessary to function in the zone of proximal development described by Vygotsky. As students identify the goals of mastery defined by the standards, they have more resources to regulate their own learning and build self-efficacy. The continuum implies an expectancy of progression in learning. The awareness of self-efficacy leads to higher levels of motivation and thereby results in enhanced growth and achievement.

Failing grades as punishment. Hypothesis 2: Grading systems that eliminate failing grades result in higher levels of motivation, self-efficacy, and growth for students who are at risk for learning, therefore producing a median SGP that is higher than systems depicting failure as a grade.

In traditional grading practices it is not unusual for teachers to assign zeros or Fs to student work that is missed, neglected, or turned in late. Doug Reeves (2004) affirms this process noting that "evidence to the contrary notwithstanding, there is an almost fanatical belief that punishment through grades will motivate students" (para. 5). If grades should represent how well a student has mastered the standards, then assigning zeros for learner behaviors or social behaviors is not conducive to best practice (Guskey, 2002). "Just two or three zeros are sufficient to cause failure for an entire semester, and just a few course failures can lead a student to drop out of high school, incurring a

lifetime of personal and social consequences” (Reeves, 2004, p. 134). There are no studies that support the use of zeros or low grades as effective punishments for students. Instead of prompting greater effort, zeros and the low grades they yield more often cause students to withdraw from learning. To protect their self-images, many regard their low mark or grade as irrelevant and meaningless. Other students may blame themselves for the low grade but often feel helpless to make improvements (Guskey, 2004b).

Traditional report cards do not promote self-efficacy or motivation for the majority of children. Poor grades do not build a student’s belief in his or her own ability to learn content; do not create a sense of self-efficacy; and will result in decreased motivation to continue striving to learn. Yet teachers continue to issue failing grades with the belief that the fear of a failing grade will somehow motivate students to improve. A system of standards-based report cards that removes the concept of failure is one positive step toward sustaining a student’s belief in his or her capacity to learn. “Grades have some value as a reward, but no value as a punishment” (Guskey, 2010, p. 10). Educational leaders struggle to direct teachers away from using grades as a method of punishment for students.

At a recent presentation to middle school teachers to introduce a new system of standards-based grading, a prevailing concern with the newer performance levels was how to give a student a zero or an F for not participating or completing homework. At a subsequent staff meeting to review the results of student performance on the Massachusetts Comprehensive Assessment System (MCAS) for middle school students several teachers were surprised by the level of growth and proficiency demonstrated by some of the students that they perceived as low achievers because they had failed their courses. These teachers have become accustomed to using grades to punish students rather than solely for reporting out on their level of academic

knowledge and skills. The use of grades as punishment ultimately feeds in a student's weak perception of self-efficacy and deteriorates his or her motivation. This can be especially true for students who are at risk for learning.

Separation of learner behaviors. Hypothesis 3 - Standards-based report cards that separate mastery of content knowledge and skills from learner behaviors lead to higher levels of growth than amalgamated grades on traditional report cards grades, therefore producing a higher median SGP and higher CPI for schools using them.

The letter grades on report cards are not always reflective of a student's mastery of content but may be more of an assessment of compliance and social skills. Students who return homework, participate in class, please the teacher, and comply with social norms generally receive good grades under traditional grading practices. Students who do not return homework, don't raise their hands in class, don't please the teacher, and don't interact well with peers will generally receive poor or failing grades under traditional grading practices.

Traditional letter grades are assigned with the intention of reflecting student progress and informing families and students, but an examination of grading practices and policies indicates inconsistencies in what those grades have come to mean. Most systems of traditional grading include a single letter grade or percentage that encompasses assessment of study skills, persistence on task, work habits, and participation that sometimes supersede academic achievement in the content (Freedman, 2005; McTighe & Bailey, 1996). In order to assign a single letter grade to students for each subject studied, teachers must combine evidence from a multitude of diverse sources into that one mark. This makes the grade "a confusing amalgamation that is impossible to interpret and rarely presents a true picture of a student's proficiency" (Brookhart, 1991; Cross and Frary 1996 in Guskey and Jung, 2006). Traditional

grading often consists of “hodgepodge grades” including elements of achievement, attitude, effort, and behavior into one amalgamated single symbol assigned to describe a student’s level of mastery of course content (Guskey, 2006, p. 671). Although stakeholders may protest that they easily recognize what a letter grade represents, there is so much disparity in grading that while a C is supposed to represent average achievement “in many districts, a C may effectively be a failing or at least below-average grade” (Zirkel, 2007, para. 19). Indeed, a local college informs its instructors that if any student is receiving below a B in a course they should be referred to the Dean, implying that a C would mean failure in mastering the content.

McMillan, Myran, and Workman (2002) collected an inventory of the factors influencing teacher’s traditional grading practices. They found a fairly large standard deviation in teacher’s indications of the amount that they factored ability, effort, and improvement into overall grades, ranging from *very little* to *quite a bit*. To accurately interpret any reporting form, parents need to know precisely what the grades mean and what is expected of students. Darling Hammond (as cited in Hearn, 2004) made the observation that the more teachers understand about what students know and how they think, the more capacity they have to reform their pedagogy, and the more opportunities they create for student success. Teachers following standards-based practices find themselves transforming their teaching as ongoing assessment reveals how students approach tasks, what helps them learn most effectively, and what strategies support their learning. When all groups understand what grades or marks mean, and how they can be used to improve student learning, the true value of a standards-based approach will be realized (Guskey & Jung, 2006, p. 3).

Stiggins and colleagues found that contrary to recommended practice teachers value student motivation and effort and set different levels of expectation on the basis of their

perception of a student's ability (Stiggins, Frisbie & Griswold, 1989 in McMillan, Myran & Workman, p. 204). A more recent trend in systems of grading is the separation of assessment of student academic progress from the evaluation of the learning behaviors such as motivation, effort, homework completion, collaboration, and cooperation with others. An important distinction between more traditional report cards and standards-based report cards is the separation of academic topics from these learner behaviors and/or life skills (Marzano, 2006). In order to effectively report academic progress in a manner that promotes learning "educators must take special care to ensure that essential cognitive or achievement goals are kept distinct from specific aspects of students' behavior – such as effort, responsibility, punctuality, participation, respect for classmates, and the like – which also may be considered important goals in the teaching and learning process" (Guskey & Bailey, 2010, Chapter 2, "Developing the Reporting Standards", para. 2).

Standards and indicators. Hypothesis 4: Standards-based report cards that clearly identify grade level power standards in math, specific to Grade 4, result in higher levels of growth and performance than report cards with one overall content area grade, therefore producing a median SGP that is higher and a CPI that is higher.

Students deserve a high quality education that assures that they will attain proficiency in an established set of knowledge and skills that will prepare them for an uncertain future. The competitiveness of the U.S. in the global economy is inherently dependent upon the assurances that students are leaving public schools with the skills necessary to think critically and creatively for the 21st century (Friedman, 2005). Bandura's theory that self-efficacy and motivation are pathways to higher achievement demands that the school community exhaust all research-based

practices to provide students with effective feedback that builds their sense of accomplishment and achievement.

Larry Ainsworth (2003) points out that expecting students to learn all of the concepts and skills embedded in a state's standards may not be an achievable goal within a school year of instructional time. Ainsworth sets the stage for identifying grade level power standards at the local level. "Learning that endures should be an essential criterion that we never lose sight of" (Chapter 1, "Are these criteria equal" para. 3). Best practices in standards-based reporting call for a moderate middle ground between the amalgamated grades of traditional report cards and the cumbersome standards-based report cards that report performance levels in every one of the dozens of indicators on state standards. Power standards developed at the local level provide a focus for instruction and learning on the essential knowledge and skills that students must acquire to be successful in successive grades and endeavors. A report card that reflects power standards provides information that is easily understood by parents and students and provides information they need in order to identify the next steps in the learning process. "Educators must develop reporting standards that are specific enough to communicate the knowledge and skills students are expected to acquire but not so detailed that they lose their utility when shared with parents" (Guskey & Bailey, 2010, Chapter 2, "Developing the Reporting Standards", para. 1). Ainsworth (2003) reports that when Robert Marzano was asked what conditions were needed to implement standards effectively, he replied "Cut the number of standards and the content within standards dramatically" (Chapter 1, "The Rationale for Power Standards", para. 2). Ainsworth recommends a prioritized set of standards and indicators that show that students can utilize higher-order thinking skills and can integrate new learning with prior knowledge. The recent introduction of the more focused Common Core State Standards reflects the national recognition

of the problem with the previous standards of many states which were often *a mile wide and an inch deep*.

Ainsworth (2003) suggests use of the following questions when determining the essential indicators that teachers must teach their students: “What knowledge and skills must I impart to my students this year so that they will enter next year’s class with confidence and a readiness for success?” (Chapter 1, “How Educators Define Essential”, para. 2); “What do your students need for success – in school (this year, next year, and so on), in life, and on your state tests?” (Chapter 1, “School, Life”, para. 2). Ainsworth cites the work of Doug Reeves in developing identification criteria for power standards. Reeves suggests three distinct criteria:

1. Endurance – Will this standard or indicator provide students with knowledge and skills that will be of value beyond a single test date? For example, proficiency in reading will endure throughout a student’s academic career and professional life.
2. Leverage – Will this provide knowledge and skills that will be of value in multiple disciplines? For example, proficiency in creating graphs, tables, and charts and the ability to draw accurate inferences from them will help students in math, science, social studies, and language arts. The ability to write an analytical and persuasive essay will similarly help students in every academic discipline.
3. Readiness for the next level of learning – Will this provide students with essential knowledge and skills that are necessary for success in the next grade or the next level of instruction? For example, fourth grade teachers are unanimous that reading comprehension and math facts recall are essential for third graders who wish to enter the fourth grade confidently and pursue fourth grade studies

successfully. Those same fourth grade teachers are not unanimous that the ability to assemble a leaf collection, identify dinosaurs, or know the state capitals are required knowledge for entry into fourth grade (Chapter 1, “Power Standards Identification”, para. 4-6).

There are differences of opinion on the number of power standards that should be included on standards-based report card. Ainsworth suggests that the power standards within one content area for a grade level should be approximately one-third of the total number. Thomas Guskey suggests 4-6 power standards within a content area. Doug Reeves’ rule of thumb is 7 to 12 power standards. These researchers agree that teachers can’t possibly assess dozens of standards on a regular basis in order to determine student mastery (Ainsworth, 2003, Guskey & Bailey, 2010, Reeves in Ainsworth, 2003). “Often all standards are considered equal when, in fact, certain standards are more important than others in terms of overall student success” (Ainsworth, 2003, “Introduction”, para. 4). Fewer reporting standards may be easier for both families and students to retain in order to focus on continuous improvements” (Guskey & Bailey, 2010, p. 43). School districts strive to determine for themselves which standards should be the focus of instruction and are essential to inform students and families of progress in learning.

It is important for educators to be mindful of the purpose of reporting grades; to inform students and families of the student’s progress toward mastery of standards. Standards-based report cards should present a picture of student’s performance level along a continuum toward mastery of a grade level standard. The report of performance should connect that standard to the essential skills the student will need to be college and career ready upon graduation. In order to meet that purpose, report cards should be presented in parent-friendly and student-friendly

language that clearly identifies achievement targets and promotes parent partnership in future learning (Stiggins, 2005, Lalley & Gentile, 2009).

Resistance: Transforming the Culture

While educational leaders approach innovation with a vision for improvement in student achievement, stakeholders sometimes approach innovation with a sense of cynicism, apathy, or even dread. New experiences are always reacted to initially in the context of some “familiar, reliable construction of reality” (Fullan, 2001, p. 21). Teachers and families bring a vast array of realities to the context of changes to report cards. The introduction of standards-based report cards puts teachers and families into a state of imbalance. Parents, teachers, students, and community members do not want to release their hold on the familiar letter grading system in spite of the evidence that letter grades are arbitrary (Marzano & Kendall, 1996). As Marzano comments:

Ultimately, to make these kinds of significant changes in report cards, a district or school must transform its culture from one in which individual teachers develop their own idiosyncratic methods of grading to a culture in which grading and reporting are uniform from teacher to teacher and subject to subject. (Marzano, 2006, p. 132)

The reporting of student performance on individual standards does make for a dramatic change in the type of report cards that are issued by a school (Marzano & Kendall, 1996). In some cases, this change has been met with such resistance that districts have reverted back to the traditional grading systems (Marzano & Kendall, 1996). While research indicates a need for a change, the reform presents a unique challenge for leaders. Engaging teachers, families, and students in a process to embrace that challenge is difficult. However, the urgency is paramount;

the educational landscape is evolving rapidly in the 21st century. Students are being educated in traditional methods that ill-prepare them to enter a society that will employ most of them in jobs that don't even exist today. *A Nation at Risk* cites "knowledge, learning, information and skilled intelligence" as the "new raw materials" of international commerce (US National Commission on Excellence in Education, 1983, para.7). Given the level of resistance to standards-based report cards for parents, teachers, students, and communities there is a need for more evidence that connects this type of grading to student performance improvements on accountability measures of those same standards.

Literature Review Conclusion

School districts can differ drastically on the reporting tool used to provide information to families. The report card can take the form of age- or grade-appropriate expectations expressed in benchmarks, complete standards, open-ended statements, or abbreviated focal points (Salend, 2005). There may be several performance levels, one overall grade or average, or a multitude of individual standards. Report card formats range from the traditional archaic form of the past 40 years, which depicts the familiar letter grade in each of the content areas of Math, Reading, Writing, Science, and History, to the multi-page standards-based reports that indicate student performance on multiple indicators within each strand and content area. McTighe and Thomas (2003) report that their analysis of grading practices revealed a process that varied so widely that report cards were not considered to be reliable measures of achievement and progress. School district leaders are struggling with the development of one easily understood tool that gives a meaningful report on learning to parents, students, teachers, and the community.

Standards-based educational practices for the twenty-first century must be embraced to promote an informed student body capable of accomplishing future-directed goals (Bagley,

1939; Bandura, 2001; Friedman, 2005; Hirsch, 2007; Stiggins, 1997; Stiggins, 2005; Stiggins & Chappuis, 2008; Vygotsky in Grusec, 1992). As districts ask parents, teachers, students, and the community to support a change in a reporting tool that is so engrained in society, it is essential to have data that indicates that the tool will impact student learning and student performance on accountability measures. Impending realignment to the newly revised *Massachusetts Curriculum Frameworks Incorporating the Common Core State Standards* for both Mathematics, and English Language Arts and Literacy, will mean that schools will revisit standards-based report cards for revisions that reflect alignment to the new curriculum and the new national assessments. The use of standards-based report cards from the earliest grades to promote self-efficacy and motivation in children will build children's perceptions of themselves as capable learners and inspire hopefulness for continuing along a procession toward higher levels of achievement. All students must come to believe that they can succeed at learning if they try. They must be continuously informed of the steps that they have taken toward academic proficiency in order to build "confidence, optimism, and persistence --for all, not just for some" (Stiggins, 2005, "A New Mission," para. 5).

Chapter 3: Research Design

This is a causal-comparative study of the effect of standards-based report cards on student growth percentiles and composite performance index scores on the Massachusetts Comprehensive Assessment System. "In causal-comparative research, investigators attempt to determine the cause or consequences of differences that already exist between or among groups of individuals" (Fraenkel & Wallen, 2009, p. 363). Both the effect and the cause have already occurred in causal-comparative studies. These studies are also referred to as *ex post facto* research (Fraenkel & Wallen, 2009). This study researches the effect of student feedback in the

form of report card type (a difference that already exists) on student performance (consequence of group membership that already occurred); it is therefore a suitable study for a causal-comparative research design.

There is one categorical independent variable - report card type. A product rating scale developed for the study is used to delineate the categorical variable of report card type according to quantitative variables in a criterion-group design. The study is conducted at the organizational level and examines the effects of standards-based report cards on growth and performance of Grade 4 students in mathematics in southeastern Massachusetts elementary schools. The study controls for several covariates among the categorical independent variables: percentage of students in the low income, special education and limited English proficient categories; student-teacher ratio; and per pupil expenditure.

The hypothesis that drives this research study is: Fourth grade results for students in the schools with standards-based report cards will have higher percentages of growth and higher indices of performance as measured by the MCAS Mathematics Test than schools with non-standards-based report cards.

Variables

Independent variable. The independent categorical variable for this study is the type of report card used by each elementary school. For the purpose of this study three groups of report card type were established using the following operational definition of standards-based report card as guidance: an alternate way of reporting student progress to parents which involves assessing student proficiency in alignment with the state/local standards and benchmarks. As the questionnaires and report card samples were analyzed, a two-step process for determining membership of each school in the three categories of the independent variable was developed.

Each report card was rated according to an ordinal rating scale first and then categorized using a grouping rubric as standards-based, mixed with general reporting standards, or non-standards-based. This process is described in detail under the measures section of this chapter.

Dependent variables. This study utilized Student Growth Percentiles (SGP) as the measure of student progress and the Composite Performance Index (CPI) as the measure of performance. These variables were described fully in Chapter 1.

Student Growth Percentile. The Massachusetts Department of Elementary and Secondary Education (2009) describes the SGP:

“A student growth percentile is a measure of student progress that compares changes in a student’s MCAS scores to changes in MCAS scores of other students with similar achievement profiles. The model establishes cohorts of students with “similar performance profiles” by identifying all students with the same (or very similar) MCAS scores in prior years; all MCAS data for a student since 2006 are used (where available) to establish academic peers” (p.1).

Composite Performance Index. The Composite Performance Index (CPI) is defined by the DESE as a measure of the extent to which students are progressing toward proficiency in English Language Arts (ELA) and Mathematics, respectively. The CPI is a 100-point index that combines the scores of students who take standard MCAS tests (the Proficiency Index) with the scores of those who take the MCAS-Alternate Assessment (MCAS-Alt).

The Student Growth Percentile (SGP) and Composite Performance Index (CPI) for Grade 4 Mathematics for each elementary school were used as the measures of growth and performance respectively for comparison as the dependent variables across the categorical independent variable as shown in Table 1.

Table 1

Causal-Comparative Criterion Group Design for Report Card Type

(a)	Group	Independent Variable	Dependent Variable
	I	Standards-Based	SGP
	II	Non-standards-based	SGP
	III	Mixed	SGP
(b)	Group	Independent Variable	Dependent Variable
	I	Standards-Based	CPI
	II	Non-standards-based	CPI
	III	Mixed	CPI

Covariates. The data collected from elementary schools in the southeast region of Massachusetts was controlled for the following moderator variables: percentage of students categorized as low income; percentage of the population designated as receiving special education services; percentage of population that is Limited English Proficient; class size as reported as a general statistic by the DESE; and per pupil expenditure (as a relative representation of level of funding). The statistical data for each covariate was gathered from the publicly available data on the DESE website under School Profiles.

Threats to Internal Validity

The major threat to internal validity of a causal-comparative design is the possibility of a subject characteristics threat (Fraenkel & Wallen, 2009). The most serious limitation is the lack of control over threats to internal validity. Because the manipulation of the independent variable has already occurred, many controls cannot be applied. Relationships can be identified but causation cannot be fully established. This threat can be reduced by matching subjects from the comparison groups on the extraneous variable that may be influencing the results, finding or

creating homogeneous subgroups, or statistical matching. Additional threats include “loss of subjects, location, instrumentation, and sometimes history and maturation” (Fraenkel & Wallen, 2009, p. 368).

To control for a subject characteristics threat, this study created three groups of relative homogeneity. The members of the three groups of report cards for this study each include the following characteristics: participants are from southeastern Massachusetts and MCAS results are reported for Grade 4 Mathematics MCAS in 2010. In order to reduce the loss of subjects threat, a gift card was included in the mailing to all elementary principals along with the questionnaire to solicit participation in the study. An internal threat to the validity of the study that was considered was the reduced number of schools in southeastern Massachusetts that continue to use traditional report cards. The actual return for non-standards-based report cards ($N = 30$) is considered a valid number for comparison to the number of standards-based report cards ($N = 52$) and the number of mixed report cards ($N = 20$).

This study includes several covariates as an additional method to control for internal validity. The impact of socio-economic status, special education needs, limited English proficiency, student-teacher ratio, and per pupil expenditure is evident in the results of high-stakes testing. Each of these factors is included in the MANCOVA as a covariate to reduce the likelihood that the causal relationship of each factor may be influencing the results for the independent variable.

Research Questions

The primary research question for this research study is: What effect do standards-based report cards have on the growth percentiles and performance indices of fourth grade students on the MCAS Mathematics Test? Based on the theoretical framework of this study and an initial

review of the literature related to the independent variable, four research questions and hypotheses were formulated to focus this study:

- Research Question 1: Do standards-based report cards that report student performance on a continuum of progress in mathematics rather than a cumulative grade average result in higher levels of growth and performance on the MCAS? Hypothesis 1: Standards-based reporting methods that provide feedback to families and students of progress along a continuum sustain motivation toward attainment of proficiency more than traditional grades, therefore producing a higher median SGP and CPI.
- Research Question 2: Do standards-based report cards that remove indications of failure such as Fs and zeros in mathematics result in higher levels of growth on the MCAS for students at risk for learning? Hypothesis 2: Grading systems that eliminate failing grades result in higher levels of motivation, self-efficacy, and growth for students who are at risk, therefore producing a higher median SGP for schools using them.
- Research Question 3: Do standards-based report cards that separate academic progress in mathematics from social and learner behaviors result in higher levels of growth and performance on the MCAS? Hypothesis 3: Standards-based report cards that separate mastery of content knowledge and skills from learner behaviors lead to higher levels of growth and proficiency than amalgamated grades on traditional report cards, therefore producing a higher median SGP and higher CPI for schools using them.
- Research Question 4: Do standards-based report cards that grade in power standards result in higher levels of growth and performance than report cards with a cumulative grade or standards-based report cards that do not provide power standards? Hypothesis 4: Standards-based report cards that clearly identify power standards in math, specific to

Grade 4, result in higher levels of growth and performance than report cards with one overall content area grade, therefore producing a higher median SGP and a higher CPI.

Methodology

A causal-comparative research methodology was utilized to examine the effect of standards-based report cards on student growth percentiles and composite performance indices on the Massachusetts Comprehensive Assessment System. The basic causal-comparative design, sometimes called a criterion group design, involves selecting two or more groups that differ on a variable of interest and comparing them on variables. No manipulation is involved. The groups differ in one of two ways: One group either possesses a characteristic (often called a criterion) that the other does not, or the groups differ on known characteristics (Fraenkel & Wallen, 2009). This study involved one categorical independent variable on which the groups differ – report card type.

The study was conducted at the organizational level and examined the effects of standards-based report cards on growth and performance of students in Grade 4 in elementary schools in southeastern Massachusetts on the 2010 Mathematics MCAS. The study was conducted at the organizational level in order to protect the anonymity of students who would have taken the Grade 4 MCAS. MCAS data is publicly available on the DESE website at the school and district level and was readily available to this research study. School level results for Grade 4 are reported as both the SGP and the CPI. A brief questionnaire was constructed to facilitate a timely response from elementary principals and gather report cards from each school. Since this study sought to compare the results for schools using different types of report cards rather than individual student results, working at the organizational level provided the most accessible and reliable results.

Participants and population sample. The target population of interest in this study is elementary students across the state of Massachusetts. The accessible population for the study is the 214 public elementary schools educating students in Grade 4 in the southeastern region of Massachusetts (see Appendix B). Questionnaires were sent to 214 principals of elementary schools. The return rate for the questionnaires was 51% which resulted in 109 (N = 109) questionnaires returned. Of the returned questionnaires 103 (N = 103) were viable as participants for this study. This nonrandom purposive sample population of 103 elementary schools (N = 103) represents 48% of the 214 elementary schools that were targeted for the study in the southeastern region of Massachusetts and 9% of the 1138 elementary schools in the state.

Measures. The coding of the independent variable in this study utilizes a criterion group design. A product rating scale developed for the study was used to sort report cards according to four quantitative criteria. The results of the rating scale were then applied to a grouping rubric developed for this study to establish assignment of each school to one of the three categories of report card types: standards-based, non-standards-based, and mixed with general reporting standards.

Rating Scale. The results from the questionnaires issued for this study was disaggregated in two steps. Initially two groups were formed: a group representing schools using standards-based report cards and a group of schools using report cards that were not standards-based. The rating scale was then applied to each report card that accompanied the questionnaire. The Standards-Based Report Card Product Rating Scale developed for the document analysis in this study is shown in Table 2.

Table 2

Standards-Based Report Card Product Rating Scale

	Rating			
	1	2	3	4
Performance Levels	1 overall grade or level, or more than 5 levels are used, to indicate progress; Language is ambiguous or unclear on levels	3 levels: Progress levels are limited to only 1 level below proficiency; Language requires more definition or uses comparative language such as below average or superior	4 or 5 levels: Levels indicate progress toward proficiency, and above proficiency; Parent friendly language	4 levels: 2 levels to indicate progress toward proficiency, 1 for proficiency and 1 for advanced or excels; Parent friendly language
Failing Grades	Displays an F, zero or other failing designation	Displays indications of no growth or minimal growth in terms that may deter motivation	Language depicts a progression of learning that supports motivation	All language depicts a progression of learning that promotes high achievement as a possibility for all students
Separation of Learner and Social Behaviors	One amalgamated grade/level for student Learner and Social Behaviors	Learner and Social Behaviors are reported as comments on conduct and effort	Multiple Learner and Social Behaviors are reported separately from academics	Learner and Social Behaviors are reported within each content level separately from academics
Standards and Indicators	One content area reported for mathematics	Mathematics reported as strands only, or as more than 12 indicators	Mathematics reported as 4 - 12 power standards that target critical focus areas for Grade 4	Mathematics reported as 7-12 power standards that feature action words and understandable indicators for students and families

Note. Criteria development was based on the research of Thomas Guskey and Jane Bailey (2010) in *Developing Standards-Based Report Cards* and Douglas Reeves and Larry Ainsworth in Ainsworth's 2003 text *Power Standards, Identifying the Standards that Matter Most*.

This rating scale is a significant limitation of the study as it has not been utilized prior to this study. The researcher developed the rating scale for the purpose of setting criteria for

sorting the report cards used in the study into independent variable groups. The study was designed in anticipation of a clear delineation of elementary schools using standards-based report cards and those using traditional report cards based on responses to the questionnaire. However, the rating of the report cards indicated that there were a number of report cards that did not fit clearly into either group. The rating scale then became a quantitative measure used to develop a grouping rubric.

Grouping rubric. A grouping rubric was developed using the product rating scale to facilitate placing each report card into one of four types for the purpose of analysis. As the report cards were rated and catalogued it became evident that adjustments were needed in two of the criteria. The second criterion, Separation of Learner and Social Behaviors, was a common practice on report cards regardless of type. The researcher eliminated it as factor for rating because the separation of learner and social behaviors appeared on the majority of both standards-based and non-standards-based report cards. The rating for the standards and indicators criteria was structured in a way that it presented a problem in discerning the data. Report cards with more than 12 specific standards for Grade 4 received the same score (2) on the rating scale as report cards that only reported on 3 or 4 general topics. The rubric established a criterion to differentiate these features. As a result a third group was established that consisted of report cards that met the criteria for standards-based reporting in three of the criteria but did not report progress on grade-specific standards. Those report cards with grade-specific standards were included in the standards-based group. The rubric shown in Table 3 distinguished between report cards that had all of the features of standards-based report cards, including the use of grade-specific standards, and those that had all of the features of standards-based report cards but used general non-specific indicators of performance. A fourth category was

established for report cards that did have specific indicators for Grade 4 mathematics but did not meet the other criteria for standards-based as established by this study. Once all of the report cards were catalogued by type there was only one report card that belonged to this fourth category so it was eliminated for comparison in the study.

Table 3

Standards-Based Report Card Grouping Rubric

	Group 1	Group 2	Group 3	Group 4[^]
	Standards-based	Non-standards-based	Mixed with general reporting standards	Mixed with grade-specific indicators
Performance levels	Score of 2 or better with required scores in criterion 2 and 4	Score of 1 or 2	Score of 2 or better with additional scores in criteria 2 and 4	Score of 1 or 2
Depiction of failure	Score of 2 or better	Score of 1	Score of 2 or better	Score of 1
Separation of learner and social behaviors	Not applied	Not applied	Not applied	Not applied
Grade-specific standards and indicators	Score of 2 or better and depicts grading standards or indicators specific to Grade 4	Score of 1 or 2 with standards or indicators that are not specific to Grade 4	Score of 1 or 2 with standards or indicators that are not specific to Grade 4	Score of 2 or better with grading standards or indicators specific to Grade 4

[^] The group depicted as ‘mixed with grade-specific standards’ was eliminated from analysis due to the fact that only 1 school’s report card qualified as a member of this group.

The three report card types included for this study’s analysis were: Group 1 – Standards-Based, Group 2 - Non-Standards-Based, and Group 3 – Mixed with General Reporting Standards.

The report cards reviewed as part of this study displayed a varied lexicon to depict progression in learning which is described in the document review section of Chapter 4. Surprisingly some report cards retained the traditional ABC grading systems but redefined the terminology. Several of the report cards that used a traditional grading system of ABC were categorized as standards-based because the schools had modified the performance meaning of the letter grades. In all cases where the schools using a traditional ABC were included in the standards-based report cards group, the letter F for failing had been eliminated from the performance level report.

A distinguishing factor for standards-based practices is the removal of depictions of failure from report cards. The product rating scale utilized to assign a quantitative representation of the criteria for standards-based report cards depicts four ratings for the report card language regarding failure. The removal of failure as a determination of student progress on report cards is essential to instill a sense of hopefulness and a vision of ultimate achievement. Designating a child as failing can have a devastating impact on their future success. Albert Bandura highlights the necessity of informing struggling students of progress, not failure. Motivation to achieve goals and expect positive outcomes are centrally influenced by efficacy beliefs (Bandura, 2001). Research has established that enhancement of perceived self-efficacy could improve children's cognitive skill development and their intrinsic interest in academic subjects (Grusec, 1992). As report cards were rated and then categorized into one of the three independent variable groups, the researcher considered only those report cards receiving a score of 2 or better, indicating no use of language that depicted failure, to be true representations of standards-based grading practices.

A report card that reflects power standards provides information that is easily understood by parents and students and provides information they need in order to identify the next steps in the learning process. Ainsworth (2003) recommends a prioritized set of standards and indicators that show that students can utilize higher-order thinking skills and can integrate new learning with prior knowledge. Power standards developed at the local level provide a focus for instruction and learning on the essential knowledge and skills that students must acquire to be successful in successive grades and endeavors. Few schools fit the definition of power standards established in the literature. Schools that received a 2 on the criterion for standards and indicators and had grade-specific standards for math were included in the standards-based group. All schools that received a 3 or 4 on the rating scale for standards and indicators were assigned to the standards-based group.

For the purpose of this study inclusion in the standards-based report cards group required that a report card had standards specific to Grade 4 mathematics curriculum frameworks, eliminated grading language that indicates failure, separated grading for learner behavior and social behavior from academic achievement, and included at least three levels of performance reporting.

Procedures

Data collection. This research study employed an unobtrusive collection of data throughout the collection of data on all variables. Data was collected from questionnaires and report cards returned from elementary schools, the Standards-Based Report Card Product Rating Scale, the demographic profiles of each school on the DESE website, and the publicly available assessment and accountability results of the MCAS on the DESE website.

Questionnaires. The questionnaire was mailed to all school principals in the southeast region of Massachusetts during the first week of June 2011. The mailing included a stamped self-addressed envelope for returning the requested report card copies, a copy of the questionnaire, and a small gift card as a thank-you for completing the questionnaire. The questionnaire included an option to submit a bill for the cost of copying the report cards as well as an option to receive the results of the study when it is completed. The timing of the mailing was designed to avoid the demanding weeks of the administration of the MCAS in elementary schools in order to promote a higher response rate.

This study utilized a brief unobtrusive questionnaire to participants to gather data about the use of a standards-based report card and to collect copies of report cards. The questionnaire and accompanying cover letter are included as Appendix A. The questionnaires were sent to school principals to inquire whether or not their elementary school used a standards-based report card. If the responding school was not using a standards-based report card, they were instructed to indicate ‘no’ on the questionnaire and return it along with a copy of their traditional report card. For those responding schools that were using a standards-based report card, there were additional questions inquiring as to when it was implemented and in what grades. Participants indicating that they did have a standards-based report card were also asked to send a copy of their school’s report card to the study along with the completed questionnaire.

Report card coding. As completed questionnaires and samples of report cards were received in late June of 2011, the report cards were rated using the Standards-Based Report Card Product Rating Scale described in detail earlier in this chapter and then coded using the Grouping Rubric. A cover sheet was attached to each report card that captured all of the scores on the rating scale as well as the code for report card type. The codes utilized were 1 for

standards-based, 2 for non-standards-based and 3 for mixed. For each school the rating on each of the criteria and the report card type code were entered into an Excel database along with the principal's responses to the questionnaire and indications of the principals who wanted to receive copies of the study. If principals indicated that their school had a standards-based report card they were coded as a 1; if they indicated that they did not, they were coded as a 2. As stated earlier the researcher made the decision to rate each report card and code it as one of three types of report cards regardless of the participant's response to the first question. However, the response of the principals to the first question regarding type of report card was maintained as a data field because it provided an interesting insight into the understanding of what constitutes a standards-based report card. The data on the additional questions on the questionnaire regarding years of implementation and grades of implementation were entered into data fields as well but were not utilized in the study because the information was missing for many of the respondents.

Growth and performance. As each questionnaire and accompanying report card was catalogued, the school's performance results on the criterion-referenced MCAS were accessed through the public database on the DESE website. To access data, the researcher selected school/district profiles, selected the district name, and then selected the specific school name. Grade level data on the student growth percentile and composite performance on the Mathematics MCAS was obtained by accessing the assessment tab.

Covariates. In order to control for covariates, the percentage of students who qualify as low income, the percentage of students receiving special education, the percentage of students with limited English proficiency, the teacher-student ratio, and the per pupil expenditures were collected for each school that responded to the survey. Data collection utilized the school and district profiles data on the website of the Massachusetts Department of Elementary and

Secondary Education. To access the data, the researcher once again selected school/district profiles, selected district name, and then selected the specific school name. To obtain the demographic data for each school, the student tab and the selected populations tab were used to collect data on percentage of low income, percentage of special education, and percentage of limited English proficient. For each school the teacher tab was used to access the student-teacher ratio. The finances tab of each district level profile was used to access the per pupil expenditure.

Data fields. The following elements represent the compilation of data collected and entered into data fields in Excel for purposes of analysis: school name; district; request for copy of study; question 1 response (traditional report card or standards-based report card); years of implementation; length of implementation; percentage of low income; percentage of special education; percentage of limited English proficient; per pupil expenditure; student-teacher ratio; Grade 4 math SGP for 2010; Grade 4 math CPI for 2010; the product rating on each of the 4 features of report cards: performance levels, designation of failure, separation of learner and social behaviors from academic grades, use of grade-specific standards and indicators for mathematics; and the type of report card coded as the independent variable.

Data Analysis: Multivariate Analysis of Covariance

A statistical analysis of the student growth percentiles (SGP) and composite performance index (CPI) on the Grade 4 MCAS for Mathematics for the responding schools was conducted to determine a causal-comparative relationship with the use of standards-based report cards. The data collected was analyzed for statistical significance through a general linear model using the Statistical Package for the Social Sciences (SPSS). The fixed factor entered for analysis was standards-based report cards. A simple plan comparison was utilized in a multivariate general linear model to examine the statistical difference in growth and performance of students.

A multivariate analysis of covariance (MANCOVA) is used when two or more dependent variables are incorporated thereby allowing for a more powerful test of differences among means (Fraenkel & Wallen, 2009). A Factorial MANCOVA was employed to determine whether there are statistically reliable mean differences among schools using standards-based report cards and schools using non-standards-based or mixed report cards after adjusting the dependent measure on several covariates. The covariates that were controlled for the study were: percentage of students in the low income category; percentage of students in special education; percentage of students with limited English proficiency; teacher-student ratio; and per pupil expenditure. Statistically significant differences among the student growth percentiles and composite performance indices in relation to the independent variable of standards-based report cards were calculated and reported as the value of Wilks' lambda (Brace, Kemp & Snelgar, 2009). The results of the study are reported as statements of probability relating the effect of standards-based report cards on student growth and performance on the MCAS.

Protection of Human Subjects

This project abided by the policy of Northeastern University for the protection of human subjects. The project used a non-obtrusive questionnaire and publicly accessible data available from the Massachusetts Department of Elementary and Secondary Education for the analysis. The brief questionnaire that was sent to all elementary schools in southeastern Massachusetts requests general information that would also be publicly available at each school. The questionnaire includes a statement of assurance of anonymity of participants as reported in the study. No individual school is named in the study.

The growth data collected from the Massachusetts Department of Elementary and Secondary Education website is reported as aggregate data for the students in Grade 4 in each

participating school. The MA DESE does not report results for subgroups of less than 40 students to protect confidentiality of data. There is no identification of student-specific data or school-specific data at any point in the study. No individual student data was accessed or reported for the purpose of this study.

The questionnaire used in this study has been approved by the Northeastern University Internal Review Board. The researcher has a certificate of completion for training on protecting human research participants in research. The researcher will not use, discuss, or benefit from the data gathered at work, with colleagues, or any other public venue without the informed consent of the individuals providing such data.

All participants were given the option of providing contact information on the questionnaire in order to receive the results of this study once it is completed.

Timeline

The questionnaire was mailed to the principal of each elementary school in the southeastern region of Massachusetts during the week of June 2, 2011. Responses were immediately catalogued as they arrived. Report cards were rated during the months of June and July. All report cards were re-rated by the researcher in August and checked for consistency. The report cards were sorted into three groups of the independent variable during August. Data fields were entered into SPSS in August 2011 and reviewed for accuracy. Data analysis was conducted in August 2011. Conclusions were derived in August and September 2011.

Chapter 4: Report of Research Findings

This study examines the causal-comparative relationship between schools using standards-based report cards and levels of student performance and growth on the criterion-referenced Massachusetts Comprehensive Assessment System (MCAS). The study reviews the presence of components of standards-based report cards in samples collected and makes inferences regarding the impact of standards-based report cards on improving outcomes for students by gathering data on Student Growth Percentile and the Composite Performance Index on the Grade 4 MCAS for Mathematics.

Independent Variable

Questionnaires. The data gathered from the questionnaire included the answer to whether or not the school used a school-based report card. There were 32 schools that responded negatively on this first question. However, an initial review of the report cards that accompanied the questionnaires revealed that those schools who indicated that they did not have a standards-based report card in many cases had a report card that had standards-based features. Schools did not consistently provide years of implementation or grades of implementation for their report cards per the questionnaire directions. Principals who did not consider their report cards to be standards-based also omitted this information per the questionnaire directions; therefore, data on the years of implementation was not considered valid for use as a covariate for the study as originally proposed. The study focused on the results for 2010 only. The researcher made the decision to apply the rating scale to each report card returned regardless of the participant's response on the first question based on this initial review.

Report card coding. The product rating scale developed for the document analysis in this study is included as Appendix C and described in detail in Chapter 3. The researcher

developed the rating scale for the purpose of setting criteria for sorting the report cards used in the study into groups. Each report card was evaluated on the four criteria and received a rating of 1 to 4. As described in Chapter 3 the results from the rating scale were then applied to the Standards-Based Report Card Grouping Rubric that is included as Appendix D to code each report card for sorting into one of the three groups of the independent variable. The three groups created for the study were: Group 1 – standards-based; Group 2 – non-standards-based; Group 3 – mixed.

- Report cards in Group 1 – Standards-Based had the following features: These report cards used at least 3 performance levels to delineate progress. They removed grades that depicted failure and included language depicting a progression of learning that supports motivation. They had indicators of proficiency in mathematics that were specific to Grade 4.
- Report cards in Group 2 – Non-Standards-Based had these features: These report cards had either 1 overall grade or exceeded 5 levels to indicate progress or included language that was ambiguous or unclear. They displayed an F, zero, or other failing designation or displayed indications of no growth or minimal growth in terms that may deter motivation. They reported with general standards or indicators in mathematics that were not specific to Grade 4.
- Report cards in Group 3 – Mixed with General Reporting Standards had these features: Similar to the standards-based group these report cards used at least 3 performance levels to delineate progress. They removed grades that depicted failure and included language depicting a progression of learning that supports motivation. However, these report cards differed from standards-based report cards in that they did not have indicators of

proficiency in mathematics that were specific to Grade 4 and instead had general standards or indicators.

The number of standards or indicators was not considered as a factor in assigning group membership. Researchers specify a limited number of indicators in parent-friendly language as qualifying features of power standards (Ainsworth, 2003, Guskey & Bailey, 2010). There were only 9 schools that had indicators that fit the definition of power standards in the research. This study included report cards in the standards-based category if they had indicators of proficiency for math that were specific to Grade 4. The number of indicators in this grouping ranged from 7 to 26. The indicators were not reviewed for parent-friendly language. As a result of coding according to the grouping rubric 12 of the 32 report cards from schools indicating on the questionnaire that they did not use standards-based report cards were categorized either as standards-based or mixed.

Continuum of progress. The review of the report cards conducted as part of the coding for this study supports research findings that there continue to be great inconsistencies in what the familiar letter grades represent for student progress. The conclusion of McTighe and Thomas that “analysis of grading practices varied so widely that report cards were not reliable measures of achievement and progress” (2003, p. 3) could easily be applied to standards-based report cards as well as traditional report cards. Vast inconsistencies now extend to the performance levels used by many schools as part of the system of grading that is present in standards-based report cards (see Schworm, 2006). Thomas Guskey has been expounding on the need for report cards that clearly communicate progress and achievement to parents in a manner that they can readily understand.

This review of report cards found a variety of performance level systems and vast differences in what the levels mean from one school to the next. For example, four schools in the same region of the state define performance that receives a letter C as a grade in four distinctly different ways:

School 1: C indicates needs improvement

School 2: C indicates average progress

School 3: C represents a 70-79% average

School 4: C indicates “The student has demonstrated most of the required knowledge and skills.”

One of these four schools was included in the group that indicated on the questionnaire it did not have a standards-based report card. However, a subsequent review of the attached report card using the product rating scale revealed that this school had all the features of a standards-based report card. The only difference was the use of letter grades instead of performance levels. This particular school used the following definition of its grades:

A- to A+ The student has demonstrated the required knowledge and skills. Achievement meets district grade level standards with distinction.

B- to B+ The student has demonstrated the required knowledge and skills.

C- to C+ The student has demonstrated most of the required knowledge and skills.

D- to D+ The student has demonstrated some of the required knowledge and skills in limited ways. Achievement falls below district grade level standard.

The school did not use the letter F or any other term to define failure. This school’s grading system is in keeping with research on standards-based grading because it provides more descriptive depictions of progress and achievement while also eliminating failing grades. The

fact that this school did not have a performance level system that used a different letter or number representation rather than the traditional ABC may have led this school leader and others to determine that his or her report card was not standards-based because of that one feature. However it is in accordance with Guskey and Bailey's recommendation that standards-based report cards include 4 to 5 performance levels with "2 distinct lower levels of performance" to provide scaffolding opportunities for those students who have not yet reached proficiency and one level higher than proficiency to "recognize those students who display truly exceptional accomplishment or skill with regard to the standard" (2010, p. 43).

This study's review of report cards revealed that some schools indicate that a student has to 'exceed the grade level standards' in order to receive a grade that reflects performance that is better than proficiency in the standards. Guskey recommends a common set of descriptors that matches performance levels 1, 2, 3, and 4 with the achievement labels *beginning*, *progressing*, *proficient*, and *exceptional* and includes performance criteria for an *exceptional* or *advanced* level of achievement or performance. Some schools are providing a performance level that depicts 'meeting the grade level expectation/standard at an exceptional level' or 'advanced at the grade level standard.' Such language suggests that advanced performance is within the grasp of students who are excelling within the standards for their grade without having to demonstrate some acquisition of above grade level standards. The use of descriptors such as *exceptional* or *advanced* are preferable to *exceeds* or *extending* to distinguish students who are attaining the highest levels of performance. Terms such as *exceeds* or *extending* leave students, families, and teachers questioning what they need to do to exceed or extend. "Descriptors should be clear, concise, and directly interpretable" (Guskey, 2001, p. 25). The schools meeting this criterion are following the best practices supported in research.

Marzano and Kendall (1996) remark that the use of letter grades is firmly entrenched in today's society. It seems that the continued use of ABCs or percentages for reporting student progress and achievement has precluded some administrators from viewing their report card as standards-based. What is indicated by the study is that some administrators believe that simply switching to a performance level system that does not use letter grades is enough to qualify their report card as standards-based. This adjustment is only one component of standards-based grading and arguably not necessary. This study proposes that the labels affixed to performance levels should not be the prominent factor in determining whether or not a report card is standards-based. Several of the report cards returned continued to use ABCs for grading but by redefining the representation of each grade they still qualified as standards-based report cards. These report cards are examples of what may be a more effective transition to standards-based practices by removing determinations of failing, separating learner and social behaviors, and reporting on grade-specific standards of performance while also maintaining a system of grading that is familiar to students, families, and teachers. If the "purpose of the report card is to communicate to parents the achievement status of students, then parents must understand the information on the report card and know how to use it." (Guskey, 2001, p. 23). The document review conducted for this study provides further evidence of the lack of clarity the meaning of grades.

Removal of failing grades. This study's review of report cards indicates that 54% (56 out of 104) schools had taken steps to remove designations of failure from their report cards. Examples of grading systems that removed failing grades are depicted in Table 4.

Table 4

Report Card Performance Levels that Remove Failing Grades

School 1		School 2	
E	Exceeds the Year End Standards	A	90% - 100%
M	Meets the Year End Standards	B	80% -89%
P	Progressing Towards the Year End Standards	C	70%-79%
B	Beginning to Develop the Year End Standards	N	69% or below
School 3		School 4	
4	Exceeding the grade level standard and consistently producing the highest quality of work	4	Demonstrates proficiency
3	Meeting the grade level standard and producing quality work	3	Shows steady growth
2	Progressing toward the standard and producing required grade level work with assistance	2	Progressing with support
1	Not meeting the standard and not yet able to produce required grade level work	1	Behavior/Skill not yet evident

An additional feature that was noted during the document review is the notation of when students had received modifications to the curriculum on which the grade was reported. Modification of curriculum was depicted on 25% of the report cards. This may be another factor that impacts motivation for the population of students at risk for learning. Students from families qualifying as having limited English proficiency and those receiving special education may find themselves perpetually in the lowest level of performance without this notation.

The document review component of this study supports research findings that schools need to move from traditional grading systems which ‘fail’ students starting in the early grades to a more motivating system of grading that places students along a continuum toward mastery of knowledge and skills. Beginning in elementary school, grades on report cards contribute to

students' views of self and their perception of their ability to succeed in their environment. Students experience grades as a reward or a punishment that shape their perceptions and behaviors. In order to become more confident, proficient learners, students need grading procedures that explicitly provide opportunities for improvement in order to strengthen images of self-efficacy and motivation (Strobino, et. al., 2002). This document review revealed that 54% of schools had removed designations of failure from their report cards.

Separating learner and social behaviors. The review of report cards included in this study found that all three types of report cards separated content grades from information on students' readiness for learning, interactions with peers and staff, and acquisition of skills for collaboration and communication. All but 8 schools received a 3 on the product rating scale for this third criterion indicating that 92% of participating schools were separating learner and social behaviors from content assessment. The rating scale did not distinguish between the learner and social behaviors reported or the method in which they were evaluated. This criterion was eliminated as a determining factor in categorizing the report cards. It appears that separating learner and social behaviors from content assessment has become a widely accepted practice for reporting progress to families regardless of report card type.

Standards and indicators. The grouping rubric developed for this study was used to make a distinction between those schools that reported on specific Grade 4 standards in mathematics and those that reported on a few general strands or standards that spanned grade levels. This study strays from the research in that all schools that reported on grade-specific standards were included in the standards-based group regardless of how many indicators for mathematics were depicted on their report card. The study did not develop further criteria to evaluate if the indicators were in parent-friendly language as recommended in the literature.

Some report cards specified Grade 4 on the letterhead or cover but a review using the product rating scale showed the indicators were not specific for Grade 4 proficiency in mathematics. These report cards with general indicators that spanned grades were entered into the mixed group.

The number of indicators reported out on the report cards in the standards-based group collected by this study varied from 7 to 26. An example of some of the specific Grade 4 report card indicators of student learning is depicted in Table 5.

Table 5

Example of Grade-Specific Standards from a Grade 4 Report Card

Strand	Indicators
Number Sense and Operations	Mastery of multiplication through 10 x 10 and related division facts
	Solve multiplication and division number stories
	Demonstrate an understanding of and the ability to use the conventional algorithm for division of up to a three-digit whole number with a single-digit divisor (with or without remainders)
	Select and use a variety of strategies to estimate quantities, measures, and the results of whole number computations up to three-digit whole numbers and amounts of money up to \$1,000
	Demonstrates the ability to explain mathematical thinking both orally and in writing
	Round whole numbers through 100,000 to the nearest 10, 100, 1,000, 10,000 and 100,000

The range of general indicators reported out on the report cards in the mixed group was from 3 to 8. These report cards included indicators such as understands concepts, solves problems using skills and facts, and works accurately with numbers. The report cards in the non-

standards-based group had 1 to 3 indicators that were highly general such as mathematics, computation, and problem solving. This language was not considered to be specific enough to qualify for membership in the standards-based or the mixed group. The literature highlights the primary purpose of report cards, which is to communicate with parents. Clarity and consistency are essential if the report card is to have meaning to those who receive it. A report card that reflects power standards provides information that is easily understood by parents and students and provides information they need to identify the next steps in the learning process. “Educators must develop reporting standards that are specific enough to communicate the knowledge and skills students are expected to acquire but not so detailed that they lose their utility when shared with parents” (Guskey & Bailey, 2010, Chapter 2, Section 2, para. 1). The document review did not attempt to differentiate whether or not the indicators were in parent-friendly language as Ainsworth has indicated is a qualifier of power standards (2003).

Independent variable statistics. The coding of the report cards using the Standards-Based Report Card Product Rating Scale resulted in the assignment of each report to 1 of 4 initial groups. The summary of the cases included in the analysis for this study is represented in Table 6.

Table 6

Descriptive Statistics for the Independent Variables

		Valid		<u>Cases</u> Missing		Total	
	<u>Report Card Category</u>	N	Percent	N	Percent	N	Percent
2010 SGP	Standards-based	48	92.3%	4	7.7%	52	100.0%
	Non-standards-based	30	100.0%	0	.0%	30	100.0%
	Mixed using general reporting	19	95.0%	1	5.0%	20	100.0%
	Standards						
	Mixed using specific standards	1	100.0%	0	.0%	1	100.0%
2010 CPI	Standards-based	48	92.3%	4	7.7%	52	100.0%
	Non-standards-based	30	100.0%	0	.0%	30	100.0%
	Mixed using general reporting	19	95.0%	1	5.0%	20	100.0%
	Standards						
	Mixed using specific standards	1	100.0%	0	.0%	1	100.0%

The group depicted as ‘mixed using specific standards’ in Table 6 was eliminated from analysis due to the fact that only 1 school’s report card qualified as a member of this group. The analysis of results for this study includes 48 valid cases for standards-based report cards (N = 48), 30 valid cases for non-standards-based report cards (N = 30) and 19 valid cases for the mixed group (N = 19).

Population and Sample Data

The population for this study is the 214 public schools in the southeastern region of Massachusetts as defined by the DESE (see Appendix B). Of the 109 questionnaires and report cards returned to the study, data for five of the schools was eliminated for one of the following reasons: the questionnaire was not accompanied by a copy of the current report card for Grade 4 (N = 2), the MCAS results for the school were under review (N = 2), or the population of

students in Grade 4 was too small to report results ($N = 2$). The sample size for this study is 103 elementary schools ($N = 103$) in the southeastern region of Massachusetts. There was one report card that did not fit into the three groups established for the study. This outlier had specific standards reported for Grade 4 but scored low on each of the other three criteria. Since it was the only report card in an initial fourth category for independent variables, the group was eliminated from further analysis. Three groups were established for the independent variable of report card type. The sample for each of the independent variable groups is 52 schools using standards-based report cards ($N = 52$), 30 schools using non-standards-based ($N = 30$) and 20 schools using a report card that is a mixed representation ($N = 20$). This third independent variable group includes report cards with general standards and indicators that are not specific to Grade 4 mathematics. It was established to test the significance of reporting standards specific to Grade 4 (standards-based group) as opposed to reporting general standards or strands of mathematics (mixed group). The participant samples analyzed from schools in the 3 independent variable groups ($N = 102$) represent approximately 48% of the elementary schools in the southeastern region of Massachusetts that were the targeted population for the study and 9% of the 1138 elementary schools in the state.

Data Analysis

The hypothesis that drives this research study is: Fourth grade results for students in the schools with standards-based report cards will have higher percentages of growth and higher indices of performance as measured by the MCAS Mathematics Test than schools with non-standards-based report cards.

Analysis was conducted using the software program Statistical Package for Social Sciences (SPSS), Version 19. The independent variable of standards-based report card was

labeled *report card type* and established as a nominal measure with three values: 1 – standards-based; 2 – non-standards-based; 3 – mixed using general reporting standards.

The dependent variables of student growth percentile and composite performance index were labeled SGP and CPI respectively and established as scale measures. The value for SGP represents a percentile ranking for growth in comparison to peer groups with similar scores on the Grade 3 Mathematics MCAS in 2009. The value for CPI represents a point value on a scale of 0 to 100 that is calculated from the performance of all students in Grade 4 in the participating school. The covariates of low income percentage, special education percentage, limited English proficient percentage, student-teacher ratio, and per pupil expenditure were each labeled LI, SPED, LEP, teacher to student, and per pupil expenditure respectively. All covariates were established as scale measures.

Each case was assigned a case number in both SPSS and the original Excel file in order to maintain confidentiality of information and integrity of data. A split file was created to sort the variable standards-based report cards by the three groupings developed for the study: standards-based, non-standards-based, and mixed.

Covariate Statistics. A split file was created using SPSS in order to allow for subsequent analysis to be organized by report card group. The population and sample data of the three independent variables used in the study is depicted in Table 7.

Table 7

Range of Covariates for Study Participants

Covariates	N	Range	Minimum	Maximum
Low Income	103	88.00	1.80	89.80
Special education	103	20.40	6.80	27.20
Limited English proficient	103	30.10	.00	30.10
Student-teacher ratio	103	13.70	7.50	21.20
Per pupil expenditure	103	12128.00	9548.00	21676.00
Valid N	103			

Covariates were reviewed using the guidelines suggested by Brace, Kemp and Snelgar (2009). They identify 6 guidelines for choosing covariates:

1. A covariate should be chosen on the basis of existing theory and research.
2. A covariate should ideally be measured using a scale at ratio, interval, or ordinal level.
3. Ideally, a covariate should be measured before the experimental manipulation takes place, as it is assumed that the covariate is independent from the experimental manipulation.
4. A covariate should be measured reliably, that is if it were measured several times a cross a time period, there would be a high correlation between the scores.
5. The relationship between a covariate and the dependent variable must be linear (straight line). If there is more than one covariate they should not be strongly correlated with each other.
6. There should be homogeneity of regression.

The covariates included in this study met each of the first five guidelines. The covariates are commonly accepted influences on student performance based on test results and research. All covariates are measured on an ordinal scale. The covariates were measured and established

independently from this study. The covariates are reliable measures as established by the DESE. The relationships between the covariates and the dependent variables were linear.

In preparation to meet the sixth guideline for establishing covariates the researcher conducted a test of the homogeneity of regression through SPSS. A significant correlation between the two of the covariates was found with one of the dependent variables. The percentages of low income and limited English proficient percentage were found to have a significant correlation to the CPI. This correlation is reported in the results for the CPI independent variable.

The largest variation in the covariates among the three groups exists in the low income percentage with a range of 86.90 within the 52 schools with standards-based report cards. The ranges are significantly less for the non-standards-based group and the mixed group at 64.80 and 46.00 respectively. The complete table of covariate descriptives separated by independent variable grouping is included as Appendix F. A review of the frequency data for the low income covariate indicates that only 13 schools had more than 50% of the populations qualify as low income. A cross-tabulation analysis was conducted to determine if greater percentages of schools with high levels of students qualifying as low income were centralized within the data for a single independent variable. The cross-tabulation reveals that 11 of 13 schools with percentages of low income above 50% are within the standards-based group. The remaining 2 schools are within the non-standards-based group. This variance may influence the statistics reported for the low income percentage covariate.

Multivariate analysis of covariance. This study utilized a full factorial one-way Multivariate Analysis of Covariance (MANCOVA) with a simple contrast to analyze the results. A MANCOVA removes the effects due to a covariate thereby reducing error variance and

leading to a larger F-value (Brace, Kemp & Snelgar, 2009). MANCOVA is an extension of analysis of covariance (ANCOVA) methods to cover cases where there is more than one dependent variable and where the dependent variables cannot simply be combined. MANCOVA is similar to multiple analysis of variance (MANOVA), but controls for the effects of supplementary continuous independent variables or covariates. If there are some covariates, MANCOVA should be used instead of MANOVA (Brace, Kemp & Snelgar, 2009).

Controlling for assumptions of MANCOVA. This study addresses the assumptions of MANCOVA to reduce Type 1 error as follows: 1) Dependent variables are measured using an interval scale to determine the linear relationship; 2) The number of cases in each cell is greater than the number of dependent variables; 3) The assumption of the homogeneity of variance – covariance matrices is checked; 4) A histogram is generated for each dependent variable to assess for normality of distributions (Brace, Kemp & Snelgar, 2009).

This study has two dependent variables: SGP (student growth percentile) and CPI (composite performance index) for Grade 4 students in public schools on the 2010 Mathematics MCAS. These dependent variables cannot be combined as they represent two different measures. The SGP measures growth from 2009 to 2010 as compared to peer groups with similar performances on the MCAS in 2009. The CPI measures the grade level performance for each school as a cumulative representation of all student performance and progress along a trajectory toward 100% proficiency. This study includes five covariates; therefore, the MANCOVA is the most efficient method of data analysis.

In keeping with recommended practices there is one between-subjects factor (report card type) as the independent variable. There are three groups within the independent variable: standards-based report card, non-standards-based report card, and mixed report card with general

standards reported. This study attempts to interpret output from more than one covariate. The primary covariates are those that identify students at risk for learning. The low income percentage, special education percentage, and limited English proficiency percentage are analyzed to inform the original theoretical framework establishing standards-based report cards as a method to improve self-efficacy, motivation, and achievement. The covariates for this study have been reliably determined in advance by the DESE and are publicly available on the DESE website as part of district and school profiles.

The check for homogeneity of variance indicates that there is the likelihood of a Type 1 error. The p-value is the probability level that indicates the level of statistical significance or the probability that the results are a function of chance. Researchers generally recognize a p-value of less than .05 as significant. There is strong interaction between the covariates of low income and limited English proficient and the composite performance index ($p < .05$) so the assumption of MANCOVA is violated for these two groups for the CPI data as shown on the table for between-subjects factors in Appendix G. The interactions are not statistically significant for the SGP ($p > .05$) so the data does not violate the assumption of homogeneity of regression slopes for the SGP. The results for the SGP dependent variable are analyzed using the MANCOVA. The results of the test of homogeneity of variance indicate there is strong correlation between the CPI and the two covariates of low income and limited English proficiency, therefore the results of the MANCOVA for these two covariates on the CPI dependent variable may be invalid. Histograms generated demonstrate normality of distributions.

MANCOVA test of significance. A MANCOVA was conducted to determine if there were differences of effect of the independent grouping variable (report card type) on the two dependent variables (Composite Performance Index and Student Growth Percentile). The

standard significance level for social science research was used; the alpha was set to .05 for the overall MANCOVA. The Wilks' lambda statistic is reported to indicate significance.

A correlation test was conducted between the two dependent variables. If the significance is less than .001 ($p < .001$) the assumption of homogeneity of covariance has been violated. The results indicate that $p = .008$ ($p > .001$) therefore this assumption has not been violated in this study. There is no significant relationship between the two dependent variables as shown in Table 8.

Table 8

Correlation Coefficient for Relations Among the Dependent Variables

Measure	1	2
1. Student Growth Percentile	--	.008
2. Composite Performance Index	.008	--

Note: $p < .01$

Results indicate that there was not a significant effect of the type of report card (standards-based, non-standards-based, or mixed with general standards) on the dependent variable of student performance, $F(4, 176) = .784$, Wilks' lambda = .965; partial $\eta^2 = .17$. Although there was no significant difference as indicated by the MANCOVA a separate ANCOVA test was performed on each independent variable to determine if there was a modest correlation with report card type. As presented in Table 9 the F Ratios show that no significant differences were found in the effect of the report card type on either of the dependent measures, $p > .05$ (see also Appendices I and J).

Table 9

Multivariate F Ratios

	MANCOVA F(4,176)	CPI F(2, 90)	SGP F(2, 91)
Report Card Type	.537	0.487	0.826

Note: $p > .05$

Mean and standard deviation. The mean and standard deviation for the two dependent variables, SGP and CPI for the independent variable of report card type from the MANCOVA analysis are shown in Table 10.

Table 10

Means and Standard Deviations by Dependent Variables for the Three Groups for MANCOVA Analysis

Group	SGP		CPI	
	M	SD	M	SD
Standards-based	50.02	15.54	80.99	9.01
Non-standards-based	52.50	16.30	82.64	7.88
Mixed using general reporting standards	49.42	10.38	83.47	4.34

Note: Study sample: $N = 97$: Standards-based: $N = 48$; non-standards-based: $N = 30$; mixed: $N = 19$)

The non-standards-based report card group had a higher mean SGP ($M = 52.5$, $SD = 16.3$) than either the standards-based report card group ($M = 50.02$, $SD = 15.54$) or the mixed group ($M = 49.42$, $SD = 10.38$). The standards-based report card group had the lowest mean CPI ($M = 80.99$, $SD = 9.01$) of the three types of report cards. The mixed using general reporting standards group had a higher CPI ($M = 83.47$, $SD = 4.34$) than either the standards-based report card group ($M = 80.99$, $SD = 9.01$) or the non-standards-based report card group ($M = 82.64$, $SD = 7.88$).

= 7.88). As noted earlier the difference between the mixed and standards-based group is the lack of specificity of Grade 4 indicators for mathematics.

Rejection or Acceptance of Hypotheses

The following section focuses on the results of the findings of each of the hypotheses based on the quantitative findings. The summary of results section which follows synthesizes the findings of the document review, the quantitative findings, and the findings of additional tests conducted in this study.

The hypothesis that drives this study is: Fourth grade results for students in the schools with standards-based report cards will have higher percentages of growth and higher indices of performance as measured by the MCAS Mathematics Test than schools with non-standards-based report cards. To explore the causal-comparative relationship between report card types and the measures of growth and performance a multivariate analysis of covariance (MANCOVA) was utilized to determine a statistical difference in effect. A significant difference between the groups would be indicated by a p-value of less than .05. Results of the MANCOVA indicate that there was not a significant effect of the type of report card (standards-based, non-standards-based, or mixed with general standards) on the dependent variable of student performance, $F(4, 176) = .784$, Wilks' lambda = .965; partial $\eta^2 = .17$; therefore the researcher rejects the primary hypothesis that focuses this study.

Hypothesis 1: *Standards-based reporting methods that provide feedback to families and students of progress along a continuum sustain motivation toward attainment of proficiency more than traditional grades, therefore producing a higher median SGP and CPI.*

The first hypothesis focused on comparing the effect of report cards that provided performance levels to those report cards that have one overall amalgamated grade for

mathematics. A review of the mean growth percentiles of the three report card types for the SGP suggests a minor difference in growth percentiles achieved by the schools in each group during the year of the study. Both the standards-based report cards and the mixed with general reporting standards groups utilized performance levels that were judged to be methods that provide feedback to families and students along a continuum of progress. The standards-based report card group ($M = 50.02$, $SD = 15.54$) and mixed report card group ($M = 49.42$, $SD = 15.54$) did not have a higher mean SGP than the non-standards-based group ($M = 52.5$, $SD = 16.3$). In fact the non-standards-based report cards had the highest SGP. The findings for the CPI suggest that the schools using a standards-based report card actually had the lowest mean ($M = 80.99$, $SD = 9.01$). The mixed report card group had the highest mean CPI ($M = 83.47$, $SD = 4.34$). The non-standards-based report card group ($M = 82.64$, $SD = 7.88$) also had a higher mean CPI than the standards-based group. As stated in the introduction to this section, this difference was not significant when examining the variance, $F(4,176) = .784$, $p > .05$. No significant difference of effect of type of report card on the measures of growth and performance was found; therefore, the researcher rejects Hypothesis 1.

Hypothesis 2: *Grading systems that eliminate failing grades result in higher levels of motivation, self-efficacy, and growth for students who are at risk for learning, therefore producing a higher median Student Growth Percentile (SGP).*

This second hypothesis focused on the effect of report cards that remove failing grades on growth for students who are at risk for learning when compared with report cards that continue to fail students. The two report card groups coded as standards-based and mixed with general reporting standards both had the removal of failing grades as a criterion for group membership. An examination of the student growth percentile for both of groups indicates that the non-

standards-based report cards, which include failing grades, produced a higher median student growth percentile than either of the groups that removed failing grades. In order to inform this hypothesis further additional analysis was conducted to look at the effect of the independent variable related to covariates representing students at risk. This additional analysis is discussed later in the next section. The MANCOVA performed for this study suggests that the difference among the effects of report card type was not significant when examining the variance, $F(4,176) = .784, p > .05$. No significant difference of effect of removing failing grades on the student growth percentile was found; therefore, the researcher rejects Hypothesis 2.

Hypothesis 3: *Standards-based report cards that separate mastery of content knowledge and skills from learner behaviors lead to higher levels of mastery than amalgamated grades on traditional report cards grades, therefore producing a higher median SGP and higher CPI for schools using them.*

This third hypothesis focused on the practice of separating grades that represent academic achievement from grades that assess student learner behaviors and social behaviors. Traditional report cards have included homework completion, class participation, and collaboration in a group as part of the overall grade. Standards-based report card practices separate academic achievement grades from learner and social behaviors. This study sought to identify the effect of this practice on student learning. The review of report cards included in this study indicated that the separation of learner and social behaviors has become a widely accepted practice regardless of report card type. All three groups of the independent variable had separation of learner and social behaviors as a characteristic. The researcher cannot reject or support Hypothesis 3 based on the data collected and analyzed.

Hypothesis 4: *Standards-based report cards that clearly identify power standards in math, specific to Grade 4, result in higher levels of growth and performance than report cards with one overall content area grade, therefore producing a higher median SGP and CPI.*

The fourth hypothesis focused on comparing the report cards that identify grade-specific standards and indicators with report cards that have one overall grade or a few grades on generalized topic areas in mathematics. The standards-based report card group had specific Grade 4 indicators in mathematics that ranged in number from 7 to 26. The report cards in the non-standards-based report card group ranged from 1 overall grade to 3 overarching topics. The mixed group with generalized standards had 3 to 5 general reporting standards that were not specific to Grade 4. The standards-based report card group ($M = 50.02$, $SD = 15.54$) did not have a higher SGP than the non-standards-based group ($M = 52.5$, $SD = 16.3$) or the mixed with general reporting standards ($M = 49.42$, $SD = 15.54$). The findings for the CPI suggest that the schools using a standards-based report card actually had the lowest mean ($M = 80.99$, $SD = 9.01$). The mixed report card group ($M = 83.47$, $SD = 4.34$) and the non-standards-based report card ($M = 82.64$, $SD = 7.88$) both had a higher mean CPI than standards-based. Schools using grade-specific indicators of performance on standards-based report cards had similar results to schools with report cards that have only one overall mathematics grade or several generalized indicators of performance on the Grade 4 Mathematics MCAS. As stated in the introduction to this section, the difference of effect was not significant when examining the variance, $F(4,176) = .784$, $p > .05$. No significant difference was found between report cards that used grade-specific indicators of performance and those report cards using generalized overall grades or one amalgamated grade when examining the effect on measures of growth and performance; therefore, the researcher rejects Hypothesis 4.

Summary of MANCOVA results. No significant differences of effect between standards-based report cards, non-standards-based report cards, or mixed report cards with general reporting standards were found using a MANCOVA. The primary hypothesis and three other hypotheses were rejected by this study. The research into the third hypothesis of this study did not result in a finding. It is important to note the second hypothesis that focused on the effect of failing grades on students who are at risk for learning was informed by grade level results only. The subgroups representing low income, special education, and limited English proficient in many of the schools were too small for results to be publicly reported by the DESE. The DESE does not report results for groups less than 40 in order to maintain confidentiality. In order to further inform this hypothesis additional analysis of the data for students who are considered at risk for learning was conducted.

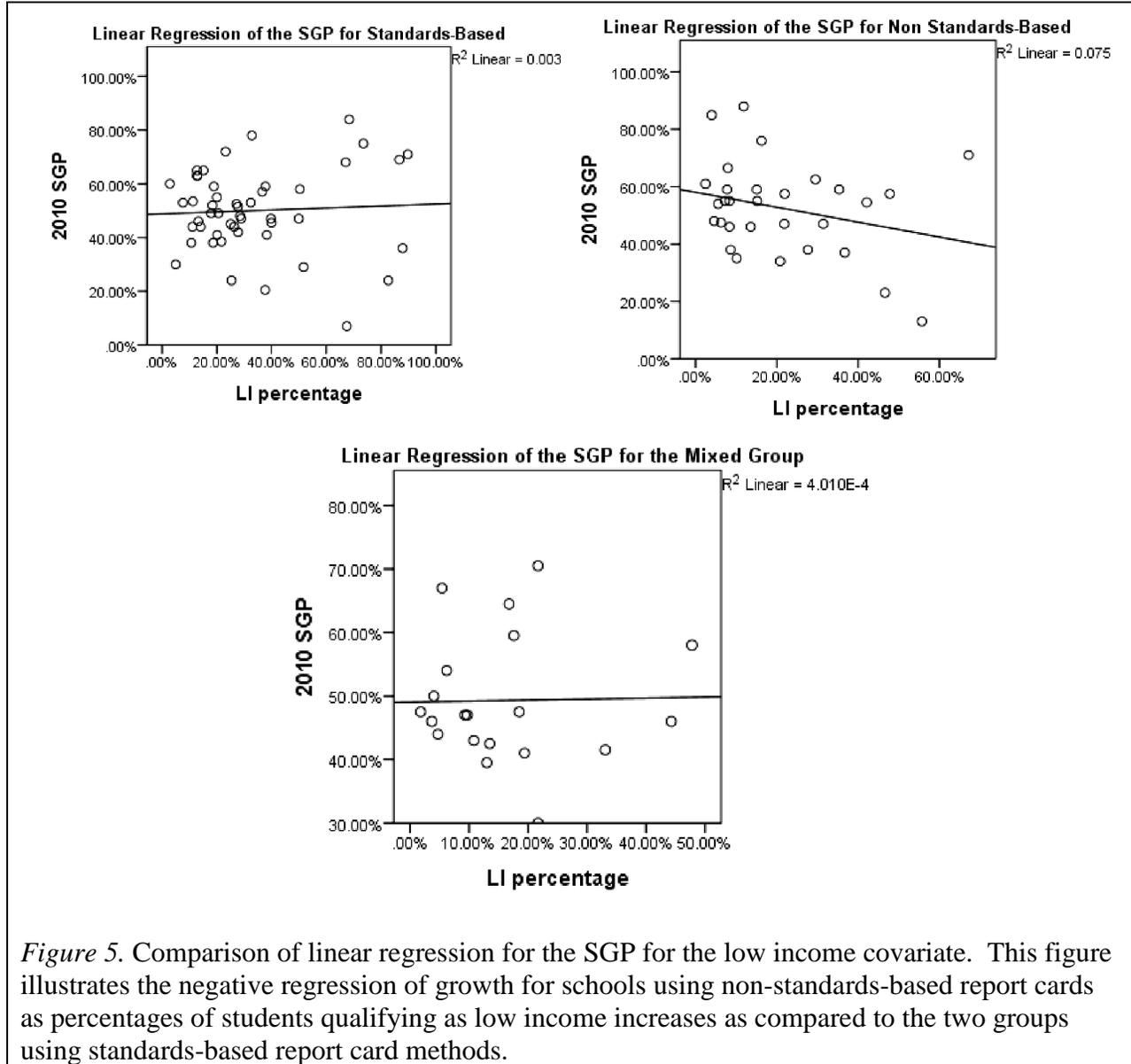
Additional Analysis of Results

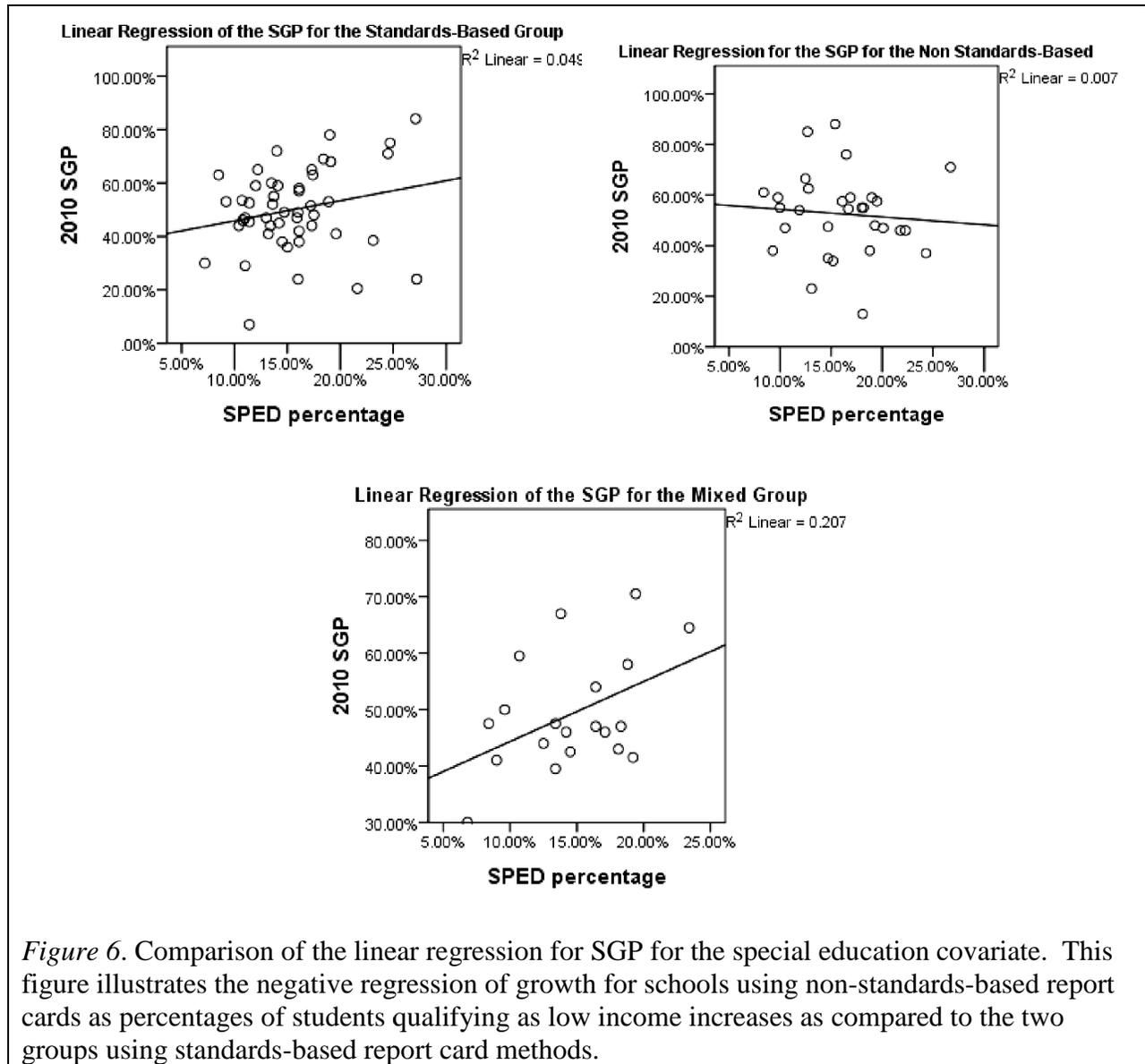
Tests of linear regression. As an additional measure to discern the presence of a relationship between standards-based report cards and the performance of students who are at risk for learning scattergrams were created for the SGP for each of the independent variable groups incorporating each covariate and the dependent variables. Figure 5 shows that the linear regression line is relatively flat for changes in growth for schools with increasing percentages of students qualifying as low income when schools have a standards-based report card or a mixed report card. The regression line in the non-standards-based group scattergram does indicate a correlation between declining growth percentiles and increasing percentages of students qualifying as low income. These findings indicate that there is no correlation of effect of standards-based and mixed report cards on the growth of students in the low income subgroup. However, the results for the non-standards-based group indicate a negative impact on growth for

populations of students qualifying as low income. These findings provide some support for the hypothesis that standards-based grading improves self-efficacy, motivation, and achievement, particularly for students who are most at risk for learning when compared to non-standards-based report cards.

The scattergrams created for the covariate of special education percentages for each category of standards-based report cards depicted in Figure 6 similarly show an effect of declining growth on the dependent variables when non-standards-based report cards are used in schools. When schools use a standards-based report card as defined by this study the regression line for the special education population shows an actual increase in growth as populations of students receiving special education increases. The scattergrams support a correlation between growth percentages and standards-based report cards for populations of students receiving special education. Scattergrams created for the limited English proficient subgroup did not indicate a difference and showed similar regression lines for all three report card types.

These findings suggest support for the second hypothesis that standards-based grading which includes the removal of failing grades, improves self-efficacy, motivation, and achievement, particularly for students who are most at risk for learning when compared to non-standards-based report cards.





Additional analysis of Hypothesis 2: *Grading systems that eliminate failing grades result in higher levels of motivation, self-efficacy, and growth for students who are at risk for learning, therefore producing a higher median Student Growth Percentile (SGP).*

The removal of failing grades was common to both the standards-based report card group and the mixed with general reporting standards group. The additional analysis of the linear regression for two of the covariates- students qualifying as low income and students receiving special education - suggests that there is a correlation between the removal of failing grades and improved student growth. The use of standards-based report cards may help to counteract the negative impact on growth for populations of students qualifying as low income and students receiving special education services; therefore the researcher accepts this hypothesis as having some support through additional analysis. Further study is needed in order to fully accept this hypothesis.

Summary of Research Findings

The primary research question for this study is: What effect do standards-based report cards have on the growth and performance of fourth grade students on the MCAS Mathematics test?

This study used the following definition of standards-based report cards: a standards-based report card has standards specific to Grade 4 mathematics curriculum frameworks, eliminates grading language that indicates failure, separates grading for learner behavior and social behavior from academic achievement, and includes at least three levels of performance reporting.

Continuum of progress. As noted in the document review, the array of performance levels terminology that is used to inform students and families of student progress along a

continuum of learning is vast. The document review of the 103 report cards included in the study revealed reporting systems that are common neither in descriptors of performance nor in labels of achievement. The terminology used in report card performance levels should be consistent, recognizable, and easily understood, but the labels or letters used as representation of those levels may be irrelevant. This study found that it is possible for school districts to maintain the use of ABCs for student grading by redefining descriptions to reflect the best practices of standards-based grading.

The current assortment of representations of progress from letter grades to percentages to ordinal scales to abstract letter representations has not clarified student or family understanding of student learning progress. The growth and performance of students in schools using standards-based report cards was not significantly different from students in schools using non-standards-based report cards. This study does not support the hypothesis that providing feedback to students and families along a continuum of progress, as opposed to traditional letter grades or percentages, results in higher levels of growth or achievement.

Failing grades. The MANCOVA conducted on the data for this study did not find any significance in effect of the types of report cards on fourth grade students' growth and performance. Additional analysis was conducted to examine the effect of the independent variables on the performance of schools related to the percentage of covariates. This study included covariates of students at risk for learning. The three covariates included that reflect the subgroups of students who are risk are: low income, special education, and limited English proficiency. Analysis was conducted using tests for linear regression.

The growth of schools using non-standards-based report cards showed a negative regression as the percentage of students in the special education group increased. However, the

growth and performance of schools using standards-based report cards or mixed report cards showed a flat regression line indicating that the negative impact of increases in percentages of students receiving special education in schools may be neutralized by the practice of grading along a continuum of progress (see Figures 5 & 6). Results for increasing percentages of students qualifying as low income revealed a less dramatic but similar finding. Schools using non-standards-based report cards showed a negative regression as percentages of students qualifying as low income increased. The schools using standards-based and mixed report cards did not parallel that pattern and maintained a constant flat regression line as percentages increased. The results for the percentages of students who are limited English proficient did not indicate a notable difference between the three report card types. Further study on the impact of standards-based practices on the performance of students who are identified as at risk for learning may provide additional insight into this finding.

The results of the linear regression tests suggest that the elimination of failing terminology from report cards may have a positive impact on the growth and performance of students who are at risk for learning and neutralize the decline in growth that exists when non-standards-based report cards are used. These findings support the hypothesis that standards-based grading improves motivation, self-efficacy, and growth, particularly for students who are most at risk for learning when compared to non-standards-based report cards.

Separating learner and social behaviors. This study did not examine this third hypothesis through data analysis once the review of report cards indicated that the separation of learner and social behaviors has become a widely accepted practice regardless of report card type.

Standards and indicators. The report cards included in the standards-based group for this study did not reflect the true definition of power standards as defined in the literature review (Ainsworth, 2003, Guskey & Bailey, 2010). Only 9 of 103 report cards from area schools met the researchers' definition of power standards and received a 4 on the Standards-Based Report Cards Product Rating Scale. The report cards coded as standards-based according to the grouping rubric had indicators of performance that were specific to Grade 4 mathematics. The two other report card types had either one overall standard or very generalized topics or standards. A MANCOVA performed did not show significance of effect for the type of report card. The confidence interval did not approach 95% on the results for either the SGP or the CPI. The growth and performance of students in schools using grade-specific indicators of performance was not significantly different from the growth and performance of students in schools using one overall grade or general standards.

In summary, the results of this study indicate that there is no statistical difference in the improvement of student growth and performance in mathematics for the general population of Grade 4 students regardless of a schools type of report card.

Chapter 5: Discussion of Research Findings; Implications for Educational Practice

This study examined the causal-comparative relationship between the use of standards-based report cards to provide student feedback and corresponding student growth and performance on the Massachusetts Comprehensive Assessment System (MCAS), a criterion-referenced assessment. The practical goal of this research project was to build a quantitative foundation for the expansion of the use of standards-based report cards to the secondary level in public schools.

This study examined the components of standards-based report cards and made inferences regarding the impact of standards-based report cards on improving outcomes for student learning by gathering data on the Student Growth Percentile (SGP) and the Composite Performance Index (CPI) on the MCAS. The results of this study may contribute to reflection and discussion as district leaders prepare to facilitate the cultural shift from traditional grades to standards-based report cards at the secondary school level.

Research Questions

The primary research question for this study is: What effect do standards-based report cards have on the growth percentiles and performance indices of fourth grade students on the MCAS Mathematics Test?

The additional research questions that inform the hypotheses presented in this research study are:

1. Do standards-based report cards that report student performance on a continuum of progress in mathematics rather than a cumulative grade average result in higher levels of growth and performance on the MCAS?
2. Do standards-based report cards that remove indications of failure such as Fs and zeros in mathematics result in higher levels of growth on the MCAS for students at risk for learning?
3. Do standards-based report cards that separate academic progress in mathematics from social and learner behaviors result in higher levels of growth and performance on the MCAS?

4. Do standards-based report cards that grade in power standards result in higher levels of growth and performance than report cards with a cumulative grade or standards-based report cards that do not provide power standards?

Interpretation of Findings

This study is framed by the theory that effective feedback, in the form of standards-based report cards, promotes self-efficacy and motivation and supports more opportunities for learning within the zone of proximal development, especially for students who are at-risk for learning. The review of the literature, the coding of 103 elementary report cards, and the data analysis involved in this study have led the researcher to one overall conclusion and four specific conclusions as follows:

Overall conclusion: Student feedback in the form of standards-based report cards may be having little impact on improving the growth and performance of students in elementary mathematics.

Specific conclusions:

1. There is a general lack of shared understanding among school leaders about what constitutes a standards-based report card.
2. Given that this study found no significant difference in the performance or growth of schools whether or not they use one overall grade in a content area or several newly developed performance levels, the efforts of school leaders to engage stakeholders in this significant change may not be producing the desired improvements in learning.
3. Report cards that grade on multiple grade-specific indicators of proficiency in mathematics did not produce higher levels of growth and performance for Grade 4 than report cards that provided a single overall content grade, therefore school leaders' should

question the format of report cards that may burden families with having to interpret achievement and progress based on dozens of performance indicators.

4. Preliminary results of additional tests conducted in this study suggest that the elimination of failing grades may have a positive impact on growth for students who qualify as low income or receive special education.

In the following sections each conclusion is discussed in detail related to the review of the literature, the process of collecting data collection and the analysis of data conducted in this study.

Shared understanding. The questionnaires that were used in this research study began with what seemed to be a straightforward question: Does your school use a standards-based report card? (see Appendix A). When the questionnaire was developed the researcher expected that the response to this question would be the initial step in sorting the responses into standards-based report cards and traditional report cards. The very first, highly anticipated response was a harbinger of the lack of clarity that exists in discerning what makes a standards-based report card. This first respondent to the questionnaire indicated that their elementary school did not use a standards-based report card and thus did not continue to answer the remaining questions. The respondent did include a copy of their report card along with the questionnaire as requested. The researcher reviewed this report card despite the respondent's answer to the first question. This first report card actually met 2 of the 4 criteria for a standards-based report card. As the cataloging of report cards proceeded it was determined that this first report card received a higher rating than some of the report cards that were returned with an affirmative response to the question asking if the school had a standards-based report card. The researcher made the

decision to review every respondent's report card using the product rating scale created for this study regardless of the school leader's response to the first question.

Twelve of the report cards returned by school leaders who reported that they did not have standards-based report cards actually qualified for membership in either the standards-based group or the mixed group for the purpose of the study. While not an original proposal of the study, the researcher reviewed the pattern of responses on the questionnaires and determined that the primary factor that appeared to be employed by participants to discern whether or not their report card was standards-based was the removal of traditional grades and the replacement of those grades with a system of performance levels. A more refined and consistent application of the definition of standards-based report cards is needed.

Performance levels. The impetus for standards-based grading has been the urgency to identify student strengths and weaknesses in terms of content knowledge and skills. The more defined standards-based indicators of learning give much more information than the overall 'C in Math'. Guskey and Bailey make a recommendation that standards-based report cards include 4 to 5 performance levels with "2 distinct lower levels of performance" to provide scaffolding opportunities for those students who have not yet reached proficiency and one level higher than proficiency to "recognize those students who display truly exceptional accomplishment or skill with regard to the standard" (2010, p. 43). Standards-based report cards are designed to inform students and families of their progress along a continuum of skills and builds self-efficacy and motivation. One overall letter grade representing a cumulative average in Mathematics could not possibly represent a student's knowledge and skills across five strands and dozens of standards.

In efforts to move away from the practice of averaging, educators have adopted such a variety of reporting systems that it is almost impossible to make comparisons on student

performance from one school to the next. The document review completed as part of this study found that there were numerous performance level labels used and that the labels represented different levels of performance from one school to the next. In one school a performance level of 2 on a report card may mean the student is at grade level proficiency. It may be below grade level proficiency at another school and approaching grade level proficiency with support at yet another school. A letter grade of C may mean average performance, meets grade level expectations, needs improvement, or below average depending on which school a student attends. Some schools use both a letter grade to report an overall grade for a subject and then performance levels to report on specific indicators within the content level. It is no wonder that families and teachers are resistant to the change to standards-based grading.

The results of this study indicate that there is no significant difference in the results for schools on the Grade 4 mathematics MCAS whether they use one overall letter grade on non-standards-based report cards or newly developed performance levels on standards-based or mixed report cards. Schools that continue to use an ABC grading system or percentage system on non-standards-based report cards performed similarly to schools using multiple performance levels and descriptors in measures of both growth and performance. As educators move forward in extending or revising standards-based report cards there is support for those who question the adoption of a different grading system. It may be that the redefinition of ABC to a consistent terminology that provides a continuum of progress and achievement may be sufficient.

Multiple indicators of proficiency. Standards-based report cards have been established around the premise that students need to have learning goals and progress toward those goals identified on their report card. A crucial consideration in identifying learning goals or standards is determining the degree of specificity. This study examined the difference in results for schools

using specific Grade 4 standards (standards-based group) and schools using general reporting topics for mathematics (mixed with general reporting standards group) or one overall content grade (non-standards-based).

Standards that are too specific make reporting forms cumbersome to use and difficult to understand. Standards that are too broad or general, however, make it hard to identify students' unique strengths and weaknesses. Most state-level standards, for example, tend to be broad and need to be broken down or "unpacked" into homogeneous categories or topics (Marzano, 1999). For grading and reporting purposes, educators must seek a balance. The standards must be broad enough to allow for efficient communication of student learning, yet specific enough to be useful (see Gronlund, 2000; Marzano & Kendall, 1995; Wiggins & McTighe, 1998). Developing a new report card is more a challenge in effective communication than simply documenting or quantifying student achievement (Guskey & Jung, 2006).

This study reviewed the reporting standards or indicators for 103 report cards and found that the number of standards ranged from 1 overall grade in math to 26 specific indicators of performance in math. Thomas Guskey (2001) describes the enormous amount of work that educators must spend identifying the learning goals or standards as a shortcoming of standards-based report cards and remarks on the considerable addition to the workload of teachers and school leaders. He notes another shortcoming is the sometimes complicated report cards that represent standards-based grading. Guskey observes that the efforts of leaders can sometimes result in cumbersome and time-consuming tools that are difficult for families to understand. There must be a balance in order to fulfill the purpose of report cards - to provide efficient and effective communication of student progress and learning to families.

This study did not find a significant difference in either the student growth percentile (SGP) or the composite performance index (CPI) for the groups of schools using three different types of report cards identified for the study. The schools that are asking families to decipher learning progress based on their child's performance level on up to 26 indicators for mathematics (and additional indicators for other content areas) should reflect on the time and effort that is required for this exercise given the results of this study. "Inundating parents and others with information that does not make sense or that they cannot use defeats the purpose of reporting" (Guskey & Bailey, 2009, 623 of 2391).

Failing grades. Failure must be eliminated from any new system of reporting progress and achievement. Failing grades are not often reflective of a student's mastery of content but may be more an assessment of compliance and social skills. These poor grades do not build a student's belief in his or her own ability to learn content; do not create a sense of self-efficacy; and will result in decreased motivation to continue striving to learn. Yet almost half of the schools included in this study continue to issue failing grades. Failing grades have been used historically as a punishment for students who do not meet the expectations of the school environment. The impact of failing grades on student motivation to learn and their view of themselves as learners is documented in the literature (Reeves, 2004, Guskey, 2002, Guskey, 2004b, Guskey, 2010). Schools that have adopted standards-based grading systems have embraced this research and have removed determinations of failure from elementary report cards. The review of report cards conducted as part of this study found that 54% of the schools participating have adjusted language for performance that is below proficiency to provide an expectancy of success at some point along a progression of learning. These report cards include

language such as *beginning*, *approaching*, or *not yet*, and provide students with statements that show they are moving along a continuum and have not failed to achieve or make progress.

Students who are at risk for learning may be especially victimized by traditional forms of grading because traditional grades are often more representative of work habits and conformity than mastery of key concepts and skills. Antiquated methods of assessing and reporting are now being actively revised to “permit the possibility that all students could succeed at some appropriate level” (Stiggins, 2005, “A New Mission,” para. 3). This study found that student growth was negatively impacted when the percentage of students qualifying as low income increased if non-standards-based report cards were utilized by schools. The study found that there was an absence of impact as low income percentages increased when standards-based or mixed report cards were utilized by schools. Both the mixed group and the standards-based group had removal of failure as a grade as one of the criteria for group membership. These findings suggest that the removal of failing grades and the practice of grading along a continuum of progress may have a positive impact on growth for students in at risk populations. This result supports the theory that standards-based report cards promote self-efficacy and motivation and facilitate learning in the zone of proximal development for students at risk for learning. Standards-based grading practices that remove failure as a grading option may improve outcomes for students in at risk populations.

Effect of student feedback on performance. This study sought to make a quantitative connection between providing students with effective feedback about what they know and are able to do through standards-based report cards and corresponding improvements in growth and performance. The findings of this study indicate that there is no significant difference in either student growth or performance on the Grade 4 MCAS for mathematics for 2010 among schools

using standards-based report cards, non-standards-based report cards, or mixed report cards.

While further studies are needed, the results of this study should give pause to educators who are preparing to engage in several years of research and planning in order to adopt standards-based grading practices at the secondary level.

The results of this study do not deter from the findings of well-known researchers that standards-based reporting methods better inform families, students, and teachers of student progress along a continuum of learning. The body of research supports the implementation of standards-based grading based on three primary elements that impact classroom practice: improved communication of student progress, the standardization of grading benchmarks, and the centralization of assessment by the teacher (Aidman, Gates & Deterra Sims, 2001; Guskey, 2004; Hargis, 1990; Marzano, 2003; Marzano & Kendall, 1996; McTighe & Bailey, 1996; Salend, 2005; Stiggins, 2005, Ward, 2004). Numerous studies support the statement that standards-based report cards improve communication of what students know and are able to do (Aidman, Gates & Deterra Sims, 2001; Ainsworth, 2003; Guskey, 2004; Guskey & Bailey, 2001; Hargis, 1990; Marzano, 2003; Marzano & Kendall, 1996; McTighe & Bailey, 1996; Salend, 2005; Stiggins, 1996; Stiggins, 2005; Ward, 2004). This study examined only the impact of report card type on student and growth and performance for one year of MCAS in mathematics and did not examine classroom practices that inform those report cards.

Implications

Implications for practice. The findings of this study generate many questions related to the impact of standards-based report cards. The study found that there is great variation in the types of report cards used and disparity among district leaders' understanding of what constitutes a standards-based report card. The coding of report cards that was the initial step in conducting

this study revealed that educational leaders did not properly identify their report cards as standards-based and further definition of what constitutes a standards-based report card is needed. Analysis of the report cards according to the rating scale developed for the study revealed that districts required teachers to report on up to 26 different indicators within mathematics for students in Grade 4. Educational leaders should seriously consider if the wave to adopt standards-based report cards has resulted in such a cumbersome tool that the report cards have lost their purpose of communicating effectively and efficiently to parents. Does the time, effort, energy, and anguish expended for developing a standards-based report card result in a value-added communication tool?

This study conducted a document analysis of 103 report cards and rated them across 4 criteria based on the research of Thomas Guskey and Jane Bailey (2010) in *Developing Standards-Based Report Cards* and Douglas Reeves and Larry Ainsworth in Ainsworth's 2003 text *Power Standards, Identifying the Standards that Matter Most*. Analysis of the report cards using the rating scale showed that most districts across the three independent variable groups had report cards that separated learner and social behaviors from academic progress. Of the 103 schools responding, 54% have removed failing grades from report cards. It appears that educational leaders are recognizing that student motivation is enhanced by a continuum of progress, not a designation of failure.

The results of the rating scale also revealed a wide variation in the two remaining criteria for report cards: performance levels and standards and indicators. Schools vary not only in the amount of performance levels used but also in the representation of grades as letters, numbers, or descriptive phrases that were used to define those performance levels for students and families. This variation makes it difficult to relate the performance at any one school to performance at

another school or to relate the report card grade to performance on the MCAS. Schools also varied widely in the number of indicators reported in mathematics for Grade 4 students, ranging from 1 overall grade to 26 specific indicators of performance. It is time to revisit the motivation for standards-based report cards and ensure that schools are developing tools that are consistent in their efficiency and effectiveness in reporting progress and achievement to families. As the new Massachusetts Curriculum Frameworks for Mathematics Incorporating the Common Core is fully implemented in public schools in the coming year educational leaders must reconsider the best method of reporting progress and the most essential indicators of learning to include on student report cards.

This study may assist educators in identifying the essential elements of standards-based report cards: grading along a continuum of progress; elimination of failing grades; separation of learner behaviors from academic achievement; and the development of power standards, in order to build a foundation of self-efficacy, motivation, and achievement for students. The Standards-Based Report Card Product Rating Scale may be useful as a rubric to guide committees of educators, families, and students as they embark on the long journey from traditional grading to standards-based report cards.

Additional analysis conducted through the study indicated that schools using standards-based report cards had some success in counteracting the negative impact of increasing populations of students from low income families and students receiving special education services on grade level growth percentiles. Based on this finding, students who are at risk for learning appear to benefit from standards-based report cards. This finding raises questions for additional inquiry: What specific elements of standards-based practice may be impacting growth and performance for students in at risk populations? How do district leaders frame report card

practices to support the foundation of self-efficacy and motivation that is so necessary for the continued growth and improved performance of students who are at risk for learning?

Implications for further research. This study analyzed results for elementary schools within the southeastern region of Massachusetts and found that there was no statistical significance in the causal relationship between standards-based report cards and student growth and performance on the 2010 Grade 4 mathematics MCAS. The study included an additional analysis of the results for schools on the basis of covariates of low income and special education and found some indications that a causal relationship may exist for these at risk populations. A future study could focus analysis on the impact of standards-based practices on the performance of students who are identified as at risk for learning across urban schools which tend to have higher populations of students from low income homes and higher populations of other demographic groups that are at risk for learning.

This study initially attempted to collect information from participants on the length of time that standards-based report cards had been utilized within their school district. This analysis could not be included in the final study because 20% of the schools categorized as standards-based did not recognize that they were using standards-based practices and therefore did not respond to length of utilization questions. In addition, several schools were unable to supply the length of utilization information because they reported no access to that historical information. A longitudinal study that includes growth data and performance data for schools over a period of years of implementation of standards-based report cards could reveal additional insights into their impact on learning.

During the coding of report cards completed for this study it was noted that numerous schools had adopted the practice of indicating that students were graded on a modified

curriculum. The modification of the curriculum generally occurs only when a student has a disability or receives services in English Language Education. Since grading on a modified curriculum may provide more opportunities to depict growth and success for students who are not achieving grade level mastery, there may be some impact of this practice on students' motivation and sense of self-efficacy. A causal-comparative study into whether or not grading students on a modified curriculum has any impact on student performance and achievement when compared to schools that do not have such distinction on report cards would have implications for the expansion of this practice to all elementary and secondary schools.

This study examined the difference in growth and performance for schools using more traditional, non standards-based practices of providing a typical overall grade of ABC with those schools using performance levels in multiple indicators of proficiency for mathematics. The study did not look specifically at the performance level systems used but identified samples of schools that continued to use the familiar ABC system of grading by providing updated definitions of those grades and removing failing options. The practice of removing the familiar ABC and shifting to a grading system that is less recognizable may also be having an impact on high-achieving students and their sense of motivation and self-efficacy that has not been explored in the literature or this study. Given the resistance of families, educators, and students to the shift away from familiar ABC grades it would be beneficial to the educational environment to focus a study on the difference in growth or performance when grades depicted are ABC or 123 or Advanced, Proficient, Needs Improvement, Warning to see if educators are needlessly expending effort on revising a system that needs only minimal adjustment.

The product rating scale developed for this study was useful in applying the best practices reported in the research to the evaluation of the report cards included in the study. However,

since the report cards varied so widely in the application of these best practices, it was necessary to develop a secondary grouping rubric to code the report cards into three manageable groupings. These groupings developed out of the ratings that were assigned during the coding process. The process for coding the report cards was therefore time-consuming and required many re-assessments of the rating data. The development of a more streamlined cataloguing system would be beneficial to any subsequent studies using similar methodology to this study.

This study explored the impact of four best practices defined in the literature: use of performance levels to identify progress and achievement; removal of failing grades; separation of achievement from grades for learner and social behaviors; and the use of power standards. This study captured results for Grade 4 mathematics MCAS for the spring 2010 test. The results of the study are not conclusive. Comprehensive studies of the impact of standards-based report cards on the growth and performance of several grade levels and several content areas could determine a greater indicator of statistical significance of the causal relationship. Studies that examine the impact on students in urban districts or students in ethnic minority groups may provide additional insight into the impact on children who are most at risk for learning. These additional studies are needed to conclusively ascertain the relationship between standards-based report cards and student growth and achievement.

Conclusion

Leaders of elementary schools are adopting standards-based report cards in order to align local forms of reporting achievement with the results reported on high-stakes testing mandated by No Child Left Behind. School leaders, families, and educators have been asked to accept that the traditional reporting tools that have been used for the past century to report student progress provide minimal representation of student achievement. These stakeholders have gradually

made the adjustment from report cards that depict achievement as an amalgamated ‘A in Math’ to the current standards-based report cards that report progress in anywhere from 3 to 26 indicators of learning proficiency.

This study did not find a significant relationship between report card type and growth or performance for fourth grade students in mathematics on the 2010 MCAS. Perhaps the reason that the schools with multiple indicators of proficiency did not score higher than the schools with general standards is because they have become victims of the “mile wide, inch deep” problem with the standards. By state law, educators are required to teach all of the standards. The sheer number of standards in the example of the school district with a report card requiring that teachers must issue a performance rating on 26 difference indicators in mathematics each term shows that there is little time to focus on the important power standards of mathematics. “The issue here is not whether students should learn all the concepts and skills embedded in all the state standards, but whether or not this is an achievable goal within the instructional time available in any given school year” (Ainsworth, 2003, Chapter 1, “Respect for State Standards,” para. 4). The efforts undertaken to develop a common core of standards nationwide are a response to the problem of too many standards. The recent release of the Massachusetts State Curriculum Incorporating the Common Core in March of 2011 appears to be a step in the right direction toward focusing instruction in Mathematics. The new frameworks align with the recommendations of Ainsworth, Reeves, and others to prioritize mathematical learning throughout the grades (Ainsworth, 2003). There are domains which focus student learning and provide a much clearer progression for learning through the grades. Early elementary grades will now have a more defined focus on domains such as *Counting and Cardinality* while domains such as *Expressions and Equations* will be addressed later in middle school and high

school (MA DESE, 2011). These adjustments to the curriculum should help to focus both the classroom instruction and the pending revision to the Massachusetts assessment system to better identify student learning in mathematics. As school leaders readjust their report cards to align with the new frameworks and assessment they should take into account the recommendations from researchers to include indicators of enduring understandings rather than encompassing every indicator in the frameworks (Ainsworth, 2003).

The development of a standards-based report card takes considerable time and effort on behalf of teams of educators and community members to identify the learning goals or standards and to develop performance levels to replace grades. Given the results of this study and the considerable resources that must be expended to implement standards-based report cards, careful reconsideration should be given to how those efforts are resulting in improved student learning.

References

- Aidman, B. J., Gates, J. M., & Deterra Sims, E. (2001, January). Building a better report card. *The Education Digest*, 66(5), 49-53.
- Ainsworth, L. (2003). *Power standards: Identifying the standards that matter most*. [Kindle edition]. Retrieved from http://www.amazon.com/Power-Standards-Identifying-that-Matter/dp/097094554X/ref=sr_1_1?s=books&ie=UTF8&qid=1299443453&sr=1-1
- American School Board Journal. (2006, October). An A+ report card. *American School Board Journal*, Vol. 193, Issue 10, p.10.
- Allen, D. (1998). *Assessing student learning from grading to understanding*. Columbia University, New York and London: Teachers College Press.
- Bandura, A. (1993, March). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, <http://www.informaworld.com/smpp/title~content=t775653642~db=all~tab=issueslist~branchs=28-v28>28(2), 117 – 148.
- Bandura, A. (1977). *Social learning theory*. New York: General Learning Press.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review Psychology*, 52, 1–26.
- Bagley, W. C. (1939). The significance of the essentialist movement in educational theory. *The Classic Journal*, 34(6), 326-344.
- Brace, N., Kemp, R. & Snelgar, R. (2009). *SPSS for psychologists, 4th edition*. New York, NY: Taylor and Frances Group.

- Butin, D.W. (2010). *The education dissertation: A guide for practitioner scholars*. [Kindle edition]. Retrieved from http://www.amazon.com/Education-Dissertation-Guide-Practitioner-Scholars/dp/1412960444/ref=sr_1_2?s=books&ie=UTF8&qid=1299443703&sr=1-2
- Carter, L. (2007). Total instruction alignment: from standards to student success. In *Comprehensive Appraisal for District Improvement, Windsor Weld County RE-4, Process Conducted September 27 through October 1, 2010, Weld County RE-4 School District*. Bloomington, IN: Solution Tree Press.
- Fraenkel, J. R. & Wallen, N.E. (2009). *How to design and evaluate research in education, seventh edition*. New York, NY: McGraw-Hill.
- Friedman, T. (2005). *The world is flat*. New York, NY: Picador/Farrar, Straus and Giroux.
- Fullan, M. (2001). *Leading in a culture of change*, San Francisco, CA: Jossey-Bass.
- Gredler, M. (2009). Hiding in plain sight: the stages of mastery/self-regulation in Vygotsky's cultural-historical theory. *Educational Psychologist*, 44(1), 1-19.
- Grusec, J. (1992). Social learning theory and developmental psychology: the legacies of Robert Sears and Albert Bandura. *Developmental Psychology*, 28(5), 776-786.
- Guskey, T. R. (2001, September). Helping standards make the grade. *Educational Leadership*, 59(1), 20-27.
- Guskey, T. R. (2002, June). Computerized gradebooks and the myth of objectivity. *Phi Delta Kappan*, 83(10), 775-780.
- Guskey, T. R. (2004a). The communication challenge of standards based reporting. *Phi Delta Kappan*, 86(4), 326-329.
- Guskey, T. R. (2004b, October). Zero alternatives. *Principal Leadership*, 5(2), 49-53.

- Guskey, T. R. (2006, May). Making high school grades meaningful. *Phi Delta Kappan*, 87(9), 670-675.
- Guskey, T. R., & Bailey, J.M. (2001). *Developing grading and reporting systems for student learning*. Thousand Oaks, CA: Corwin Press.
- Guskey, T. R. & Bailey, J. M. (2010). *Developing standards-based report cards*. Thousand Oaks, CA: Corwin Press.
- Guskey, T. R. and Jung, L. A. (2006, Winter). The challenges of standards-based grading: A well-planned report card can help parents relate standards to their children. *Leadership Compass*, Vol. 4, No. 2.
- Hearn, J. (2004). *Assessment as a tool for learning*. Education Oasis. Retrieved from http://www.educationoasis.com/curriculum/assessment/assess_article.htm
- Hirsch, E.D., Jr. (2007). *The knowledge deficit: closing the shocking education gap for American children*. Boston, MA: Houghton Mifflin Company.
- Hursh, D. (2005, October). The growth of high-stakes testing in the USA: accountability, markets and the decline in educational equality. *British Educational Research Journal*, 31(5), 605-622.
- Lalley, J. P. & Gentile, J. R. (2009). Classroom assessment and grading to assure mastery. *Theory into Practice*, 48:28-35. The Ohio State University, doi:10.1080/00405840802577577.
- Marzano, R. (2000). *Transforming classroom grading*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Marzano, R. J. (2003, February). Using data: two wrongs and a right. *Educational Leadership*, 60, 56-61.

Marzano, R. (2006). *Classroom assessment and grading that work*. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. J. & Kendall, J. S. (1996). *A comprehensive guide to designing standards-based districts, schools, and classrooms*. Aurora, CO: Mid-Continent Regional Educational Laboratory.

Massachusetts Department of Elementary and Secondary Education (2002). *No Child Left Behind*. Retrieved from <http://www.doe.mass.edu/nclb/>

Massachusetts Department of Elementary and Secondary Education (2006, August). *School leaders' guide to the 2006 cycle IV accountability and adequate yearly progress (AYP) reports*. Retrieved from <http://www.doe.mass.edu/sda/ayp/cycleIV/glossary.html>

Massachusetts Department of Elementary and Secondary Education (2009, October). *MCAS student growth percentiles: Interpretive guide*. Retrieved from <http://search.doe.mass.edu/search.aspx?q=student+growth+percentile>

Massachusetts Department of Elementary and Secondary Education (2010). *MA DESE school leader's guide to 2010 school and district accountability reports*. Retrieved from <http://www.doe.mass.edu/sda/ayp/2010/default.html>

Massachusetts Department of Elementary and Secondary Education (2010). *2010 Lists of Massachusetts Schools and Districts by NCLB Accountability Status and Accountability and Assistance Level*. Retrieved from <http://www.doe.mass.edu/sda/ayp/2010/>

Massachusetts Department of Elementary and Secondary Education (2011). *Massachusetts Curriculum Framework for Mathematics, March 2011*. Retrieved from <http://www.doe.mass.edu/candi/commoncore/>

- McMillan, J.H., Myran, S. & Workman, D. (2002, March/April). Elementary teachers' classroom grading practices. *The Journal of Educational Research*, 95(4), pp. 203-213.
- McTighe, J., & Bailey, J. (1996). Reporting achievement at the secondary level: what and how. [Yearbook] *Association for Supervision and Curriculum Development* (pp. 119-140).
- McTighe, J. & Thomas, R. S. (2003, February). Backward design for forward action. *Educational Leadership*, 60(5), p.52.
- National Governor's Association Center for Best Practices (NGA Center) and the Council of Chief State School Officers (CCSS) (2010). *Common Core State Standards Initiative English Language Arts and Mathematics Curriculum*, Retrieved from http://www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf
- Olson, K.E. (2005, August). *Standards-based report cards: Accountability at the classroom level* (Doctoral dissertation, University of Southern California). UMI Number: 3196865.
- Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: a review of the literature. *Reading & Writing Quarterly*, 19, 139-158.
- Reeves, D. (2004). The Case Against the Zero. *Phi Delta Kappan*, 86(4), 324. Gale Document Number:A125847487.
- Salend, S. J. (2005, March/April). Report card models that support communication and differentiation of instruction. *Teaching Exceptional Children*, 37(4), 28-34.
- Schworm, P. (2006, February 16). Getting good grades is as easy as S-M-P (That's A-B-C to most of you). *Boston Globe*, pp. 1, 3.
- Silberglitt, B., Burns, M. K., Madyun, N. H., & Lail, K. E. (2006, May). Relationship of reading fluency assessment data with state accountability test scores: a longitudinal comparison of grade levels. *Psychology in the Schools*, 43(5), 527-535.

Stiggins, R. (1997). *Student-centered classroom assessment, 2nd edition*. Columbus, OH: Merrill.

Stiggins, R. (2005, December). From formative assessment to assessment FOR learning: a path to success in standards-based schools. *Phi Delta Kappan*, 87(4), 324. Gale Document Number A139834752.

Stiggins, R., & Chappuis, J. (2008). Enhancing student learning. *District Administration*, 44(1), 42–44.

Strobino, J., Gravitz, K. & Liddle, C. (2002, Spring) Student Perceptions of Grades: A Systems Perspective. *Academic Exchange Quarterly*, pp. 145 -150.

United States National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. (A report to the Nation and the Secretary of Education, United States Department of Education). Retrieved from www2.ed.gov/pubs/NatAtRisk/index.html. Govt. Doc. Number ED 1.2: N 21

Zirkel, Perry A. (2007) "Much Ado about a C?" *Phi Delta Kappan* 89(4), 318(2). Gale Document Number:A172516997.

Bibliography

- Briggs, K. (2010). Individual Achievement in an honors research community: teaching Vygotsky's zone of proximal development. *National Collegiate Honors Council*. Farmington Hills, Michigan: Gale Group.
- Chatterji, M. (2002, Autumn). Models and methods for examining standards-based reforms and accountability initiatives: Have the tools of inquiry answered pressing questions on improving schools? *Review of Educational Research*, 72(3), 345-386.
- Dewey, J. (1902). *The child and the curriculum*. Chicago, Illinois: University of Chicago Press.
- Dewey, J. (1915). *The school and society*. Chicago, Illinois: University of Chicago Press.
- Lilley, I. (1967). *Friedrich Froebel, A section from his writings by Irene M. Lilley*. Cambridge, MA: Cambridge at the University Press.
- Fullan, M. (2007). *The New Meaning of Educational Change*, New York, NY: Teachers College Press.
- Guskey, T. R. (1996). Reporting on student learning: Lessons from the past – Prescriptions for the future. In T. R. Guskey (Ed.) *Communicating Student Learning, 1996 ASCD Yearbook* (pp. 13-24). Alexandria, VA: Association for Supervision and Curriculum Development.
- Hargreaves, A., & Fullan, M. (2009) *Change wars*. Bloomington, IN: Solution Tree.
- Herbart, J. H. (1896). *The science of education, its general principles deduced from its aim and the aesthetic revelation of the world*. Boston, MA: D.C. Heath & Co., Publishers.

- Jung, L.A. & Guskey, T. R. (2007, April). *Determining Fair Grades for Students with Special Needs: A Standards-Based Model*. Online Submission; Paper presented at the Annual Meeting of the American Educational Research Association.
- Lave, J., & Wenger, E. (1991) *Situated learning: legitimate peripheral learning*. Boston, MA: Cambridge University Press.
- Marzano, R. (2007). *The art and science of teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- McGehee, J. J., & Griffith, L. K. (2001, Spring). Large-scale assessments combined with curriculum alignment: Agents of change. *Theory into Practice*, 40(2). 137-144.
- McMillan, J. H. (2009). Synthesis of issues and implications for practice. In T.R. Guskey (Ed.), *Practical solutions for serious problems in standards-based grading* (pp. 105-119). Thousand Oaks, CA: Corwin Press.
- Melograno, V. J. (2007, August). Grading and report cards for standards-based physical education: Methods of grading have become more complicated—and more appropriate! *The Journal of Physical Education, Recreation & Dance*, 78(6). 45-54.
- Muir, M. (2005, June 6). Standards based report cards [Research Brief]. *The Principals Partnership*, Retrieved from <http://www.principalspartnership.com/>
- O'Connor, K. (2009). *How to grade for learning K-12* (3rd ed.). Thousand Oaks, CA: Corwin Press.
- Reese, W. J. (2001). The origins of progressive education. *History of Education Quarterly*, 41(1), 1-24.

- Reeves, D. (2002). *Making standards work, 3rd edition: How to implement standards-based assessments in the classroom, school, and district*. Denver, CO: Advanced Learning Press.
- Reeves, D. (2009). Level-five networks: Making significant change in complex organizations. In Hargreaves, A. & Fullan, M. (Eds.), *Change wars* (pp. 237-258). Bloomington, IN: Solution Tree.
- Schmidt, R. (2008). *From report card to standards-based progress report: The teacher's change process* (Doctoral dissertation, Capella University). Retrieved from Dissertations & Theses: Full Text. Publication No. AAT 3320347.
- Scriffiny, P. L. (2008, October). Seven reasons for standards-based grading. *Educational Leadership*; 66(2), 70-74.
- United States Department of Education. (2006, March). *Federal Non-Regulatory Guidance on Designing Schoolwide Programs*. Retrieved from Massachusetts Department of Elementary and Secondary Education, <http://www.doe.mass.edu/titlei/programs/>.
- United States Department of Education. (2006, July 21). *LEA and school improvement non-regulatory guidance*. Retrieved from Office of Elementary and Secondary Education, <http://www.ed.gov/policy/elsec/guid/schoolimprovementguid.doc>
- United States Department of Education (2007). *Trends in international mathematics and science study*. Retrieved from <http://nces.ed.gov/timss/results07.asp>
- Zambo, D. (2009). Gifted students in the 21st century: Using Vygotsky's theory to meet their literacy and content area needs. *Gifted Education International*, 25, 270-280.
- Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992, Fall). Self-Motivation for

academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal*, 29(3), 663-676.

Appendix A

Cover Letter and Questionnaire

Date:

To:

Dear Principal,

I am a student in the Northeastern University Doctoral Program in Educational Leadership. I would greatly appreciate your participation in a study that I am conducting on the effect of standards-based report cards on student growth percentiles on the MCAS Mathematics Test. The study will specifically analyze the results for fourth grade students in the low income subgroup across southeastern Massachusetts. If you could take a few minutes to complete the brief questionnaire and return it along with a copy of your current report card for grade four, it would contribute immensely to my study and the resulting research. No individual school data will be reported in the study.

Please return the questionnaire and the report card copies by June 6. If you or your superintendent would like to receive a copy of the results of this research study please indicate the name and contact information of the interested party on the back of the questionnaire.

I have included a \$5.00 gift card to Dunkin Donuts as a token of my appreciation of your willingness to review this request. Please accept the gift card even if you choose not to participate in the study. Thank you very much for your time. If you have any questions or concerns, please contact me at 508-947-6117 or craig.t@husky.neu.edu.

Sincerely,

Theresa Craig
37 Stone St.
Middleborough, MA 02346
Advisor: Dr. Claire Jackson

District: _____ **School:** _____

- I.** Does your school use a standards-based report card? ___YES ___NO

NO - If you answered “NO”, please return this questionnaire along with **a copy of your 4th grade report card** in the attached self-addressed envelope. Please accept the gift card as a token of appreciation. **Thank you very much for your participation.**

YES - If you answered “YES”, please provide three more items of information. Your responses will contribute immensely to this study:

- II.** What grades in your elementary school have a standards-based report card? _____

- III.** Please provide the month and year that the standards-based report card was introduced in your elementary school. _____

- a. If the report card was phased in at different grade levels in different school years provide the years and corresponding grades below:

Please return this questionnaire along with a **copy* of your grade four report card** in the attached self-addressed envelope. Please keep the gift card as a token of appreciation.

Thank you very much for your participation.

**If your district requires the payment of a fee for copies of documents, please return the questionnaire with an invoice or contact me according to the cover letter. Payment will be sent promptly. Thank you, Theresa Craig*

Please return in the envelope provided to: Theresa Craig, 37 Stone St., Middleboro, MA 02346

Please send a copy of the results of this research study to:

School District: _____

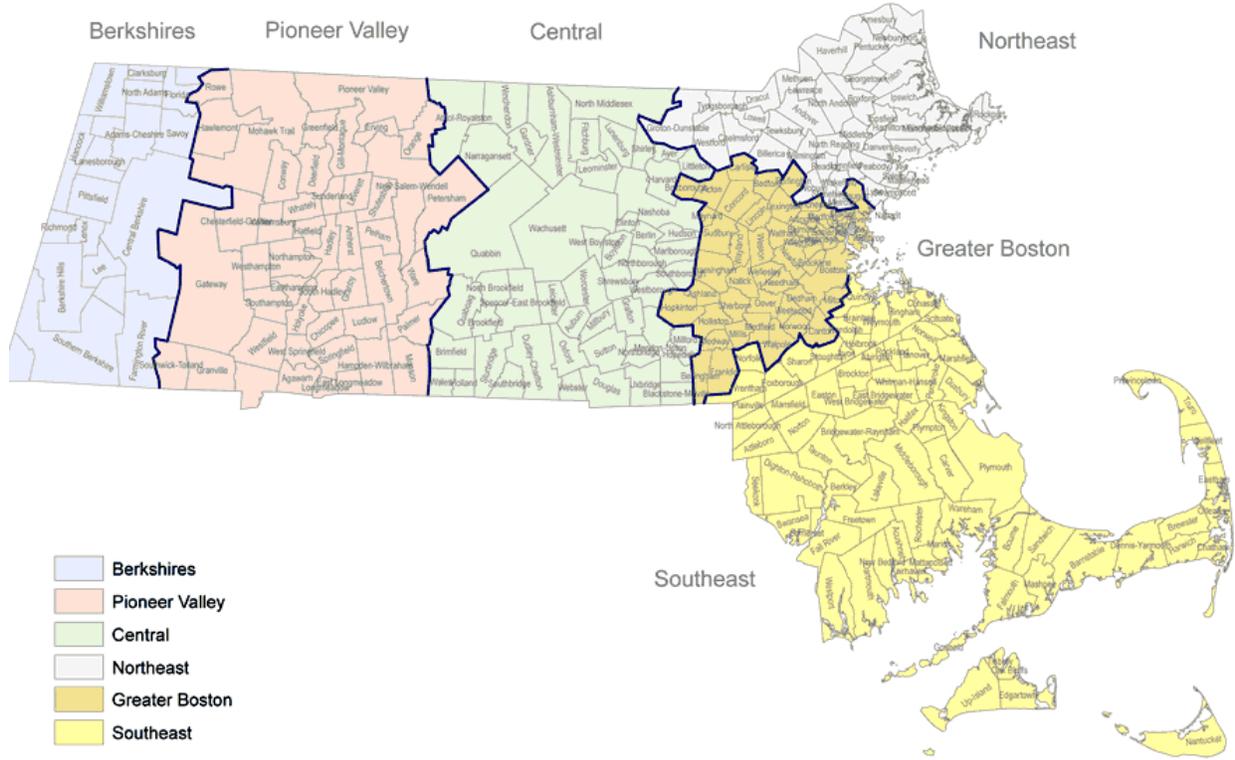
Address: _____

Attention: _____

The anticipated date for submission of this Doctoral Research Project is July 2011.

Appendix B

Map of 2010-11 District and School Assistance Centers Regions



Massachusetts Department of Elementary and Secondary Education

Last Updated: 7/9/2010

<http://www.doe.mass.edu/sda/regional/regions.html>

Appendix C

Standards-Based Report Card Product Rating Scale

	Rating			
	1	2	3	4
Performance Levels	1 overall grade or level, or more than 5 levels are used, to indicate progress; Language is ambiguous or unclear on levels	3 levels: Progress levels are limited to only 1 level below proficiency; Language requires more definition or uses comparative language such as below average or superior	4 or 5 levels: Levels indicate progress toward proficiency, and above proficiency; Parent friendly language	4 levels: 2 levels to indicate progress toward proficiency, 1 for advanced or excels; Parent friendly language
Failing Grades	Displays an F, zero or other failing designation	Displays indications of no growth or minimal growth in terms that may deter motivation	Language depicts a progression of learning that supports motivation	All language depicts a progression of learning that promotes high achievement as a possibility for all students
Separation of Learner and Social Behaviors	One amalgamated grade/level for student Learner and Social Behaviors	Learner and Social Behaviors are reported as comments on conduct and effort	Multiple Learner and Social Behaviors are reported separately from academics	Learner and Social Behaviors are reported within each content level separately from academics
Standards and indicators	One content area reported for mathematics	Mathematics reported as strands only, or as more than 12 indicators	Mathematics reported as 4 - 12 power standards that target critical focus areas for Grade 4	Mathematics reported as 7-12 power standards that feature action words and understandable indicators for students and families

Appendix D

Standards-Based Report Card Grouping Rubric

	Group 1	Group 2	Group 3	Group 4[^]
Criteria	Standards-based	Non-standards-based	Mixed*	Mixed 2 **
Performance levels	Score of 2 or better with required scores in criterion 2 and 4	Score of 1 or 2	Score of 2 or better with additional scores in criteria 2 and 4	Score of 1 or 2
Depiction of failure	Score of 2 or better	Score of 1	Score of 2 or better	Score of 1
Separation of learner and social behaviors	Not applied	Not applied	Not applied	Not applied
Grade specific standards and indicators	Score of 2 or better and depicts grading standards or indicators specific to Grade 4	Score of 1 or 2 with standards or indicators that are not specific to Grade 4	Score of 1 or 2 with standards or indicators that are not specific to Grade 4	Score of 2 or better with grading standards or indicators specific to Grade 4

*General standards reported but other criteria meet standards-based

**Specific standards reported for Grade 4 but other criteria do not meet standards-based

[^]Group 4 was eliminated from the study. There was only 1 entry in this group.

Appendix E

CPI Performance Rating and Range

Performance Rating	CPI Range
<i>Very High</i>	90 – 100
<i>High</i>	80 – 89.9
<i>Moderate</i>	70 – 79.9
<i>Low</i>	60 – 69.9
<i>Very Low</i>	40 – 59.9
<i>Critically Low</i>	0 – 39.9

(MA DESE, August 2006, p. 3)

Appendix F

Descriptive Statistics for Covariates Separated by Independent Variable Grouping

	N	Range	Minimum	Maximum	Mean
Standards-Based					
Low income percentage	52	86.90	2.90	89.80	34.4231
Special education percentage	52	20.00	7.20	27.20	15.2250
Limited English Proficient percentage	52	30.10	.00	30.10	3.1981
Student-teacher ratio	52	11.80	8.60	20.40	15.1154
Per pupil expenditure	52	11976.00	9700.00	21676.00	12153.1923
Valid N	52				
Non-standards-based					
Low income percentage	30	64.80	2.40	67.20	21.2200
Special education percentage	30	18.30	8.40	26.70	16.1100
Limited English Proficient percentage	30	12.30	.00	12.30	1.8800
Student-teacher ratio	30	7.80	12.50	20.30	15.4300
Per pupil expenditure	30	4845.00	9548.00	14393.00	11349.7333
Valid N	30				
Mixed					
Low income percentage	20	46.00	1.80	47.80	16.1485
Special education percentage	20	16.60	6.80	23.40	14.6700
Limited English Proficient percentage	20	10.60	.00	10.60	2.3250
Student-teacher ratio	20	13.70	7.50	21.20	15.1200
Per pupil expenditure	20	11923.00	9682.00	21605.00	11908.2500
Valid N	20				

Appendix G

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III		Mean Square	F	Sig.	Partial		
		Sum of Squares	Df				Eta Squared	Noncent. Parameter	Observed Power ^b
Corrected Model	2010 SGP	2032.268 ^a	7	290.324	1.350	.236	.096	9.453	.546
	2010 CPI	2871.999 ^c	7	410.286	11.419	.000	.473	79.930	1.000
Intercept	2010 SGP	2922.369	1	2922.369	13.594	.000	.133	13.594	.954
	2010 CPI	5189.134	1	5189.134	144.418	.000	.619	144.418	1.000
Low income	2010 SGP	17.888	1	17.888	.083	.774	.001	.083	.059
	2010 CPI	744.314	1	744.314	20.715	.000	.189	20.715	.994
Special Education	2010 SGP	314.423	1	314.423	1.463	.230	.016	1.463	.223
	2010 CPI	15.964	1	15.964	.444	.507	.005	.444	.101
Limited English	2010 SGP	416.326	1	416.326	1.937	.168	.021	1.937	.280
	2010 CPI	621.985	1	621.985	17.310	.000	.163	17.310	.984
Student-teacher	2010 SGP	802.326	1	802.326	3.732	.057	.040	3.732	.481
	2010 CPI	76.772	1	76.772	2.137	.147	.023	2.137	.304
Per pupil expenditure	2010 SGP	55.074	1	55.074	.256	.614	.003	.256	.079
	2010 CPI	46.995	1	46.995	1.308	.256	.014	1.308	.205
Report card type	2010 SGP	88.787	2	44.394	.207	.814	.005	.413	.081
	2010 CPI	49.159	2	24.580	.684	.507	.015	1.368	.162
Error	2010 SGP	19133.176	89	214.980					
	2010 CPI	3197.886	89	35.931					
Total	2010 SGP	270209.000	97						
	2010 CPI	658183.090	97						
Corrected Total	2010 SGP	21165.443	96						
	2010 CPI	6069.885	96						

a. R Squared = .096 (Adjusted R Squared = .025)

b. Computed using alpha = .05

c. R Squared = .473 (Adjusted R Squared = .432)

Appendix H

MANCOVA Results for the Independent Variable

Effect		Value	F	Hypothesis		Sig.	Partial	Noncent. Parameter	Observed Power ^b
				df	Error df		Eta Squared		
Intercept	Pillai's Trace	.650	81.718 ^a	2.000	88.000	.000	.650	163.436	1.000
	Wilks' Lambda	.350	81.718 ^a	2.000	88.000	.000	.650	163.436	1.000
	Hotelling's Trace	1.857	81.718 ^a	2.000	88.000	.000	.650	163.436	1.000
	Roy's Largest Root	1.857	81.718 ^a	2.000	88.000	.000	.650	163.436	1.000
Report card type	Pillai's Trace	.035	.788	4.000	178.000	.535	.017	3.151	.249
	Wilks' Lambda	.965	.784 ^a	4.000	176.000	.537	.017	3.134	.248
	Hotelling's Trace	.036	.779	4.000	174.000	.540	.018	3.117	.247
	Roy's Largest Root	.033	1.449 ^c	2.000	89.000	.240	.032	2.898	.303

a. Exact statistic

b. Computed using alpha = .05

c. The statistic is an upper bound on F that yields a lower bound on the significance level.

Appendix I

Univariate Test for the SGP Dependent Variable

	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^a
Contrast	81.272	2	40.636	.191	.826	.004	.382	.079
Error	19142.140	90	212.690					

The F tests the effect of report card type. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Computed using alpha = .05

Appendix J

Univariate Test for the CPI Dependent Variable

	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^a
Contrast	60.104	2	30.052	.724	.487	.016	1.449	.169
Error	3775.509	91	41.489					

The F tests the effect of report card type. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Computed using alpha = .05