It's More Than Money

Teacher Incentive Fund–Leadership for Educators' Advanced Performance Charlotte-Mecklenburg Schools

February 2013

COMMUNITY TRAINING AND ASSISTANCE CENTER

ABOUT CTAC:

The Community Training and Assistance Center is a national not-for-profit organization with a demonstrated record of success in the fields of education and community development. CTAC builds district, state, and community capacity by providing technical assistance, conducting research and evaluation, and informing public policy. It focuses on developing leadership, planning and managerial expertise within school systems, community-based organizations, collaborative partnerships, state and municipal governments, and health and human service agencies. Since 1979, CTAC has provided assistance to hundreds of community-based organizations, coalitions and public institutions in the United States and several other countries.

CTAC's staff is comprised of nationally recognized executives, educators, policy makers, researchers and organizers who have extensive experience working with city, county and state agencies, educational institutions, federal legislative bodies, not-for-profit organizations, philanthropic institutions and the private sector.

It's More Than Money

Teacher Incentive Fund–Leadership for Educators' Advanced Performance Charlotte-Mecklenburg Schools

February 2013

CTAC | COMMUNITY TRAINING AND ASSISTANCE CENTER

Acknowledgements

The Community Training and Assistance Center (CTAC) wishes to thank the Charlotte-Mecklenburg Schools (CMS) Board of Education and former School Superintendent, Peter Gorman for their commitment to bringing quality and focus to this evaluation and initiative. They have maintained critical attention to the development and implementation of the Teacher Incentive Fund-Leadership for Educators' Advanced Performance (TIF-LEAP).

CTAC would like to acknowledge the many people who have contributed both to the initiative and to making this study possible. The leadership of the Board of Education has been essential to the initiative. A special thank you goes to Trent Merchant, TIF-LEAP Board of Education Liaison to the initiative, for his commitment and dedication. CTAC would also like to thank the following board members for their leadership during the five years of the initiative: Eric Davis, George Dunlap, Ericka Ellis-Stewart, Larry Gauvreau, Ken Gjersten, Molly Griffin, Vilma Leake, Rhonda Lennon, Mary McCray, Richard McElrath, Kaye McGarry, Kimberly Mitchell-Walker, Timothy Morgan, James Ross, Amelia Stinson-Wesley, Tom Tate, Joyce Waddell, and Joe White.

We have been assisted by many current and former members of the leadership team of CMS, including Maurice Ambler, Charity Bell, Nora Carr, Ann Clark, Maurice Green, Daniel Habrat, Hugh Hattabaugh, LaTarzja Henry, Susan Johnson, Scott Muri, Ruth Perez, Jonathan Raymond, Tyler Ream, Sheila Shirley, and Denise Watts. Thank you for being willing to share this professional challenge with CTAC. Your time and effort have been appreciated.

The CTAC evaluation provides a solid contribution to the ideas and desires of those who wish to transform teaching and learning in public education. We wish to thank those individuals who, during these past five years, have contributed so much to this effort, including Curtis Carroll, Christopher Cobitz, Dennis Covington, Shelton Jefferies, Carrie Miranda, Mike Raible, Vincent Smith, and Barbara Ann Temple.

The coordination and implementation of the many facets of this evaluation could not have been conducted without the assistance of the TIF-LEAP team: Susan Kirkendol, Jan Millard, Susan Norwood, Scott Reynolds, Theresa Shipman, and Carol Turner. Their persistence, expertise, and dedication have been exemplary.

We would like to extend a very special thank you to the principals, assistant principals, teachers and other staff at the following TIF-LEAP schools: Berryhill ES, Billingsville ES, Druid Hills ES, Highland Renaissance ES, Lincoln Heights ES, Reid Park ES, Shamrock Gardens ES, Bishop Spaugh MS, John Taylor Williams MS, J.W. Wilson MS, Martin Luther King, Jr. MS, Sedgefield MS, E.E. Waddell HS, Business and Finance @ Garinger HS, International Studies @ Garinger HS, Leadership and Public Service @ Garinger HS, Math and Science @ Garinger HS, New Technology @ Garinger HS, West Charlotte HS, and West Mecklenburg HS. They have made special efforts on behalf of students.

We have been supported by many current and former members of the Steering Committee and four TIF-LEAP working groups: Human Resources/Finance, Principals' Working Group, Instructional Support, and SLO Design Team. We would like to give special recognition and appreciation to the following committee members for their generous time and attention: Darius Adamson, Andrew Baxter, Christina Efird, Randolph Frierson, Steve Oreskovic, Paul Pratt, and Julie Stafford.

CTAC is grateful for all those listed above, and the many other people who helped to make this study a reality.

This evaluation is made possible through a Teacher Incentive Fund grant from the U.S. Department of Education.

The findings, analyses and recommendations expressed in this study are those of the Community Training and Assistance Center.

Credits

This study was conducted and prepared by the Community Training and Assistance Center of Boston, Massachusetts.

Principal Study Authors
William J. Slotnik
Maribeth D. Smith

Contributing Authors
Barbara J. Helms, Ph.D.
Zhaogang Qiao, Ph.D.

Field Operations Coordinators
Geraldine Harge, Ed.D.
Richard Larrabee

Study Team Members
Peggie L. Brown
Judith A. Clary, Ph.D.
Ileana Cintrón
William Eglinton
Elizabeth Gutierrez
Jiefang Hu
Elizabeth Larrabee
Cathi Leone
Guodong Liang, Ph.D.
Natalie Nier
Camelia Rosca, Ph.D.
Sylvia Saavedra-Keber
Terrence Lee-St. John

Chapter Reviewers
Donald Gratz, Ph.D.
James Wyckoff, Ph.D.

Contents

4
Executive Summary

10
Preface

14
CHAPTER I
The Beginnings and Governance of the TIF-LEAP Initiative

28
CHAPTER II
TIF-LEAP Performance-Based Compensation Approaches: Design, Development, and Implementation

42
CHAPTER III
Evaluation Design and Methods of Analysis

54
CHAPTER IV
Student Learning Objectives: Quality and Attainment

62
CHAPTER V
Impact of TIF-LEAP on Student Achievement

80
CHAPTER VI
Perspectives on the TIF-LEAP Initiative and Performance-Based Compensation

98
CHAPTER VII
National Implications for Performance-Based Systems

Executive Summary

Teacher Incentive Fund–Leadership for Educators' Advanced Performance (TIF-LEAP), a multi-year performance-based compensation initiative in Charlotte-Mecklenburg Schools (CMS), focused on improving teaching and learning in a select group of high need schools. Benefiting from a community culture that supports using monetary incentives to encourage and reward employee performance, the district partnered with the Community Training and Assistance Center (CTAC) in 2007 to seek, obtain and implement a Teacher Incentive Fund (TIF) grant from the U.S. Department of Education.

The TIF-LEAP initiative introduced two approaches to performance-based compensation—Student Learning Objectives (SLOs) and a value-added measure (VAM)—into a total of twenty schools over the course of implementation. The initiative established a structure through which teachers and principals earned bonuses for demonstrated increases in student academic growth.

Already complex and multi-layered, the initiative was seriously affected in years four and five as the recession roiled the district budget and accordingly district schools. CMS maintained its matching fiscal commitment to the TIF-LEAP initiative, but planning for and implementing teacher layoffs, principal changes, school closures, and school reorganizations influenced outcomes in the final years of implementation.

After a peak performance year (2009-10), with all twenty schools phased in and both approaches implemented, the following year (2010-11) turned into what evaluators came to call "the perfect storm." Misunderstandings and misgivings about the VAM rankings among many participants and the implementation of a new teacher appraisal system started the year, while layoff notices and plans for closing and reorganizing selected schools dominated the spring. The TIF-LEAP schools experienced significant turnover in principals, and the superintendent left later in the spring. At the beginning of year five, the overall number of participating schools was reduced to eleven. The TIF-LEAP initiative concluded with the 2011-12 academic year.

It's More Than Money is the evaluation of the initiative, based on five years of observations, annual stakeholder surveys and interviews, and analyses of SLO artifacts and student achievement results. As the title of the report indicates, improving teaching and learning through performance-based compensation is an enterprise that does not run on the promise of monetary incentives alone. Success depends on more than money.

Promising Results

Student Academic Growth on North Carolina Assessments

The descriptive statistical analysis shows increases in student achievement attributable to the TIF-LEAP initiative. This analysis examines the North Carolina End-of-Grade student results from TIF-LEAP schools together with those of comparison schools:

- The growth rate of students in TIF-LEAP schools is greater than that of students in the comparison schools. Although the TIF-LEAP schools start with lower student performance, by the end of year four, the student test scores in both mathematics and reading are closely approaching those of the comparison schools.
- The TIF-LEAP schools show greater resilience to the negative shocks resulting from the economic recession, including teacher layoffs, and planning for school closures and restructuring that occurred in 2010-11. Student test scores in the TIF-LEAP schools grew at a lower rate in that year than in the previous school year. However, they grew at a higher rate than the comparison schools that experienced the same disruptions.

The longitudinal hierarchical linear models provide the estimated effects of the TIF-LEAP initiative on student achievement. They show that TIF-LEAP had a positive impact on the participating schools which is both statistically and practically significant. Specifically,

- In terms of mathematics achievement, students in TIF-LEAP schools on average have a growth rate 12% greater than students in the comparison schools.

  This growth difference is substantial and means that the TIF-LEAP students are growing 12% more than the 0.8% annual growth rate of the comparison schools. This growth translates into 0.34 points annual growth difference between TIF-LEAP and comparison students. As a result, at the end of year four of the initiative, the test scores of students in the TIF-LEAP schools improved, cumulatively, 1.4 points more than students in the comparison schools. This growth brings the TIF-LEAP schools close to par with the comparison schools (students in TIF-LEAP schools started 1.5 points lower than students in the comparison schools at the beginning of the initiative).

- In terms of reading achievement, students in TIF-LEAP schools on average have a growth rate 13% greater than students in the comparison schools.

  This growth difference is substantial and translates into 0.44 points annual growth difference between TIF-LEAP and comparison students. As a result, at the end of year four of the initiative, the test scores of students in the TIF-LEAP schools are only 0.7 points lower than those in the comparison schools. The initial test scores of the TIF-LEAP students started 2.5 points lower than students in the comparison schools.

Three cross-sectional HLM analyses were conducted over the course of the TIF-LEAP initiative. The findings of the cross-sectional HLM models vary by subject and year. The first cross-sectional analysis is for 2008-09, the first year of SLO implementation. The full SLO effects on student achievement were expected to phase in over several years of implementation. The findings in the first year support this expectation:

- There are positive, statistically significant associations between the attainment of Target SLOs and student achievement both in mathematics and in reading.
- There is no statistically significant association between the quality of SLOs (as indicated by the rubric rating) and student achievement in this first year.

The second cross-sectional analysis is for 2009-10. In terms of achieving higher student performance, this is the peak year of SLO implementation. The key findings are:

- There are positive, statistically significant associations between the quality of SLOs and student achievement. This finding means that a teacher's SLO rating relates positively to student achievement in elementary school mathematics, elementary school reading, and middle school mathematics.
- There are positive, statistically significant associations between the attainment of SLOs and student achievement at the elementary school level. This finding means that the students whose teachers met their SLOs achieved higher scores in elementary school mathematics and reading.

The third cross-sectional analysis is for 2010-11. In this school year, as a result of the increase in the number of students and classrooms, the investigation is conducted at the individual grade level in grades 4-8 rather than combining grades into elementary and middle school analyses. The key findings are:

- There is a positive, statistically significant association between the quality of SLOs and student achievement in mathematics in grade five.
- There is a positive, statistically significant association between the attainment of SLOs and student achievement in reading in grade six.

Student Learning Objectives

The district began implementing Student Learning Objectives in year two of the initiative as an approach to measuring the connection between teacher performance and student achievement. Specifically, SLOs were implemented in the first ten schools in 2008-09 and in ten additional schools which joined the initiative in 2009-10.

Nearly 4,000 teacher-developed SLOs were evaluated for this study, using CTAC's four-level rubric that examined content, expectations, completeness, and coherence. The study also examined (1) whether or not the growth targets set by teachers for their SLOs were attained (met or not met) and a bonus collected, and (2) the relationship between the quality of each SLO and the attainment of the growth target. The analysis showed:

- The overall relationship between the quality of SLOs and their attainment is positive. Year-by-year findings vary with the highest correlation found in 2009-10. The relationship between the quality of an SLO and its attainment (meeting or exceeding the growth target set by the teacher) is statistically significant. It shows that the higher the quality of the SLO, the greater the likelihood it will be attained.

Further, the number of years a teacher participates in SLO implementation matters.

- Teachers in the initiative for three years of SLO implementation develop higher quality SLOs and have greater success in attaining their SLOs. The relationships between the quality of an SLO and its attainment to the teacher's length of time in the initiative are statistically significant.

Setting and reaching SLO targets was a new practice for teachers and principals, and the learning curve for both participants and initiative staff is evident as teachers and TIF-LEAP staff gained experience between the first and second years of SLO implementation. The second year of SLO implementation (2009-10) is the strongest year overall for SLO-related performance, in particular, as well as for the fully implemented initiative, in general. The analysis of the third year of SLO implementation (year four of the initiative) finds decreases in some areas, as many teachers and schools were affected by the fiscal crisis which led to layoffs and school closures.

Survey and interview responses show:

- The Curriculum and Instruction Department identified SLOs as a district instructional best practice. Numerous teachers in the TIF-LEAP schools declared an intention to continue the process after the conclusion of the initiative.
- Beyond the bonus payouts, teachers valued the data analysis, planning, and instructional elements of the SLO process. These elements resonated with teachers as being significant to teaching students more effectively and to advancing their professional growth.
- The SLO process provided teachers and principals with the tools to look carefully at their students through the lens of more timely and better baseline data. They observe that they had greater capacity to analyze each student's progress and set growth targets that stretched and encouraged every student.
- The SLO baseline data analysis—prior to setting targets and planning instruction—prompted more in-depth analysis about the best instructional strategies to meet student needs.
- The new North Carolina Teacher Evaluation Process, implemented in year four of the initiative, reinforces the TIF-LEAP work with SLOs, according to teachers, principals, and the TIF-LEAP team.

Value-Added Measure

CMS leadership introduced a district-developed VAM in 2009-10. The VAM applied to teachers and administrators in the TIF-LEAP schools at the end of 2009-10. It was continued in 2010-11 and 2011-12.

Like SLOs, the district-developed VAM was piloted as a measurement of the link between teacher performance and student achievement. In TIF-LEAP, only teachers in tested grades and subject areas were eligible to receive an individual VAM bonus; that is, teachers whose students participated in the North Carolina EOG/EOC assessments. In order to receive the VAM bonus, those eligible teachers had to have (1) a VAM score at or above the 70th percentile (be in the top 30% of teachers in the district); and (2) as with SLOs, a rating of proficient or above on their performance evaluation.

- In 2009-10, 318 of the 875 teachers participating in TIF-LEAP taught in a state-tested grade or subject area. Of these eligible teachers, 34.7% (119 out of 318) received a VAM bonus.
- In 2010-11, 240 of the 733 teachers participating in TIF-LEAP taught in a state-tested grade or subject area. Of these eligible teachers, 36.6% (88 out of 240) received a VAM bonus.

In terms of the relationship between the VAM and SLOs, the analysis shows:

- Teachers in TIF-LEAP schools who received a VAM bonus are more likely to have high quality SLOs. This relationship applies to 2009-10 and 2010-11. It is statistically significant in 2009-10.

Related findings from the analysis of interviews and survey data include:

- The VAM rankings, first distributed by the Accountability Department at the close of 2009-10, became public before teachers understood what the rankings meant and how the use of the VAM information would impact them. When the VAM rankings were announced, it became apparent that many teachers had misunderstandings and/or misgivings about the accuracy and use of the value-added approach to measure teacher performance.
- Despite their perceptions of a lack of transparency surrounding the introduction and use of the VAM, most TIF-LEAP teachers and principals are more positive about the use of student academic growth as performance measures than teachers in other district schools.
- Principals require more professional development on the technical aspects of value-added measures so they can assist with messaging and addressing the concerns of teachers.

Broader Institutional Findings

In designing and implementing the TIF-LEAP initiative, CMS became benefactor and beneficiary. The district role as *benefactor* included: entering into and committing financially and organizationally to the TIF grant, designating leadership and professional staff, and implementing a senior level governance structure.

CMS became a *beneficiary* in several ways. Students benefited as evidenced by the promising results in student achievement, even in difficult times. Other benefits also accrued to CMS as a learning organization.

Key findings include:

- TIF-LEAP teachers and principals accepted accountability for student learning results, both at the school and classroom levels. When districts or states seek to measure educator effectiveness, the leading and sometimes the only story is about teacher and principal pushback against the use of student assessment results to evaluate or compensate performance. However, survey and interview data show that TIF-LEAP teachers and principals maintained a measured openness to performance-based systems throughout the initiative.

- Having a teaching staff with a high degree of professional accountability for their own performance is a substantial asset, but capitalizing on it demands a requisite investment from the district. CMS leaders perceived the potential of performance-based compensation as a lever for systemic reform. They mobilized staff and resources from a variety of departments to help bring the initiative quickly to life.
- The approach or vehicle selected to measure teacher performance, whether for compensation, evaluation, or employment, matters. There are fundamental differences among approaches, including who is eligible to participate. Implementing multiple approaches can provide more nuanced and defensible information, but requires a thoughtful design that integrates the approaches.
- Looking at students through the lens of student growth stimulates different thinking about student progress than does the use of the proficiency lens alone. Leaders indicate TIF-LEAP served as a catalyst in moving the district to having an institutional emphasis on student growth.
- A more systematic and evidence-based approach to developing, evaluating, compensating, and placing effective teachers requires the availability of a teacher database that shows how teachers move within the district and why. This is a district challenge.
- The ongoing management of priorities in a district and schools that are launching new approaches to measuring teacher performance is of high importance. With a large number of new programs on the agenda, competing priorities within the district and their impact on the TIF-LEAP schools were a recurring concern for key CMS leaders—Board members and central administrators. Their concern was referenced frequently in interviews. On the one hand, competing priorities in a large district is a given; but on the other hand, competing priorities affected the overall narrative of the TIF-LEAP initiative as well as the outcomes.
- The first rule of communications is "message times one thousand." The importance of this rule became pronounced when the district experienced difficulties in communicating effectively about the VAM and was undergoing budgetary cutbacks, school closures, and related changes.
- On-site, customized professional development—supported with timely follow-up and troubleshooting—proved an effective delivery mode for building new skills, such as developing SLOs. Teachers and principals continually praised and advocated for the use of the TIF-LEAP model for all professional development.

National Implications

The TIF-LEAP experience contributes to a broader, research-based, and practical understanding of what is required to effectively implement performance-based systems for purposes of compensation and evaluation. Specifically, the national implications drawn from the TIF-LEAP initiative are highlighted below.

- An effective performance-based system requires a dual emphasis on support and accountability.
- SLOs provide a measure of student growth and a measure of teacher practice—and quality matters with both.
- The whole process counts when implementing Student Learning Objectives.
- Learning Content and Instructional Strategies are key to effective SLOs.
- Effective SLO implementation requires distinguishing between training, professional development, and leadership development.
- The four pivotal considerations when introducing a value-added measure are role, understanding, fairness, and application.
- Performance-based systems must meet three standards of validity—statistical, educational, and political.

Summary

Increasing student achievement means identifying and fostering an outstanding teacher for every classroom. Supported by federal financial incentives, states and districts are taking up the challenge of connecting compensation and evaluation to effective teaching. In so doing, they are catalyzing a national movement towards performance-based reforms.

Charlotte-Mecklenburg Schools is an accomplished district. In its laboratory of performance-based approaches, the district experienced what worked and what did not. In navigating through a fiscal crisis, the district managed to keep the initiative going. As a result, the initiative benefitted students, teachers and administrators—thereby demonstrating that, when it comes to effective performance-based systems, more is involved than money alone.

Preface

Even through two decades of standards-based reform and accountability measures for schools and districts, there have been frequent reminders that when it comes to learning, one should never forget that it is the teacher that matters most. Indeed, "variations among teachers dominate school quality differences," overwhelming all other school inputs.[1] More recently, a study of a large number of students conducted over a longer period of time and in greater depth found elementary and middle school teachers to have a wide-ranging and long-lasting positive impact on students, even beyond academic success.[2]

The ascendant status of highly effective teaching—teaching that gets results—in the hierarchy of classroom learning variables is widely acknowledged. Who does not admire great teaching? Or have a story about an influential teacher? Yet, until recently, state and district policies and practices that recognize, assess, and systematically strengthen and reward effective teaching have been the exception. Traditionally, neither teacher performance appraisal systems nor compensation schedules have differentiated among teachers based on the outcomes of their students; and trustworthy assessments to measure the teacher effectiveness: student achievement link across the spectrum of grade levels and subjects remain scarce to non-existent.

New federal programs, such as the Teacher Incentive Fund and Race to the Top, have focused the reform spotlight on linking student results to teacher evaluation and compensation. How can states and districts improve their teacher performance appraisal and compensation systems to insure that all children have great teachers? How can teacher effectiveness be measured? How should outstanding teaching be recognized and rewarded?

Improving student learning by providing awards for effective teaching has been the focus of a five-year initiative in Charlotte-Mecklenburg Schools (CMS). A county district in North Carolina serving more than 135,000 students, CMS is well regarded nationally for its resolute commitment to educational and organizational excellence and was awarded the Broad Prize in 2011.[3] Seeking to encourage the best teaching in its highest need schools, the district partnered with the Community Training and Assistance Center (CTAC) in 2007 to create and implement a performance-based compensation initiative in 20 schools in the district through the Teacher Incentive Fund.

It's More Than Money is the evaluation of performance-based compensation in twenty schools in CMS. It examines the structure of the initiative, including the role of technical assistance and district support. It analyzes two different approaches to measuring teacher performance for the purpose of improved student learning and to awarding extra compensation: (1) Student Learning Objectives (SLOs), an instruction-based approach; and (2) a district-developed value-added measure (VAM), an assessment-based approach. Moreover, the evaluation provides a comprehensive perspective on the initiative based on five years of implementation.

Examining the genesis, development, and implementation of the SLO approach together with the incorporation of the VAM approach provides a study in the opportunities, complexities, benefits, and challenges that arise in measuring and compensating the impact of teacher effectiveness on student growth. It is at once an inspirational chronicle that draws one into the promise of performance-based compensation and a cautionary tale that alerts any district starting down this path that *significantly more is at stake than money alone.*[4]

The SLO Approach to Performance-Based Compensation

Effective teaching was very much on the minds of the members of the Board of the Denver Public Schools and the Denver Classroom Teachers Association in the fall of 1999 when they agreed during contract negotiations to pilot an approach to performance pay for teachers who demonstrated an impact on student learning. Together with the Community Training and Assistance Center, teacher and administrative leaders developed and refined a design for performance-based compensation and conducted an evaluation of how it worked. The outcome of the study and the genesis of ProComp, a fundamentally reformed teacher salary schedule, are both well documented in *Catalyst for Change: Pay for Performance in Denver* (2004)[5] and *Pay-for-Performance Teacher Compensation: An Inside View of Denver's ProComp Plan* (2007).[6]

The ground-breaking design that Denver implemented as the basis of extra compensation and later included as a component in ProComp has come to be known as Student Learning Objectives (SLOs) and has been adapted for use in performance-based compensation systems in Austin, Texas as well as CMS. More recently, SLOs increasingly are being chosen for use in teacher evaluation systems in numerous states and districts throughout the nation.

Using the foundation established in Denver, the TIF-LEAP initiative in Charlotte-Mecklenburg extended and strengthened the scaffolding of best teaching practice around SLOs, particularly in the area of direct, school-based professional development and technical support for teachers. As observed by TIF-LEAP staff members, teacher and principal participants, and district leaders, the SLO process expanded the concept and role of baseline data in planning for results and demonstrated a method by which teachers can plan, teach, and assess with greater precision and more science. Further, the SLO process ignited a redefinition of student achievement for the entire district, from a measure on a proficiency scale to a measure of student growth based on the comparison of individual student starting and ending points. Finally, the SLO approach smoothed the way in the TIF-LEAP schools for the introduction of the new North Carolina Teacher Evaluation Process, with its emphasis on collecting evidence of teaching effectiveness.

The Value-Added Approach to Performance-Based Compensation

Value-added refers to a statistical methodology[7] that accounts for prior student achievement in estimating the teacher's contribution to a student's learning—the value-added. Analyses of this type are constructed with multi-tiered statistical models that account for individual student factors and other classroom and school covariates known to influence student learning, such as socioeconomic status, class size, and attendance, in order to isolate the teacher's contribution. The availability of these sophisticated models and the access to better data enable districts and researchers to isolate the effect of the teacher from myriad other influences on student learning. It is an attractive approach to many who are responsible for evaluating and compensating teachers but also has detractors among both educators and statisticians who question its suitability for measuring individuals.[8]

The individual value-added measure in Charlotte-Mecklenburg was reported as a percentile ranking assigned to teachers (of assessed grades and subjects) and schools in the initiative based on a multi-level model. It was developed by the district's Accountability Department staff using student performance from the North Carolina End-of-Grade/End-of-Course assessments (EOG/EOC) and other selected covariates believed to be relevant to student learning in the district. The VAM approach did not require changes in teacher classroom practice as did SLOs, but the methodology of the VAM, as well as the teacher rankings, did prove to be unexpectedly complicated to communicate purposefully to participants.

Learning from Charlotte-Mecklenburg Schools

The story of implementing Student Learning Objectives in CMS is compelling. It demonstrates that when teachers have better information about their students earlier in the year, and a heuristic for using that data for better decisions, they can more effectively plan for, teach, and meet student growth targets. It shows that the investment in the implementation of new initiatives pays off when district leadership can create and sustain systemic conditions that engender greater fidelity to the design and more stability in the implementation of an initiative.

Further, by implementing two different approaches—the instruction-based SLOs and the assessment-based VAM—the TIF-LEAP initiative "provided a laboratory," according to the thinking of one district official, in the effects of teacher and principal performance bonuses and their promise for improving learning. The fact that the two approaches to performance-based compensation were piloted in the same schools in overlapping time intervals became a significant complicating factor in understanding their impact, as this evaluation will show. That does not mean that multiple approaches to compensation awards

cannot be effectively integrated, as they are today in Denver's ProComp, especially if planned co-terminously for that purpose.

New initiatives confront obstacles. How effectively participants and leaders work through the challenges to stay the course is one indicator of an initiative's resilience and potential for success. The economic recession that occurred during implementation of the TIF-LEAP initiative presented an atypical set of implementation obstacles. Maintaining district finances, unfortunately, dictated layoffs, staffing changes, and eventually school closures and reorganizations. These all impacted the spirit and shape of the initiative, particularly in the fourth and fifth years. However, the professional manner in which teachers, principals, project staff, and district leadership maintained a commitment to the initiative is a lesson for states and districts struggling to move forward during difficult times.

Ongoing Developments

Most teacher evaluation systems have not traditionally assessed individual teachers' contribution to student achievement, often for lack of effective and inclusive assessment instruments and well-integrated student and teacher data systems. As states and districts—stimulated by federal incentives—seek better methods of teacher performance appraisal with the goal of placing an outstanding teacher in every classroom, the Student Learning Objectives process, which originated as a performance-based compensation model, is being adapted for use as the student growth component of annual teacher performance reviews.

Recently, state education departments, including North Carolina, school districts, national teacher centers, and teacher organizations involved in implementing Race to the Top (RTTT) projects around the country, are joining the U.S. Department of Education in investigating and advancing Student Learning Objectives as one measure of teacher effectiveness. A major advantage of SLOs as a measure of teacher performance is their adaptability to all grades and all subjects. More importantly, SLOs originate in the instructional milieu of the classroom that teachers understand well and over which they believe they have the greatest influence. However, SLOs still require the thoughtful and strategic involvement of other district systems and resources in order to provide effective principal leadership, focused professional development, reliable student assessments, and timely technology support.

Value-added approaches, when used for compensation and evaluation purposes, are under scrutiny from (1) some researchers and statisticians/analysts[9] who point to too much variability in the VAM ratings and too many inconsistencies and inaccuracies for use as the only measure in high stakes decisions about individual teachers; and (2) teacher organizations, which promote access to compensation opportunities and fair appraisals for all of their membership. These concerns notwithstanding, value-added approaches are playing an important role in research about teacher effectiveness and other education topics and in the development of policies and practices that promote and support teacher effectiveness.

Evaluating the Impact of the TIF-LEAP Initiative

The Community Training and Assistance Center (CTAC) conducted the evaluation of performance-based compensation in CMS by examining the quality of the implementation and the impact of the TIF-LEAP initiative on teacher effectiveness and student achievement in the participant schools, using a variety of data and analyses over the life of the initiative. In addition, CTAC sought and studied the ideas and perspectives of TIF-LEAP participants, district and initiative leaders, students, and parents and community members on the progress, quality, and significance of the TIF-LEAP initiative in CMS.

This evaluation provides information about the inner workings of a many-faceted initiative and its impact on teacher performance and student achievement, all of which were identified and analyzed through the large volume of data collected over time from a variety of sources.

The nuts and bolts of the TIF-LEAP initiative are found in Chapter I, which describes the goals of the Teacher Incentive Fund and the priorities and practices of CMS with respect to incentive-based reform. It follows the actions of district leaders and CTAC in establishing a governing structure to guide and maintain the initiative; to select and support the participant

schools; and to staff and implement the project. It introduces the work of the instruction-centered SLO Design Team, which with CTAC's counsel set up the design parameters of Student Learning Objectives as the basis of the first performance-based compensation approach. Finally, it describes the inclusion of a value-added measure as a second approach through which eligible participants could earn bonuses. This chapter also includes a table which provides an overview of the initiative.

The work of the TIF-LEAP initiative provides a case study in the introduction of two different performance-based compensation approaches into multiple school settings for the purpose of measuring and rewarding effective teaching. Chapter II outlines the theoretical underpinnings, design, and implementation of each approach. It also provides a picture of the work of the TIF-LEAP team introducing a process that involved changes in both teacher and principal/supervisor practice. Finally, this chapter provides a comparison of the implementation features of the two approaches using principles of compensation reform.

The design of the evaluation is the subject of Chapter III. The discussion includes a description of the evaluation questions, the method of selecting comparison (control) schools, the data collected over the life of the initiative, the limitations presented by the data and how they were addressed, and the tools of analysis.

The next two chapters examine the relationship between teacher performance and student learning. The quality and attainment of Student Learning Objectives developed by teachers in TIF-LEAP schools is the subject of Chapter IV, together with a review of the relationship between VAM bonuses and the quality of SLOs. Chapter V investigates the impact of the TIF-LEAP initiative on student growth over three years on the North Carolina EOG tests.

As part of the evaluation of the TIF-LEAP initiative, the Community Training and Assistance Center annually surveyed teachers and school-based administrators and conducted interviews with district leaders and samples of TIF-LEAP and non TIF-LEAP teachers and principals. Additionally, questions about linking teacher performance and pay were included on the district's annual phone survey of parents and community members. Chapter VI provides an overview of the perspectives gathered from these groups and shows how opinions were influenced over time by ongoing developments in the initiative and changes in district circumstances.

Finally, Chapter VII draws on TIF-LEAP's accomplishments and presents their implications in the broader national school reform context.

Endnotes

1 Hanushek, E., Kain, J., & Rivkin, S. (1998). Teachers, schools, and academic achievement. *National Bureau of Economic Research* (Working Paper w6691). Retrieved from http://www.nber.org.

2 Chetty, R., Friedman, J., & Rockoff, J. (2011). Long-term impacts of teachers: Teacher value-added and student outcomes in adulthood. *National Bureau of Economic Research*. (Working Paper 17699). Retrieved from http://www.nber.org.

3 The 2011 Broad Prize Winner. Retrieved from http://www.broadprize.org.

4 Slotnik, W. (2009). *It's more than money: Making performance-based compensation work*. Center for American Progress at http://www.americanprogress.org/issues/2009/07/more_than_money.html.

5 Slotnik, W., Smith, M. *et al.* (2004). *Catalyst for change: Pay for performance in Denver*. Boston, MA: Community Training and Assistance Center.

6 Gonring, P., Teske, P., & Jupp, B. (2007). *Pay-for-performance teacher compensation: An inside view of Denver's ProComp Plan*. Cambridge, MA: Harvard Press.

7 Sanders, W. and Horn, S. (1994). The Tennessee value-added assessment system (TVAAS): Mixed model methodology in educational system. *Journal of Personnel Evaluation in Education*, *8,* 299-311.

8 Baker, E. *et al.* (2010, August 27). Problems with the use of student test scores to evaluate teachers. *Economic Policy Institute.* (Briefing paper 278). Retrieved from http://www.epi.org.

9 Baker, E. *et al.* (2010). See also Braun, H. I. (2005). *Using student progress to evaluate teachers: A primer on value-added models.* Retrieved from Policy Information Center, Educational Testing Service, Princeton, NJ. Website: http://www.ets.org/research/policy_research_reports/publications/report/2005/cxje and Darling-Hammond, L. *et al.* (2012). Evaluating teacher evaluation. *Phi Delta Kappan, 93*(6), 8-15.

CHAPTER I

The Beginnings and Governance of the TIF-LEAP Initiative

In 2007, Charlotte-Mecklenburg Schools (CMS) and the Community Training and Assistance Center (CTAC) sought and were awarded a Teacher Incentive Fund (TIF)[1] grant in the second cohort of districts and agencies approved by the U.S. Department of Education. The TIF initiative for CMS, *Leadership for Educators' Advanced Performance (LEAP),* got underway in the 2007-08 academic year and concluded with the 2011-12 academic year. Over the five years of the grant, an allocation of $11,880,267 from the Teacher Incentive Fund was augmented with $8,641,327 in local contributions for a project total of $20,521,594. These funds were largely targeted to performance bonuses for teachers and principals based on evidence of student growth.

For Charlotte-Mecklenburg Schools, the 19th largest district in the country in 2008-09,[2] attracting and retaining highly effective teachers in high need, difficult-to-staff schools is a serious priority. In 2006-07, nearly 70% of the district's 161 schools qualified as high need; 65% of schools did not meet Adequate Yearly Progress (AYP) goals, as required by No Child Left Behind (NCLB); and 36% of students at the high school level performed below proficient on state tests. The year prior (2005-06), CMS had been identified for NCLB District Improvement. Thus, at the time of the TIF proposal, CMS was developing a foundation, in concert with other educational initiatives, that could help engender and nourish outstanding teaching and inspire higher levels of learning in its highest need schools.

The Teacher Incentive Fund

Connecting Teacher Performance and Compensation

Created in 2006 by Congress during President George W. Bush's administration as part of an appropriations bill, the Teacher Incentive Fund was expanded and supported in 2009 through the American Recovery and Reinvestment Act (ARRA) during President Barack Obama's administration. Integral to TIF is an assumption that performance-based compensation can be linked to better teaching and better outcomes for high need students. TIF's stated goals include: (1) improving student achievement by increasing teacher and principal effectiveness; (2) reforming teacher and principal compensation systems so that teachers and principals are rewarded for increases in student achievement; (3) increasing the number of effective teachers teaching poor, minority, and disadvantaged students; and (4) creating sustainable performance-based compensation systems.[3]

The Teacher Incentive Fund provides a federal funding stream dedicated to reforming the single salary schedule, a well-recognized traditional compensation system that remunerates teachers and principals according to length of service, or longevity, and attainment of educational credits beyond the baccalaureate degree, with little or no reference to the nature, quality, or results of their work. Traditional compensation systems usually fail to account for (1) the differences in working conditions among schools; (2) the demand for scarce, specialized teaching skills; and (3) the individual teacher's effectiveness in the classroom.

Seeking New Compensation Models

The Teacher Incentive Fund has emerged as an important resource nationally for the design and study of alternative teacher and principal compensation approaches and best practices in linking those to the quality of student outcomes. It is possible that learnings from TIF programs will not only refine compensation and human capital management systems, but will also contribute to improvements in other teacher-related policies and practices, such as teacher evaluation, district placement practices, new teacher induction, improved teacher retention, and professional development.

Federal funding of compensation reform in education alleviates some of the difficulty that states and districts face in garnering adequate resources and building the political will to change entrenched practices, even where such practices are generally deemed ineffective. Providing federal funding dedicated to piloting innovative and promising practices has the potential to stimulate creativity and build a sound knowledge base about teacher performance policies that benefit teachers and students.

The establishment of and the ongoing commitment to the Teacher Incentive Fund by the U.S. Department of Education emanate from a theory that indiscriminate compensation and performance appraisal systems inhibit districts in meeting critical student learning goals. By promoting and incentivizing the development of new policies and practices for the strategic selection, placement, and retention of the best teachers with students who need them the most (but who may get fewer than their fair share), the Teacher Incentive Fund is testing its hypothesis and contributing to greater understanding of fundamental education reform.

Charlotte-Mecklenburg Schools

Planning for Student Success

By advancing the district in 2007-08 as a candidate for TIF funding, the leadership of CMS and CTAC relied on strong evidence of a readiness to succeed within the district, both in pursuing better outcomes for students and in electing compensation reform as a pivotal strategy to do so. First of all, the district's strategic plans, two of which have governed district priorities during the initiative, advocate for every school to have a strong, effective principal and for every classroom to have an effective teacher. As one component of *Strategic Plan 2010: Educating Students to Compete Locally, Nationally and Internationally*, the district established an Achievement Zone to address the needs of 10 of the district's highest need schools, identified as Low Performing or Priority schools by the state of North Carolina.[4]

Recognizing the challenge of staffing these highest need schools as well as the district's 100

other high need schools, the Superintendent of Schools and other district leaders established performance-based compensation as a core component of systemic reform and a catalyst for much needed improvements in student achievement. The learnings and gains of the TIF-LEAP initiative, which was created under *Strategic Plan 2010*, fed into the development of the district's subsequent *Strategic Plan 2014: Teaching Our Way to the Top* with its two key goals of (1) improving teaching, and (2) managing performance.

Building on Previous Efforts

Over a ten-year period prior to seeking TIF funding, CMS had experimented with a variety of bonus and incentive programs. A state program, North Carolina's ABC program[5] had been modeled on early efforts with performance pay in Charlotte-Mecklenburg. The history of that decade, outlined below, demonstrates that a district with experience in the use of incentives to strengthen the teaching force and improve student achievement existed prior to its engagement with the Teacher Incentive Fund.

- *Incentives for School-Based Performance:* The North Carolina State ABC program provided cash incentives up to $1,500 to certified staff (principals and teachers) and teacher assistants in K-12 schools making high or expected academic change as determined by the NC General Assembly, subject to budget approval. The Local Accountability Bonus (LAB) provided bonuses up to an additional $800 to staff (principals and teachers) at schools that met both Adequate Yearly Progress (AYP) and ABC goals. In 2004-05, staff at 72% of schools districtwide received the ABC awards, and staff at 54% of all schools received LAB awards.
- *Incentives for Classroom Level Performance:* During 2004-05, a pilot was conducted in 11 schools based on specific individual goals for each employee in the areas of attendance, professional development and student achievement. Twenty-three percent of instructional staff achieved stellar performance in 2004-05. The Successful Teacher and Administrator Reward (STAR), a locally funded program initiated in 2005-06, was aimed at high need schools, recognizing teachers and principals with bonuses of up to $1,400 and $5,000, respectively, based on End-of-Course or End-of-Grade state assessments, as well as teachers of grades K-2 or exceptional children.
- *Signing Bonus:* A variety of signing bonuses for teachers were available through the district in amounts ranging from $1,000 to $3,000, with the higher amounts for teachers at high need schools. Teachers of hard-to-staff subject areas were eligible for an additional $500 Critical Needs bonus.
- *Master Teacher Bonus:* To attract and retain teachers at high need schools, bonuses ranging from $1,500 to $2,500 were made available to teachers who could meet a set of criteria, including certification and licensure in the appropriate field, higher education, years of experience, and credentials for specializations (advanced placement, academically gifted, international baccalaureate).
- *Targeted Recruitment and Retention Bonus:* To increase quality teaching at four high need high schools—Garinger, Waddell, West Charlotte and West Mecklenburg—one-time signing bonuses ($10,000) and retention bonuses ($5,000) were made available to a limited number of teachers who demonstrated high score results in their classroom on EOC assessments.
- *Principal Compensation Incentives:* Annual principal salary increases were differentiated based on a performance matrix. Principals at high need schools also were eligible for Master Administrator bonuses of $2,500 based on a set of criteria demonstrating education, experience and credentials. In addition, as noted above, principals were eligible for bonuses based on student achievement, aggregated at the school level, consistent with CMS goals and the North Carolina ABC program.
- *Other:* Tuition reimbursement was provided on a limited basis in high need schools. Teachers with National Board Certification were eligible for a salary increase.

A local compensation committee composed of teacher, principal and administrative leaders was convened in February 2006 to steer an effort to build on these programs in order to develop an effective and comprehensive performance-based compensation system with more accountability and measurable results for the district.

Moving to a More Comprehensive Approach

Based on previous extra compensation efforts and the priorities of the strategic plan, a consensus had emerged among CMS leaders in favor of a more comprehensive, performance-based compensation system tied directly to the goals of increasing student achievement. The decision to move beyond the piecemeal incentive programs discussed in the previous section was based on data that demonstrated (1) the persistence of student underperformance, and (2) disappointingly low numbers of teachers and principals choosing to participate in incentive programs when given an opportunity, especially at the highest need schools.

Further, the community urged the district to address gaps in student achievement and hold teachers and administrators accountable for student achievement. This imperative was also reflected in the opinions of the Citizens Task Force on Education, a group charged with examining CMS outcomes and making recommendations to the Board and Superintendent.

Feedback from teacher focus groups in the district at this time made clear that while teachers and principals were open to incentive pay, the pre-existing programs were deemed ineffective because criteria overlapped, were duplicative, and/or were too restrictive. This feedback was an indication that the incentives of the past had not been substantial enough to drive change in the classroom, schools, and district systems and had lacked the clarity and sustainability to attract enough participants to make a difference.

Thus, district leadership, armed with a new strategic plan, community backing, and supporting data and feedback related to previous programs, was interested in 2007 in seeking support from the Teacher Incentive Fund and moving forward on a new approach to reward teachers and improve student learning.

The TIF-LEAP Initiative: Goals, Funding, and Governance

The introduction of the TIF-LEAP initiative into the priorities and culture of the district involved establishing goals commensurate with both district needs and requirements of the federal grant. Further, it was necessary to identify the source(s) of matching funds and the fiscal management that must attend the awarding of merit supplements. As the size and complexity of the grant demanded, the next task was creating and developing a governance structure to oversee the design, implementation, communication, and procedures for the initiative. Finally, an important task was the identification of participant schools in accordance with the criteria of the grant and the needs of the district's students and teachers.

TIF-LEAP Goals

Leadership for Educators' Advanced Performance (LEAP), the initiative resulting from the five-year grant from the Teacher Incentive Fund, took the form of a defined program of merit-based supplements for teachers and principals tied to student growth in, initially six and ultimately, 20 of the district's highest need schools, with the following goals:

- Provide differentiated levels of compensation based on student achievement gains and teacher/principal evaluations that include multiple classroom observations;
- Support the recruitment and retention of qualified teachers and principals in hard-to-staff schools and subjects;
- Build teacher and principal capacity to increase student achievement by aligning and improving district systems in support of the schools; and
- Develop district capacity to implement, scale-up, evaluate and sustain a performance-based compensation system, with measurable impact on student achievement.

The TIF-LEAP goals attest to the initiative's operating premise that aligning teacher and principal performance and compensation, the district's single largest expenditure area, with improved student achievement, the district's highest priority, is fundamental.

TABLE I.1

TIF-LEAP Funding, 2007-12

Year	TIF Award	CMS Match	Total Budget	Incentives Budget
2007-08	$1,987,589	$410,104	$2,397,693	$970,302
2008-09	$3,061,279	$1,211,813	$4,273,092	$2,897,648
2009-10	$3,154,594	$1,395,636	$4,550,230	$3,312,111
2010-11	$1,865,648	$2,737,415	$4,603,063	$3,737,415
2011-12	$1,811,157	$2,886,359	$4,697,516	$3,848,359
TOTALS	**$11,880,267**	**$8,641,327**	**$20,521,594**	**$14,765,835**

TIF-LEAP Funding, 2007-2012

The amounts of the TIF award and Charlotte-Mecklenburg's match over the life of the initiative, as well as the amount of funds dedicated to incentives or teacher and principal merit-based supplements, the term used for extra or bonus compensation by the Teacher Incentive Fund, are shown in *Table I.1*.

The incentives budget—the amount for teacher and principal merit supplements and benefits—accounts overall for more than 70% of available TIF-LEAP funds. Other TIF-LEAP budget items include project administration personnel and related costs, technology, technical assistance, research, and evaluation costs.

TIF-LEAP Governance Structure

The TIF-LEAP Steering Committee[6] became a standing body of representative CMS leaders convened to oversee the development and implementation of the initiative. Charged with assuring that the initiative met its stated goals and used the funds judiciously for its purposes, the Steering Committee led the planning and initial payouts of the initiative. Because the initiative, as proposed, impacted most district systems, a breadth of appointees from the following departments[7] and areas of the district served on the committee:

- Accountability
- Charlotte-Mecklenburg Association of Educators
- Communications
- Curriculum and Instruction
- Finance
- Human Resources
- Information Technology
- Participating Schools
- Planning and Project Management
- Principals: TIF-LEAP Schools (2)
- Professional Development
- School Zone Leadership
- Superintendent's Office
- Teacher Advisory Council
- TIF-LEAP Team[8]

Three members of the CTAC staff served as non-voting attendees at the Steering Committee meetings to provide research findings, technical counsel, and oversight of the federal grant. The Board of Education appointed one of its members as liaison to the TIF-LEAP initiative in order to facilitate the exchange of information and ensure a link between practice and policy.

TIF-LEAP started up in the first year under the supervision of Human Resources, but the initiative's supervision was moved to Curriculum and Instruction in the second year, commensurate with the selection of Student Learning Objectives—an instruction-focused performance-based compensation approach.

Four major working groups were charged by the Steering Committee with completing the initial development and implementation tasks as well as planning for the long-term effectiveness and sustainability of the initiative. These groups included:

- Human Resources/Finance Working Group
- Instructional Support Working Group
- Principals' Working Group
- Student Learning Objectives Design Team

While contributing to task completion, working groups also provided a vehicle to increase the number and expertise of participants involved in the planning and development of the initiative, identifying essential supports, and looking ahead at fiscal sustainability.

Communication and Public Information for the TIF-LEAP Initiative

A communication plan was developed for the initiative through the auspices of the communication and public information arm of the district. As shown in *Table I.2*, first phase communication goals and measures related to promoting awareness and building support in participant schools and the entire district, and second phase goals and measures primarily related to the awareness and support among parents and community members.

The most visible tool for the communication and public information effort, the TIF-LEAP section of the district website (www.cms.k12.nc.us), provided updates about the project and sources of further information. Annual board reports and presentations by the TIF-LEAP team were posted in the Board section of the website. Wider dissemination of information about the TIF-LEAP initiative occurred primarily through the participation of the TIF-LEAP team members and district leaders in state and national conferences on performance-based compensation, teacher evaluation, and Student Learning Objectives, which led to the widespread sharing of implementation materials developed by the TIF-LEAP team. Awareness and support among participants and constituents for compensation reform and the effectiveness of communication about the initiative was tracked largely through surveys and interviews conducted annually by CTAC, the results of which are discussed in Chapter VI.

TABLE I.2

Communication Plan Goals and Measures

Goal	Measure
Phase I	
Introduce and increase awareness of pilot compensation reform initiative within TIF-LEAP schools.	Majority will say they are aware of compensation reform initiative as measured by a teacher survey in pilot schools.
Introduce and increase awareness of pilot compensation reform initiative within CMS.	Majority will say they are aware of initiative as measured by a teacher/employee survey.
Build support for compensation reform within CMS.	Majority of CMS employees will say they favor compensation reform on teacher/employee survey.
Phase II	
Introduce and increase awareness regarding compensation reform among CMS parents, volunteers and community/business partners.	Majority will say they are aware of CMS' compensation reform initiative as measured by parent and community annual opinion poll/survey.
Build support for compensation reform among CMS parents, volunteers and community/business partners.	Majority will say they support efforts to expand LEAP/compensation reform district-wide as measured by annual parent/community opinion poll.
Position CMS as a leader in compensation reform within North Carolina, and nationally.	CMS cited as positive example or case study for successful compensation reform in K-12 education conferences, academic journals and trade press.

School Participation in the Initiative, 2007-12

A pillar of effective compensation reform is teacher involvement, such as providing potential participants an opportunity to examine and opt into a compensation initiative.[9] However, the Teacher Incentive Fund criteria, which include a goal of "increasing the number of effective teachers teaching poor, minority, and disadvantaged students," signaled to CMS leadership that TIF-supported initiatives must be implemented in schools with high need populations. Ultimately, the schools selected were ones deemed to be of highest need and priority. Even with the lack of an opt-in opportunity, survey and interview data throughout the initiative showed that administrative and teaching staff at the selected schools supported the concept of performance-based compensation, including rewards based on increased student growth.

For example, survey data from the fall of 2008 indicate that all respondent groups at the participant school (TIF-LEAP) sites favored rewards to teachers for overall school improvement (95.7% principals, 94.1% assistant principals, and 84.7% teachers), only slightly more than they favor awards for individual classroom teachers (88.6% principals, 91.2% assistant principals, and 84.8% teachers). A comparison of TIF-LEAP and non TIF-LEAP school responses on that same survey shows that members of TIF-LEAP schools agree at a higher level to both approaches than do members of non TIF-LEAP schools, but respondents were generally positive about the concept of performance awards.

The TIF-LEAP schools selected by the district were phased in over a three-year period beginning with six schools in the first year, moving to a total of ten in the second year, and reaching the maximum in the third year. Because one of the high schools originally designated for participation (Garinger) was reorganized into five small schools, the actual full slate of participant schools in the third and fourth years rose to twenty (20) rather than the projected sixteen (16), without changing the number of teachers and students included in the initiative. In the fifth and final year of the initiative, school closures, reorganizations, and repurposing reduced the number of participating schools to eleven (11), though serving mostly the same students. *Table I.3* outlines the history of school participation in the TIF-LEAP initiative.

For 2011-12, four (4) of the twenty TIF-LEAP schools were closed; one (1) was repurposed (as a language academy); three (3) were reorganized into a K-8 configuration; and five (5) small high schools were re-integrated into a comprehensive high school. The decision to make structural changes, including closings, in the last year of the initiative ensued from CMS budget cuts necessitated by the revenue impact of the national recession. However, the decision of the district and most[10] of the remaining participant schools to stay the course during the final year despite fiscal/structural changes speaks to the serious commitment that had characterized the initiative.

Layoffs of teachers, in both 2010-11 and 2011-12, also impacted the continuity of the initiative. All schools participating in the initiative were impacted by cutbacks, closures, and reorganizations in the final year (2011-12) of the initiative. Not surprisingly, district discussions in the year preceding layoffs and school changes were also distracting for the staff whose positions and school placements were at risk, as evidenced in interviews conducted in the winter of 2011.

Performance-Based Compensation Approaches in the TIF-LEAP Initiative

The performance-based compensation work of the TIF-LEAP initiative resulted in two approaches to merit-based supplements for principals, assistant principals, and teachers in the twenty participating schools. The initial approach selected for implementation in year two of the initiative is known as Student Learning Objectives, also referred to as the SLO process, and a different approach, implemented in year three is called the value-added measure, or VAM.

The design and development of SLOs was the work of the Student Learning Objectives Design Team, under the leadership of the Curriculum and Instruction Department, while the VAM was formulated by the Accountability Department and executed through an *ad hoc* group, confirmed by the Steering Committee. The approaches are described briefly here and discussed more extensively in Chapter II.

TABLE I.3

TIF-LEAP School Participants, 2007-12

School/Academic Year	2007-08	2008-09	2009-10	2010-11	2011-12
Billingsville Elementary School	X	X	X	X	X
Shamrock Gardens Elementary School	X	X	X	X	X
Bishop Spaugh Middle School	X	X	X	X	O
J.W. Wilson Middle School	X	X	X	X	O
Martin Luther King, Jr. Middle School	X	X	X	X	X
Sedgefield Middle School	X	X	X	X	X
Druid Hills Elementary School		X	X	X	K-8
Highland Renaissance Elementary School		X	X	X	X
Reid Park Elementary School		X	X	X	K-8
John Taylor Williams Middle School		X	X	X	O
Berryhill Elementary School			X	X	K-8
Lincoln Heights Elementary School			X	X	O
E.E. Waddell High School			X	X	R
Business and Finance @ Garinger High School			X	X	I
International Studies @ Garinger High School			X	X	I
Leadership and Public Service @ Garinger High School			X	X	I
Math and Science @ Garinger High School			X	X	I
New Technology @ Garinger High School			X	X	I
West Charlotte High School			X	X	X
West Mecklenburg High School			X	X	X

X = Participating School; O = Closed School; K-8 = Restructured to K-8 Configuration;
I = Integrated into Comprehensive High School; R = Repurposed to Language Academy

Approach One: Student Learning Objectives

Planned during year one and implemented beginning in year two of the initiative, a Student Learning Objective is described by the TIF-LEAP team in training materials as: "a targeted, long-term goal for advancing student learning. This data-informed process involves diagnosing and improving specific student learning needs. Each SLO includes the following six components based in effective instructional design: learning content, population, interval, assessment(s), growth expectations, and strategies."[11] In this compensation approach, teachers are active participants—analyzing data, selecting a focus and a growth target, choosing teaching strategies, and assessing students—who work with their principal/supervisor and, if they so choose, with colleagues to reach a student growth target. All classroom teachers in the participant schools, regardless of the grade level or subject taught, develop SLOs. Through collaboration, non-classroom teachers with teaching responsibilities, such as resource specialists, may participate on a team with classroom teachers.

Approach Two: Value-Added Measure

In 2009-10, CMS leadership moved forward with the implementation of a value-added measure, a statistical methodology designed to measure a teacher's effectiveness in raising the scores of his/her students on the annual North Carolina End-of-Grade or End-of-Course assessment. The

TABLE I.4

Comparison of Participant Eligibility Criteria for SLO and VAM Merit-Based Supplements[13]

	Student Learning Objectives	Value-Added Measure
Teacher Eligibility	**ELIGIBLE** • Certified and on the teacher pay schedule providing direct instruction • "Proficient" or above ratings on each overall evaluation standard • 85% attendance during interval specified on the SLO • Attend TIF-LEAP training sessions on the SLO process • Submit SLO(s) for approval to principal prior to the deadline(s) • Provide assessment data to demonstrate that at least 75% of the specified students achieved or exceeded growth expectations **NON-ELIGIBLE** • Interns • Non-certified Employees • Counselors, Social Workers, Speech Pathologists, and Psychologists • Substitute, Interim, Temporary Employees, and Tutors • Teachers arriving after the last training date	**ELIGIBLE** • Certified and on the teacher pay schedule providing direct instruction • "Proficient" or above ratings on each overall evaluation standard • Aggregate growth at or above the 70th percentile using the district's VAM for applicable courses in one of the following categories: ▪ teachers who are the primary instructors for students in content areas assessed with an EOC/EOG test, including self-contained Exceptional Children's classrooms; ▪ secondary teachers who share responsibility for instructing students are eligible for a team growth bonus (subject to approval) • Teach at least five students who have valid scores for the year in the same grade and content area. The students must have been enrolled for: ▪ 140 days at the current school for an EOG, or ▪ 70 days at the current school for EOCs taught in the semester format **NON-ELIGIBLE** • Interns • Non-certified Employees • Counselors, Social Workers, Speech Pathologists, and Psychologists • Substitute, Interim, Temporary Employees, and Tutors
School Eligibility	Not Applicable	Teachers in a school are eligible for a school growth bonus if the school's aggregate growth on all EOG/EOC tested subjects in the school falls at or above the 60th percentile using the district's value-added model for growth (and the teacher meets the evaluation criteria). This award includes "non-tested" teachers as well as primary and secondary teachers.

VAM is calculated through a multilevel model in which students are nested within classrooms, which are nested within schools. The VAM "provides a rating of the relative performance of a teacher after adjusting for conditions that are outside of a teacher's control (such as the number of student absences)."[12] As designed and implemented, the value-added process compared the effectiveness of teachers of tested grades and subjects districtwide, expressed as a percentile ranking.

Eligibility for Participation

A critical task for the governance of compensation reform programs is to create the "rules of the road." Decisions about who is eligible to participate in a compensation program and who is not, and any requirements or criteria for participation and receipt of an award, must be as clear as a teacher contract and administered in a consistent, fair, and equitable way.

Eligibility decisions are rarely straightforward, however. For example, "who is a teacher?" A teacher is not always a person with a classroom and 30 students, but may also be a resource teacher, a special education teacher, a media teacher, etc. If there is a decision to include teachers who do not have classrooms with rosters of assessed students attached to their names, then a method for evaluating their contribution will need to be devised and agreed upon. As shown in *Table I.4*, the TIF-LEAP Steering Committee decided to base eligibility on direct instruction.

Teacher awards for SLOs were based on meeting growth targets set by the teacher on one or two objectives that had been approved and verified by the principal. The TIF-LEAP teacher eligibility for an SLO compensation award was predicated on an established set of teacher criteria that included ratings of proficiency on annual performance evaluations, percentage of time teaching, and submission of required documents. Additionally, the SLO approach required participation in a training component.

The VAM approach required teachers to teach in the assessed grades and subjects[14] and to be proficient on performance evaluations. In the 20 TIF-LEAP schools, teachers whose VAM scores ranked at or above the 70th percentile using the district's VAM for applicable courses, earned individual bonuses. Additionally, teacher and principal eligibility for school level performance bonuses was earned by those TIF-LEAP schools that performed in the top 40% of the district as determined by the VAM.

Payout Structure for TIF-LEAP

Just as important as a clear delineation of eligibility is the management of the payout structure so that (1) the amount of the merit-based supplement or bonus is clear, and (2) those who do qualify receive the money in the designated pay period.

The evolving payout structures in the successive years of the initiative demonstrate the Steering Committee's responses to changes in number of participants, district leadership priorities, and available funding, as well as to the number and type of compensation approaches funded under the grant. *Tables I.5, I.6* and *I.7* show the changes in the payout structures over three years.

TABLE I.5

Payout Structure for 2008-09

Position	Criteria	Amount	Total
Principals	SLO Facilitation	$5,600	**Up to $5,600**
Assistant Principals	SLO Facilitation	$4,200	**Up to $4,200**
Teachers with/without End-of-Course/ End-of-Grade Test(s)	SLO Attainment	$1,400/SLO (two required)	**Up to $2,800**

Source: CMS Memo, Undated; TIF-LEAP Document, 3/17/2009

TABLE I.6

Payout Structure for 2009-10

Position	Criteria	Amount	Total
Principals	SLO Facilitation*	$1,000 (Tier 1) $1,500 (Tier 2) $2,000 (Tier 3)	**Up to $4,000**
	School Growth (VAM)	$500-$2,000 (in top 40% of district)	
Assistant Principals	SLO Facilitation*	$750 (Tier 1) $1,125 (Tier 2) $1,500 (Tier 3)	**Up to $3,500**
	School Growth (VAM)	$500-$2,000 (in top 40% of district)	
Teachers with End-of-Course/ End-of-Grade Test(s)	SLO Attainment	$1,400/SLO (two required)	**Up to $5,300**
	Classroom Growth (VAM)	$500-$2,500 (in top 30% of district)	
Teachers without End-of-Course/ End-of-Grade Test(s)	SLO Attainment	$1,400/SLO (two required)	**Up to $4,200**
	Team Growth (VAM)	$400-$1,400 (in top 30% of district)	

Source: TIF-LEAP Summary for Board of Education, 12/08/2009

* Tier 1 = Up to 19 teachers; Tier 2 = 20-28 teachers; Tier 3 = 29 or more teachers per administrator

TABLE I.7

Payout Structure for 2010-11

Position	Criteria	Amount	Total
Principals	SLO Facilitation	$1,800	**Up to $5,400**
	School Growth (VAM)	$2,200-$3,600 (in top 40% of district)	
Assistant Principals	SLO Facilitation	$1,800	**Up to $4,700**
	School Growth (VAM)	$1,500-$2,900 (in top 40% of district)	
Teachers with End-of-Course/ End-of-Grade Test(s)	SLO Attainment	$1,000/SLO (two required)	**Up to $7,400**
	Classroom Growth (VAM)	$2,000-$3,000 (in top 30% of district)	
	School Growth (VAM)	$1,000-$2,400 (in top 40% of district)	
Teachers without End-of-Course/ End-of-Grade Test(s)	SLO Attainment	$1,000/SLO (two required)	**Up to $6,400**
	Team Growth (VAM)	$1,000-$2,000 (in top 30% of district)	
	School Growth (VAM)	$1,000-$2,400 (in top 40% of district)	

Source: TIF-LEAP Summary for Board of Education, 12/14/2010

Summary

The Teacher Incentive Fund and Charlotte-Mecklenburg Schools bonded because the goals of the federal program and those of CMS dovetailed. CMS had a demonstrated need and commitment to improve teaching and learning in high need schools, as well as a history and a willingness to use incentive funds available at the local and state levels to meet its goals.

The thoughtful work of the Steering Committee got the initiative off the ground; developed a governance structure to establish a nd monitor the parameters and payouts of merit-based supplements; and phased in the district-selected participant schools. It adopted Student Learning Objectives as the initial approach for improving and measuring teaching effectiveness and awarding bonuses. Later, a value-added measure was added to the initiative.

Over five years, it can be expected that conditions will change in schools and in a district. Two such phenomena occurred in Charlotte-Mecklenburg Schools that had a significant impact on the shape and outcome of the TIF-LEAP initiative: (1) a structural change to the initiative, that took effect in year three, to accommodate the addition of the VAM approach in the participant schools; and (2) unusually deep fiscal cuts, necessitated by the national recession, that led to teacher layoffs and school closures and reorganizations. *Table I.8* at the conclusion of this chapter shows the major changes over the five years of the initiative.

The district, particularly through the TIF-LEAP team, worked to keep the initiative on a steady keel and maintain fidelity in the implementation of the performance-based compensation approaches.

TABLE I.8

TIF-LEAP and Charlotte-Mecklenburg Schools Overview

Year	TIF-LEAP Initiative Actions	Participating TIF-LEAP Schools			CMS Actions
2007-08 TIF-LEAP Year I	• Develop initiative goals, structure, and Steering Committee • Begin with six schools • EOG/EOC and ABC-based payouts • Convene SLO Design Team	Billingsville ES Shamrock Gardens ES Bishop Spaugh MS	J.W. Wilson MS Martin Luther King, Jr. MS Sedgefield MS		• Assign Human Resources as TIF-LEAP initiative supervision • Develop TIF-LEAP Communication Plan
2008-09 TIF-LEAP Year II	• Increase to ten schools • Year I SLO implementation • SLO-based payouts • Convene District-Developed Growth Measure (VAM) Task Force • Report to Board	Billingsville ES Shamrock Gardens ES Bishop Spaugh MS J.W. Wilson MS	Martin Luther King, Jr. MS Sedgefield MS Druid Hills ES	Highland Renaissance ES Reid Park ES John Taylor Williams MS	• Assign Curriculum and Instruction as TIF-LEAP initiative supervision • Hire TIF-LEAP staff • Principal change in one school
2009-10 TIF-LEAP Year III	• Increase to twenty schools • Year II SLO implementation • Year I VAM implementation • Modify payout structure to include VAM • SLO and VAM payouts • Report to Board	Billingsville ES Shamrock Gardens ES Bishop Spaugh MS J.W. Wilson MS Martin Luther King, Jr. MS Sedgefield MS Druid Hills ES	Highland Renaissance ES Reid Park ES John Taylor Williams MS Berryhill ES Lincoln Heights ES E.E. Waddell HS Business and Finance @ Garinger HS	International Studies @ Garinger HS Leadership and Public Service @ Garinger HS Math and Science @ Garinger HS New Technology @ Garinger HS West Charlotte HS West Mecklenburg HS	• Implement Small Schools Initiative at Garinger HS • Principal change in two schools • Send teacher layoff notices
2010-11 TIF-LEAP Year IV	• Year III SLO implementation • Year II VAM implementation • SLO audit • SLO and VAM payouts • Report to Board	Billingsville ES Shamrock Gardens ES Bishop Spaugh MS J.W. Wilson MS Martin Luther King, Jr. MS Sedgefield MS Druid Hills ES	Highland Renaissance ES Reid Park ES John Taylor Williams MS Berryhill ES Lincoln Heights ES E.E. Waddell HS Business and Finance @ Garinger HS	International Studies @ Garinger HS Leadership and Public Service @ Garinger HS Math and Science @ Garinger HS New Technology @ Garinger HS West Charlotte HS West Mecklenburg HS	• Implement NC teacher evaluation system • Principal change in seven schools • Send additional teacher layoff notices • Select district schools for closures • Begin CMS Strategic Plan 2010 • Superintendent leaves; interim assigned
2011-12 TIF-LEAP Year V	• Decrease to eleven schools • Year IV SLO implementation • Year III VAM implementation • Report to Board	Billingsville ES Shamrock Gardens ES Martin Luther King, Jr. MS Sedgefield MS	Druid Hills Academy K-8 Highland Renaissance ES Reid Park Academy K-8 Berryhill School K-8	Garinger HS West Charlotte HS West Mecklenburg HS	• Close the following schools: Bishop Spaugh, J.W. Wilson, John Taylor Williams, E.E. Waddell • Reconfigure as K-8 ESs: Reid Park, Berryhill, Druid Hills • Re-integrate: Small High Schools as Garinger HS • Principal change in three schools

Endnotes

1 The Teacher Incentive Fund is authorized in P.L. 109-149—the Departments of Labor, Health and Human Services, and Education, and Related Agencies Appropriations Act, 2006, Title V, Part D.

2 The 2011-12 enrollment is 138,012 K-12 students as reported on the CMS website at http://www.cms.k12.nc.us/cmsdepartments/StudentPlacement/PlanningServices/Pages/Enrollmentdata.aspx.

3 http://www2.ed.gov/programs/teacherincentive.

4 School zones were reorganized again in 2010-2011.

5 "A" for Accountability, "B" for an emphasis on the Basics, and "C" for increased local Control.

6 During year one of the initiative (2007-08), the TIF-LEAP Steering Committee was known as the Compensation and Support Task Force.

7 Some department names changed over the course of the initiative.

8 The district and grant provided the TIF-LEAP initiative with a staff that included an Executive Director, a Student Learning Objectives Specialist, a Compensation Specialist, a Senior Assessment Program Analyst, and a Senior Administrative Secretary, all referred to as the TIF-LEAP Team. The Executive Director, Student Learning Objectives Specialist, and Senior Assessment Program Analyst became members of the Steering Committee.

9 Slotnik, W. (2009). *It's more than money: Making performance-based compensation work*. Center for American Progress at http://www.americanprogress.org/issues/2009/07/more_than_money.html.

10 The new principal of a restructured school opted out of the initiative for 2011-12, though the teachers who wished to continue to develop SLOs could do so in cooperation with the TIF-LEAP staff.

11 http://www.cms.k12.nc.us/CMSDEPARTMENTS/TIF-LEAP/Pages/StudentLearningObjectives.aspx.

12 The VAM description is adapted from the TIF-LEAP initiative description at www.cms.k12.nc.us, a brief "CMS TIF-LEAP Value-Added Measure of Teacher Effectiveness," and "Notes for Ann Helms on Value-Added Calculation" at www.cms.k12.nc.us.

13 2011-2012 Implementation Guide for TIF-LEAP, Charlotte-Mecklenburg Schools.

14 At least two years of assessment data on a teacher's students are fundamental to the statistical analysis that determines the VAM, so having a minimum of students with two or more annual assessment scores is required.

CHAPTER II

TIF-LEAP Performance-Based Compensation Approaches: Design, Development, and Implementation

Implementing performance-based compensation is a complex undertaking. Implementing two different approaches to performance-based compensation compounds the complexities, with effects both beneficial and distractive. To understand the outcomes of the TIF-LEAP initiative, it is critical to look at the two adopted approaches to performance-based compensation, the reasoning and research related to each approach, and the features of their design, development and implementation.

Teacher performance evaluation for the purposes of either compensation or employment appraisal is based on various measures of teacher effectiveness,[1] such as: (1) observations and/or evaluations of the effectiveness of teacher *input* (what teachers do); (2) analyses of the effectiveness of teacher *output* (student performance); or (3) a *combination* of input and output measures. The district-developed value-added measure is in the second category, measuring teacher output based on analyses and ranking of teacher contributions to student growth on the

EOG/EOC assessments. Student Learning Objectives are in the third category, improving the planning and quality of the teacher-developed learning objectives (input) and measuring student growth on the SLO target (output).

A decision to provide monetary incentives is based on the assumption that they will lead to increased teacher effort and/or to changes in teacher behavior that will, in turn, increase effectiveness and result in measurable increases in student achievement. Though many dimensions of teaching may be observed and measured to judge teacher effectiveness, a teacher's contribution to student growth on state assessments is the current gold standard for policymakers in incentive-based programs.

In both approaches implemented through the TIF-LEAP initiative, teacher effectiveness is ultimately measured by student growth. However, the approaches differ in opportunity, participant input, focus, theory, comprehensiveness, professional development, and assessment tools. The design, theory, and implementation features of each approach are explored in the following discussion.

SLOs differ by involving teachers in a process that brings (1) greater precision to the instructional analyses that go into developing objectives, and (2) more science to the art of teaching through thoughtful baseline data analyses and student growth projections.

Approach One: Student Learning Objectives

Student Learning Objectives are grounded in teaching craft and practice. Teachers engage in precise instructional planning in order to gain better results with their students. In the process of developing and submitting SLOs, teachers analyze student performance data, including prior achievement data and the results of any pre-tests administered. They then compose a designated number of objectives (usually two), using a multi-component protocol. The protocol also guides teachers in selecting significant standards-based learning content and suitable assessments aligned with the content and standards.

Teachers evaluate baseline data, identify student needs, project student growth targets, and provide rationales for their decisions. Teachers also seek and plan the most effective instructional strategies to meet identified student needs and establish an appropriate interval of instructional time. At a designated time, teachers collect and present evidence of the level of attainment of their objectives.

Principals have a significant role in reviewing and approving teacher objectives, strengthening the delivery of instruction, and validating the teacher's evidence of attainment and eligibility for bonus pay.[2] Teachers may choose to collaborate with their colleagues in developing objectives but are accountable as individuals for the outcomes. For example, a group of fourth grade teachers may collaborate on the content, assessment and interval of instruction in their SLOs, but because they have different students and different baseline data, they set different growth targets.

While goals and objectives are commonplace in district and school settings, SLOs differ by involving teachers in a process that brings (1) greater precision to the instructional analyses that go into developing objectives, and (2) more science to the art of teaching through thoughtful baseline data analyses and student growth projections. The depth of thinking and the extent of original teacher work generated throughout the SLO process runs counter to many educational trends that require teachers to follow published lesson scripts.

Student Learning Objectives: Research and Best Practice

Student Learning Objectives form a solid foundation for awarding extra compensation to teachers who show evidence of a positive impact on student learning. First formulated and piloted by the Denver Public Schools (1999-2004), these teacher-developed objectives create a direct, credible link between learning and earning and remain a component of the multi-layered

compensation system in Denver, known as ProComp.[3]

From design to implementation to mid-course corrections, teacher leaders and the union in Denver did much of the heavy lifting in collaboration with the administration and Board of Education. They were also the advocates for an independent evaluation of the pilot. A four-year evaluation conducted by CTAC in Denver found that students whose teachers crafted the highest quality objectives showed more than a year's worth of gain on the *Colorado Student Assessment Program* and the *Iowa Test of Basic Skills* at all three school levels during each year of the four-year study.[4]

SLOs challenge the entire school system to perform better on behalf of teachers and students.

The evaluation found that SLOs get better results than conventional instructional objectives because they are developed and implemented through a more thoughtful, evidence-based process with a more selective use of teaching and assessment practices. While building on teacher craft knowledge and respecting the teaching realm that is the classroom, SLOs challenge teachers to (1) think critically about their instructional decisions; (2) ground decisions firmly in sound data analyses and learning theory; (3) follow stricter rules of evidence in the evaluation of student growth than may be typical of most classroom practice; and (4) be accountable for student results.

In performance-based compensation initiatives, the SLO approach enables all teachers to participate, not just those who teach the state-tested subjects and grades. This is a critical feature for teacher leadership in most districts, and nationally. Additionally, the SLO process engages participants in a cluster of teaching practices associated with student achievement,[5] increasing their likelihood of succeeding and earning their incentive. Finally, SLOs challenge the entire school system to perform better on behalf of teachers and students.

The process of developing, implementing, and assessing SLOs takes into account the strengths of teachers and still challenges them to improve. The reasons SLOs[6] are considered well suited for both promoting and measuring teacher effectiveness in the classroom are enumerated in the next column.

1. The development or selection of classroom instructional objectives is decidedly a teacher activity. The SLO process respects teachers as professionals, starting with something they know and extending their thinking about student learning.
2. Developing learning objectives is usually part of teacher "boot camp," a familiar activity.
3. Instructional (or planning) objectives are commonplace in many curriculum guides, textbooks, and other materials that teachers use in the classroom.
4. Improving the timeliness and availability of student data for teachers strengthens the quality of objectives, and can positively impact the effectiveness of the district's student data system for teachers, as the need for timely and accessible student information scales up.
5. The SLO process dovetails with and enhances other reforms. It does not impose a teaching method or conflict with state, district, or Common Core standards. It works for individual teachers, teachers in professional learning communities, and teachers of special services.
6. The implementation of SLOs promotes a "planning backwards" approach to instruction, encouraging more thoughtful and analytical uses of assessments when planning instructional strategies and measurements with student results in mind.
7. Even though objective-setting is pervasive in classrooms, schools, and districts, in-depth assessment of outcomes or evaluation of results may be either cursory or passed over all together. An objective-based compensation system influences an entire organization to become more accurate, open, and reflective about student outcomes.

Teacher quality is the most critical variable in student achievement. Using an incentive approach that maximizes a teacher's capacity to plan, focus,

problem-solve, and seek solutions in order to improve student achievement capitalizes on an under-utilized resource in education reform—a thinking classroom teacher.

In interviews and focus groups, teachers discuss why SLOs work as authentic measures of their performance. Like teachers in Denver, teachers in Charlotte-Mecklenburg describe being more focused and more knowledgeable about their students' strengths and needs as a result of the process. They attribute increased collaboration among their colleagues to the shared experiences of developing SLOs. They are also reflective about SLOs that do not succeed and what they will do differently next year. In this regard, the SLO process is not only an incentive initiative, but also a professional development initiative that is more than a one-day workshop.

Studies from the organizational psychology field shed light on why a teacher's setting an objective and then working to get a good result can be more effective than other types of organizational goal setting. One meta-study of organizational goal setting indicates that the most difficult goals produce the highest levels of effort and performance and that the more specific (what, who, when, by what standard) and personal the goals, the greater the likelihood of attaining them.[7] This analysis also finds that individual goals affect performance by: (1) directing attention and effort toward activities that are relevant; (2) energizing, or creating greater effort; (3) impacting effort (i.e., more time on task); and (4) arousing "task-relevant knowledge and strategies."

To summarize, well-crafted SLOs help teachers to improve student learning by (1) expanding on what teachers know and do best; (2) promoting greater depth and rigor in thinking about and planning instruction; (3) creating the conditions for professional learning and development; and (4) triggering changes in attention and focus, effort, time on task, and use of task-relevant knowledge and strategies. Rather than assuming that monetary incentives alone will generate changes in teacher practice that lead to student improvement, the SLO approach is intended to prompt changes in practice that result in better outcomes for students.

Student Learning Objectives: The Design for TIF-LEAP Schools

The TIF-LEAP Steering Committee evaluated various incentive and compensation programs and consulted extensively with stakeholders before selecting Student Learning Objectives to be the foundation of the performance-based compensation system in the initiative. The committee then charged the SLO Design Team to design and develop the SLO approach. The SLO Design Team was comprised of district and school site educators—teachers, principals, subject specialists, project specialists, and department heads, joined by department specialists in assessment, accountability, technology, and professional development—and supported by research and technical counsel of CTAC.

Because the SLO process is not a pre-packaged reform, it needs to be customized to the district curriculum, instruction, and assessment program and supported by other district systems such as data and technology. Thus, initiating SLOs in participant schools represented a significant commitment from many professionals in Charlotte-Mecklenburg Schools.

The original seven components selected as the core of the SLO framework were adapted from

TABLE II.1

Components of Student Learning Objectives

Component	Guiding Questions
Population	Which students are targeted and why?
Learning Content	What curricular standards are included and why?
Interval	When during the year will the instruction take place and why?
Assessment(s)	How will growth in student learning be measured and why?
Growth Expectations	What expectations are set for each student and why?
Strategies	What methods of teaching will be used and why?

ones developed in the Denver Public Schools. These included: population, learning content, rationale, strategies, interval, assessment and targeted growth. Building on the Denver design, the SLO Design Team developed an SLO Guide that elaborates each element or trait, providing "guiding questions" and "criteria" to prompt critical thought as well as clarity and quality in the implementation process. Participants indicate that a document such as the SLO Guide serves as more than a planning outline. It assures that everyone involved in the process works with common understandings of the components required and the level of rigor expected.

Over time, the traits and heuristics of the original SLO design in CMS were refined, clarified, and simplified into six traits, as shown in *Table II.1.* The original component "Rationale" was folded into all of the other components with the addition of "why?" to the questions.

In addition to designing the SLO components and a heuristic for teachers and principals to use, the SLO Design Team, with the help of technical counsel, worked to identify and address thorny system-related issues and potential challenges, including the need for (1) prompt teacher access to student data, (2) a description of the types and purposes of assessments and assessment item banks currently available to teachers, and (3) the development of an intra-district web-based platform for teachers to upload their SLOs and principals to review and approve them.

The SLO Design Team also addressed a myriad of issues related to assessments and their validity and reliability, professional development, and leadership training. From system support to project management, the SLO Design Team focused on the breadth and complexity of planning and implementation.

Student Learning Objectives: Implementation in TIF-LEAP Schools

Student Learning Objectives are, at root, an instructional reform. Because much of the required expertise of SLO implementation resides in the instructional area, placing the initiative under the direction of this unit of the district proved to be a pivotal decision in CMS. Cast in the leading roles of implementing Student Learning Objectives in the ten schools in year two of the initiative and ten additional schools in years three and four, the TIF-LEAP team from the Curriculum and Instruction Department became the "face of Student Learning Objectives" in the schools.

The priority for the TIF-LEAP team beginning in the fall of 2008 was to introduce SLOs to teachers and then assist them through the process of crafting objectives. The TIF-LEAP team assisted teachers and principals in developing SLOs aligned with district and school plans, state and local standards, district curriculum, and assessments, while assisting teachers to use data to identify, analyze, and address student needs. In implementing SLOs, teachers learned to map the destination(s) for their students and pinpoint the position of each student before starting the road trip, planning a route to get all students there.

Implementation work, no matter how well planned, is about keeping one's eye on the goal while catching myriad balls coming one's way. Over the course of the initiative, the TIF-LEAP team maintained fidelity to the SLO process while refining and making mid-course improvements to strengthen the initiative. The TIF-LEAP team's support was frequently cited by participants as a model for delivering a high standard of services to schools and teachers.

Teacher and principal interviews together with analyses of artifacts of the implementation highlight the fact that services provided by the TIF-LEAP team were of particular importance to participants:

- providing ongoing school site training and timely services to participants;
- fostering rigor in assessment choices and practices;
- increasing the level of science in teacher planning;
- linking SLO processes to teacher evaluation; and
- analyzing and auditing SLO implementation for improvement.

Providing Ongoing School Site Training and Timely Services to Participants. Teachers and principals in the participant schools identified in interviews what they valued most about the TIF-LEAP team's efforts—the hands-on training method that led to ongoing access to and support from the team. Teachers preferred not to leave their campuses for training, so the TIF-LEAP team provided training and follow up at each school, working out a schedule in collaboration with the principal and faculty.

In addition to on-site training, the TIF-LEAP team prioritized communication with teacher and principal participants, responding to requests for assistance within a day, and going out to the school to troubleshoot or to follow-up directly with individual teachers, as needed. Many interviewees, both participants and district leaders, advocated the use of the TIF-LEAP team's service model for all professional development, believing it to be a particularly beneficial method for an initiative that (1) required more expert uses of technology; and (2) was of higher stakes for teachers because compensation was connected directly to success.

The TIF-LEAP team created training and guidance materials for introducing teachers and principals to SLOs and guiding them in the process that included: (1) an annually updated how-to handbook about SLOs and district expectations and policy; (2) a walkthrough of the features of the student data system with a technical expert[8] where teachers learned to retrieve the information they needed for a baseline analysis of their students' strength and needs; and (3) a demonstration of the web-based platform for completing and managing SLOs.

Further, the TIF-LEAP team served as an interface between the schools and relevant district departments to address problems emerging in implementation. As an example, the web-based platform required several iterations to become more effective and user friendly. Interestingly, in describing their concerns related to the platform, teachers separated their dissatisfaction with the technology from their appraisal of the overall process and the implementation services of the TIF-LEAP team.

Fostering Rigor in Assessment Choices and Practices. SLOs derive much of their effectiveness and acceptance from the fact that they are grounded in classroom practice. Where classroom assessment practice is ineffective, the SLO process will reveal it. Assessment is an area of school and teacher practice that, when left to drift, may become fraught with inconsistencies, idiosyncrasies, and biases. Over the course of the initiative, the TIF-LEAP team pressed for rigor in the assessments used in SLOs, particularly by providing tools for more careful critiques and decisions regarding assessments. These efforts promoted:

- alignment between the assessment and the content/standard being taught;
- appropriate and consistent timelines for pre- and post-assessments based on the interval of instruction;
- verification of the reliability of the items on the pre- and post-assessments;
- avoidance of bias in test items;
- the use of multiple measures for purposes of validity;
- the explanation of the method of evaluating divergent (open-ended) items on assessments;
- verification that the same assessment(s), pre and post, were used; and
- rationales for all of these choices, including test modifications.

Ultimately, the web-based software guided teachers through an analysis of their selected assessment. Principals acted as the arbitrators of assessment choices and results, and teachers were required to keep their pre- and post-assessments for potential audits.

Increasing the Level of Science in Teacher Planning. Through their own commitment and effort, and with the support of the TIF-LEAP team, many teachers and principals say that they came to view SLOs as "not merely objectives as usual." They found that SLOs had the potential to impact the outcome of student learning through data-informed planning. They also felt that,

through the process, teachers generated better information about what students already knew or needed to know, thereby helping them focus and refine their teaching.

The TIF-LEAP team described[9] how SLOs measure teacher performance in areas that affect student learning, finding evidence that SLOs demonstrate how a teacher diagnoses student learning needs and progress and uses that knowledge to target effective interventions and set measurable expectations for student growth. Further, SLOs measure the depth of a teacher's knowledge of content and pedagogy as well as his or her skill in selecting and using instructional strategies that engage students, including reflecting on one's practice and adjusting, where needed.

Additionally, SLOs improve a teacher's use of measures, including the increasing use of multiple forms of measurement; the attention to the alignment and rigor of the assessments; the more frequent use of assessments to diagnose and provide student feedback; and the expanded teacher capacity to evaluate the quality and reliability of assessment instruments and results.

Observations made by the TIF-LEAP team and evaluators about the use of the SLO process and its impact on teacher practice resonate with research that finds "effective teachers are able to:

1. understand subject matter deeply and flexibly;
2. connect what is to be learned to students' prior knowledge and experience;
3. create effective scaffolds and supports for learning;
4. use instructional strategies that help students draw connections, apply their learning, practice new skills, and monitor their own learning;
5. assess student learning continuously and adapt teaching methods to students;
6. provide clear standards, constant feedback, and opportunities for revising work; and
7. develop and manage a collaborative classroom where all students have membership."[10]

Linking SLO Processes to Teacher Evaluation. During 2010-11, CMS implemented the state's new teacher evaluation system, known as the North Carolina Teacher Evaluation Process (NCTEP). This system enumerates 25 elements of teacher practice for evaluation.[11] The TIF-LEAP team conducted an analysis showing that "classroom observation captures 17 of the elements [of teacher practice], and the current SLO model in use in CMS captures 18 of the elements so that the combined usage of SLOs and classroom observations provide a more complete assessment of teacher performance."[12]

Because the new evaluation tool emphasizes the teacher's collection of evidence, the SLO process, with its focused use of baseline data and documentation of attainment, advantaged TIF-LEAP teachers and principals. Linking the SLO process and teacher evaluation practices, according to teachers and principals, enriched both efforts.

Analyzing and Auditing SLO Implementation for Improvement. As the project was underway, the TIF-LEAP team conducted various analyses and audits that provided information for discussion with the SLO Design Team, Steering Committee, and TIF-LEAP Principals' Working Group, as well as suggesting areas for school-level improvements. A principal interviewee extolled the benefits of the TIF-LEAP team's summary of the types of instructional strategies used to implement SLOs in his school. He used the information to generate a discussion in his school about the overuse of strategies that lead to convergent thinking (lecture, recitation, right answer questioning) and the underuse of strategies that develop divergent thinking (problem-based learning, collaboration, open-ended questions). Later iterations of the web application for inputting SLOs provided teachers with prompts through a pull-down menu of teaching strategies.

At the conclusion of year four, when a random sample of pre- and post-assessments of SLOs red flagged a higher than expected error rate, a formal audit of assessments was conducted by the TIF-LEAP team. As the audit showed, improving and standardizing assessment practices in schools is daunting work and maintaining fidelity to a high standard of practice during times of change requires extra vigilance from school and district leaders, or

quality assurance may be lost. A series of checks were added to the SLO software as a follow-up.

Overall. As a performance-based compensation approach, SLOs are grounded in practices associated with effective teaching (as listed above), with the intent of assisting all teachers to realize a positive result with their students—and earn a bonus award. As an instructional reform, the SLO approach is based on a design that is rooted in learning research and theory. SLO implementation relies on attention from district leaders, particularly (1) in delivering critical services to teachers, such as timely and accessible student data; (2) in improving the assessment tools and practices within schools; and (3) in providing more targeted professional development in the school setting.

Approach Two: Value-Added Measure

Following the implementation of Student Learning Objectives in 2008-09, the leadership of Charlotte-Mecklenburg Schools moved forward in 2009-10 by instituting a district-developed value-added measure, or VAM—a statistical analysis conducted to measure a teacher's contribution to the growth of his/her students on the annual End-of-Grade or End-of-Course assessments and then compare it with teachers of similar grades and subjects. The CMS Accountability Department defined the value-added measure developed in CMS as "a rating of the relative performance of a teacher after adjusting for conditions that are outside of a teacher's control (e.g., the number of student absences, the number of special needs students in a classroom, and the number of students with behavioral issues)."[13]

Value-Added Measure: Research and Best Practice

Value-added measures in education are an accountability reform, grounded in the concept of identifying teacher effectiveness in generating student growth, usually based on state assessments. Assessing teacher performance using a set of statistical protocols known as "value-added" has become an increasingly attractive tool for state and district policymakers in recent years. More than comparing a linear string of annual assessment results, value-added measures focus on the amount of academic growth of students over their prior year achievement and seek through statistical analyses to estimate teacher impact on that growth, holding constant a set of student, classroom, and school variables that also influence learning outcomes.[14]

Among the reasons for increased interest in value-added models[15] is that after more than a decade of trying to make gains on the federal adequate yearly progress (AYP) targets, educational leaders are seeking more powerful analyses regarding teacher and student outcomes, ones that go beyond measuring the proficiency level of students on state assessments. Additionally, federal and state policymakers and some districts, concerned about teacher appraisal and compensation systems that fail to discriminate among levels of teacher effectiveness, have turned to value-added with the expectation of bringing more objectivity and quantitative analysis to an evaluation process that has been subjective, anecdotal, and often disconnected from school and district personnel decisions.

There is not one value-added method. Researchers debate the validity and proper uses of the various VAM methodologies. In particular, there is continued examination of the assumption that student assessments, even those considered reliable and credible, can be used accurately to measure teacher effectiveness.[16] A value-added methodology attaches a number, in scale points, to each teacher, indicating how that teacher's performance compares to the performance of his or her peers in terms of student growth. Problems often arise from calling the value-added ranking a measure of teacher effectiveness, implying that there is a causal relationship. Some researchers question such a claim, arguing that statistical theory calls for the random assignment of students to a class and the random assignment of teachers to classes, not practices in any school system.[17]

Other issues relating to the methodology include score inconsistencies, such as the inability to adjust the VAM to account for teachers with a disproportionate number of learning challenges within a classroom population or to account for the impact of "short-run" test prep activities, conducted by many teachers and schools. Other

inconsistencies show up with the use of different assessments or statistical models.[18] However, most criticism is related to inappropriate uses of the value-added information, particularly in high stakes decisions associated with teacher evaluation and compensation. Release of scores to the news media, which have led to teachers' being confronted in their homes to explain their scores,[19] also raise concerns.

There is recognition, though, among researchers, policymakers, and educational leaders that value-added methodologies have opened a new vein of information to mine for policymaking and employment decisions, albeit with cautions. Some consensus about the use of value-added methodologies and the evaluation of teacher performance coalesces around the following: (1) student achievement, when used as a component of teacher compensation and evaluation decisions, should not be based on simple year-to-year comparisons; (2) value-added methodologies are preferable to simple comparisons; (3) value-added analyses in practice are not adequate (accurate, consistent, replicable) for use as the sole measure of individual teacher effectiveness; (4) misuses and abuses of the value-added information, such as releasing teacher names and ratings to the public, are best avoided by the profession; and (5) valued-added methodologies can be useful in evaluating the results of other measures of teacher performance.

Value-Added Measure: CMS Design for TIF-LEAP Schools

In the spring of 2009, the Superintendent asked the Steering Committee to include a value-added measure of teacher performance in the initiative. Accordingly, the Accountability Department convened a Growth Measure Task Force which met over a two-month period beginning in the late spring and included Accountability Department staff, TIF-LEAP team members, others from the Steering Committee, and a principal. Later, a TIF-LEAP teacher was added to the task force.

The task force examined a simple value-added method (using only prior year achievement) and a multilevel model (accounting for the impact of other variables) to measure student growth. Feedback on the design of the value-added approach was sought from the Steering Committee, area superintendents, and the TIF-LEAP Principals' Working Group.

Minutes from these meetings show that the concept of a growth measure based on a three-tier statistical model and a proposed altered bonus payout structure were presented, along with possible thresholds and rankings that might be used. Responses primarily centered on the potential negative effects of reducing SLO bonuses, and on teacher eligibility for VAM based on the grades and subjects taught. Because the funding was coming from the TIF-LEAP grant, the Steering Committee created a new payout schedule to include individual and school VAM bonuses, which entailed decreasing the amount of SLO bonuses for teachers and principals.

As finalized, the district-developed value-added measure is a multilevel model where students are nested within classrooms, which are nested within schools. Both the classrooms and teachers are estimated as random intercepts with covariates at the student, classroom, and school level. This VAM approach controls for factors outside the teacher arena that may influence student test scores, removing the effect of these factors in the analysis in order to isolate the teacher's contribution to student growth. *Table II.2* lists the published covariates used in the value-added measure for years three, four, and five of the initiative.

The VAM analysis is based on student performance data for all teachers of the tested subjects and grades in the district, but individual VAM bonuses were only available for eligible teachers in the twenty TIF-LEAP schools. The value-added score is expressed as a percentile ranking for each teacher indicating how he or she compares to other teachers in the district teaching the same grade and content.

Value-Added Measure: Implementation in TIF-LEAP Schools

The valued-added approach to measuring teacher performance for the purposes of compensation focuses on teacher output, which means measuring and analyzing a teacher's effect on student growth on the End-of-Grade/End-of-Course assessments. As an approach, implementing a

value-added measure of teacher performance does not call for structured changes in teacher practices. Rather, it works on the assumption that a bonus or performance rating will incentivize teachers to improve student achievement in their classes. However, the value-added measure was implemented in the TIF-LEAP schools where the implementation of SLOs—which do call for structured changes in practice—was already in progress.

Compared to the SLO implementation, the VAM implementation would appear to have been relatively uncomplicated, but a critical task in implementing the value-added approach was to communicate effectively with participants since performance ratings and compensation both are high stakes for professionals and the calculation and use of a VAM can be difficult to understand. Led by the Accountability Department, communication about the VAM in TIF-LEAP schools occurred in three phases: first, in meetings conducted with standing district groups (i.e., principals) for the purposes of educating and gathering feedback on the proposed VAM approach; secondly, with the teachers and principals in the TIF-LEAP schools where the VAM approach would be piloted; and finally, through explanations and videos posted to the district website.[20]

The TIF-LEAP schools did not self-select into a VAM-based bonus initiative. However, as detailed in Chapter VI, multiple years of survey and interview data show teachers and principals

TABLE II.2

Covariates Included in Teacher Value-Added Measure Analysis, 2009-12

Student	Classroom	School
Gender	% Male	% Male
Age	Average Age	Average Age
English Fluency	% English Fluent	% English Fluent
Exceptional Child	% Exceptional Child	% Exceptional Child
Repeating Grade	% Repeating Grade	% Repeating Grade
First Year in the School	% First Year in School	% First Year in School
Test Score (Prior Year)	Average Test Score (Prior Year)	Average Test Score (Prior Year)
Test Score (2 Years Ago)		
Days Absent (Prior Year)	Average Days Absent (Prior Year)	Average Days Absent (Prior Year)
# Discipline Incidents (Prior Year)	Average Discipline Incidents (Prior Year)	Average Discipline Incidents (Prior Year)
Days OSS* (Prior Year)	Average Days OSS (Prior Year)	Average OSS Days (Prior Year)
Days ISS* (Prior Year)	Average Days ISS (Prior Year)	Average ISS Day (Prior Year)
McKinney-Vento Status**		% School EDS***
Student Mobility		% School Mobility
Grade (e.g., 4th)	Class Size	School Type (ES, MS, HS)
Year (e.g., 2009)		School Size
Academically Gifted	% Academically Gifted	% School Academically Gifted

* Categories of Student Suspensions from School (OSS: out-of-school suspension; ISS: in-school suspension)
** Federal Assistance for Homelessness
*** Economically Disadvantaged Students

in TIF-LEAP schools were open to the concept of performance-based compensation and to the use of student academic growth measures. Even so, many teachers claim not to have understood the VAM, including how it was computed. Further, they felt that the introduction of the VAM was not transparent, of particular concern because of a commitment from district leadership that the TIF-LEAP initiative would be implemented with teachers. When individual teacher VAM rankings were made public at the close of the first year of implementation, concerns and misunderstandings about the method, as well as mistrust about the accuracy and fairness of the rankings, became evident. Additionally, many teachers observed that the rankings provided too little information for purposes of improving instruction.

Two Performance-Based Compensation Approaches: Comparison, Analysis and Summary

Student Learning Objectives required that teachers adopt a set of planning practices, using baseline data, setting growth expectations and selecting quality content and assessments. Bonuses were based on the teacher and principal's agreement that the objective had been properly assessed and the growth target met. Implementation of SLOs required a significant ongoing dedication of the initiative's resources, as well as those of several district departments. Further, it required steady participation from school site leaders.

By comparison, implementation of the value-added measure was perceived as relatively uncomplicated in school-level implementation, involving an end-of-year analysis of the spring administration of the North Carolina assessments conducted by the Accountability Department. However, the VAM required more input, communication, and buy-in than took place in the relatively small window between the decision to develop a VAM and its implementation.

Both performance-based compensation approaches in CMS were susceptible to "bumps in the road," such as software or communication issues, which affect the course of and/or the conceptual fidelity of the initiative. Other salient issues such as systemwide layoffs, school closures/reorganizations, and a new state teacher evaluation process clearly impacted participants professionally and personally. The U.S. Department of Education's evaluation of the first two TIF cohorts indicates that the TIF-LEAP initiative in CMS was not alone in implementation complexities:

> *"The experience of the first two cohorts of TIF grantees underscored the technical, cultural, and contextual complexity of compensation reform. Projects were implemented by these grantees in varying local contexts with shifting leadership, policy, and reform agendas. Many grantees reported having to rebuild their data systems, build understanding and support from educators for the new system, and add new evaluation responsibilities to administrators or accomplished teachers. In addition, many grantees had to develop support systems that would allow educators to make the changes necessary to succeed under a new compensation system. Moreover, grantees had to confront traditional attitudes and beliefs about how educators should be judged and differentiated."* [21]

While lamenting bumps along the implementation road, it is worth noting that the complications that should be given the most scrutiny—for the purpose of making corrections or modifications or not repeating mistakes—are those inherent in the design or ones resulting from inadequacies in the implementation plan. An instructive way to do this is to consider the features of the design and implementation in light of the original proposal submitted to the U.S. Department of Education. It was guided by the Six Cornerstones of Performance-Based Compensation which state that successful compensation reform: (1) is systemic reform; (2) is done with teachers and not to them; (3) must be organizationally sustainable; (4) must be financially sustainable; (5) requires a broad base of support in the district and community; and (6) goes beyond politics and finances to benefit students.[22]

A critical feature analysis (*Table II.3*) compares the design and implementation of the two approaches to compensation in the TIF-LEAP schools in relation to several of the key principles of compensation reform.

TABLE II.3

Critical Feature Analysis of Compensation Reform Principles

Feature	Student Learning Objectives	Value-Added Measure
Teacher/principal participation in the development of the compensation approach strengthens it and builds commitment.	• Focus groups conducted in year prior to adoption • Teacher/principal members on Steering Committee • Teacher/principal members on SLO Design Team and other TIF-LEAP working groups • Annual parent and community survey and parent focus groups	• A principal and later a TIF-LEAP teacher member on the Growth Measure Task Force
School opt-in also increases commitment with teachers making knowledgeable choices.	• High need schools selected for participation by district leaders; no formal school opt-in	• VAM implemented in same district-selected schools as SLOs; no school opt-in
Eligibility refers to who actually participates and receives a bonus. Opportunity, clarity, consistency and fairness are important in determining eligibility.	• All classroom teachers with satisfactory evaluations and attendance • All principals, initially, who work with SLOs and teachers	• Top 30% of classroom teachers of EOG/EOC assessed grades and subjects with satisfactory evaluation and attendance • Top 40% of schools eligible
Communication informs teacher decisions; clarifies expectations; builds trust; and when two-way, provides feedback for improvement or correction.	• Communication plan, including website, press releases • Annual parent and community survey; feedback from annual educator survey and interviews • TIF-LEAP team's 24-hour turnaround policy	• Meetings with standing groups • Meetings on campuses • VAM component added to the TIF-LEAP orientations • VAM videos on the website
Teaching/learning intervention(s) and/or professional development opportunities are designed and implemented to help teachers increase student achievement and earn bonuses.	• Teacher-developed objectives, using student baseline data, setting growth targets, planning instruction, and selecting/conducting assessments	
Quality assessments and rigorous assessment practices are critical for measuring student growth in a credible and reliable manner.	• Teacher-selected/principal-approved pre/post assessments of student results • Protocols guide assessment approval, practices • Analyses/audits conducted to provide improvement in teaching and assessment	• Uses student growth on EOG/EOC assessments as basis of district value-added analysis in years 3 and 4 • Uses NC ABC growth measure in year 5
Principal leadership is a priority, requiring high expectations for participation and adequate professional development in the administration of both approaches.	• Principal turnover and/or opt out was 40% in last three years • Professional development was informal via meetings • Supervision of the quality of the principal role in SLOs was uneven	• Principals dissatisfied with changes in bonus amounts that resulted from VAM implementation • Principals struggled to explain and work with the VAM rankings
Evaluation of a compensation reform initiative should be based on student outcomes, teacher and principal feedback, and artifacts of the process.	• Outside evaluator • Student growth compared to match schools • Rubric evaluation of SLO artifacts • HLM analysis of SLO ratings and student achievement • 5 years of survey/interview data	• Questions about VAM were added to annual CTAC surveys and interviews for feedback
Organizational and fiscal sustainability is important because it creates the conditions for success and impacts trust, commitment, and long-term reform.	• SLOs incorporated into best practices in CMS curriculum and instruction • SLOs as a measure of teaching effectiveness (talent) incorporated into new district plan and NC measures of teacher performance • Bonus payouts managed through fiscal crisis; impacted long-term sustainability	• The VAM model or district-developed growth measure was discontinued • The district will implement a VAM under development by the state of North Carolina

Summary

The discussion of implementing two approaches to performance-based compensation in the TIF-LEAP schools describes the theory, design, and implementation benefits and challenges of each approach. Both approaches have strengths and the decisions to adopt the two approaches were significant for the district and schools. However, they differed markedly in their methods of introduction to participants and in the level of acceptance by participating schools.

Endnotes

1 Chait, R. and Miller, R. (2009). *Paying teachers for results: A summary of research to inform the design of pay-for-performance programs in high poverty schools.* Center for American Progress. Retrieved from http://www.americanprogress.org.

2 The Teacher Incentive Fund refers to bonus pay as merit-based supplements.

3 See http://denverprocomp.dpsk12.org/.

4 Slotnik, W., Smith, M. *et al.* (2004). *Catalyst for change: Pay for performance in Denver.* Boston, MA: Community Training and Assistance Center.

5 See, for example: Marzano, R. *et al.* (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement.* Alexandria, VA: Association for Supervision and Curriculum Development. See also: Darling-Hammond, L. and Bransford, J. (2005). *Preparing teachers for a changing world: What teachers should know and be able to do.* San Francisco: Jossey-Bass.

6 Slotnik, W. and Smith, M. (2008). *Tying earning to learning: The link between teacher compensation and student learning objectives.* Boston, MA: Community Training and Assistance Center.

7 Locke, E. and Latham, G. (2002, September). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 706.

8 Password access to student data had been available to teachers for a time, but many teachers had not had training or practice in navigating the system and some had forgotten their passwords. Mostly, they did not expect student data to be up-to-date, or even, in some cases, that they should use data in their individual planning. Principals were more knowledgeable and skilled with data access.

9 Adapted from Norwood, S. *et al.* (2011). *The role of student learning objectives*. Charlotte-Mecklenburg Schools: Office of TIF-LEAP Initiative.

10 Darling-Hammond, L. and Bransford, J. (2005).

11 http://www.ncpublicschools.org/docs/educatoreffect/ncees/instruments/teach-eval-manual.pdf.

12 Norwood, S. *et al.* (2011). *The role of student learning objectives.* Charlotte-Mecklenburg Schools: Office of TIF-LEAP Initiative.

13 The VAM description and covariate chart (Table II.2) are adapted from (1) the TIF-LEAP initiative description at www.cms.k12.nc.us; (2) a brief "CMS TIF-LEAP Value-Added Measure of Teacher Effectiveness;" (3) "Notes for Ann Helms of the *Charlotte Observer* on Value-Added Calculation" at www.cms.k12.nc.us; and (4) TIF-LEAP teacher training materials.

14 Ewing, J. (2011). Mathematical intimidation: driven by the data. *Notices of the American Mathematical Society*, *58*(5) at http://www.ams.org/notices/201105.

15 Steele, J. *et al.* (2010). *Incorporating student performance measures into teacher evaluation systems*. Rand Corporation at http://www.rand.org.

16 Koppich, J. and Rigby, J. (2009). *Alternate teacher compensation: a primer.* Policy Analysis for California Education (PACE) at http://www.stanford.edu/group/pace.

17 Braun, H. I. (2005). *Using student progress to evaluate teachers: A primer on value-added models*. Retrieved from Policy Information Center, Educational Testing Service, Princeton, NJ. Website: http://www.ets.org/research/policy_research_reports/publications/report/2005/cxje.

18 Darling-Hammond, L. *et al.* (2012). Evaluating teacher evaluation. *Phi Delta Kappan*, *93*(6), 8-15.

19 Ewing, J. (2011).

20 Value Added Research Center. The Oak Tree Analogy. (Power Point slides). Retrieved from http://varc.wceruw.org/tutorials/oak/index.htm.

21 Humphries, D. *et al.* (2012). *Teacher incentive fund: First implementation report, 2006 and 2007 grantees*. U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service: Washington, D.C.

22 Slotnik, W. (2009, July 15). Get performance pay right: Six cornerstones of successful compensation reform. *Education Week*, *28*(36), 26-32.

CHAPTER III

Evaluation Design and Methods of Analysis

Over the five-year course of the TIF-LEAP initiative, the Community Training and Assistance Center (CTAC) annually collected and reviewed a wide array of information relating to the goals, status, and outcomes of the initiative. CTAC applied a mixed methods approach in analyzing these data for the purposes of (1) providing interim progress reports to the Superintendent of Schools; (2) generating periodic reports in compliance with regulations of the Teacher Incentive Fund; and (3) providing a final, comprehensive evaluation of the initiative's impact on teacher performance and student achievement. The design of the final evaluation, the scope and sources of the data included in the analysis, and the methods of analysis used for the study are summarized in this chapter.

Evaluation Design

The evaluation of the TIF-LEAP initiative is based on a quasi-experimental design that examines the effect of the initiative on the selected schools rather than schools randomly assigned to either the TIF-LEAP or comparison school groups. To minimize the potential threats to validity often raised when using a quasi-experimental design (i.e., the TIF-LEAP schools had incentive funds), CTAC collected a substantial volume and variety of data for both TIF-LEAP and comparison schools at regular points over five years, and an array of analyses have been employed in order to discern the impact of the initiative on teacher performance and student achievement.

Selection Criteria for TIF-LEAP and Comparison Schools

As discussed in Chapter I, district leadership had an interest in implementing a performance-based compensation initiative in its highest need schools. The participating schools selected by the district were phased in, reaching a total of 20 schools by the third year of the initiative. In the fifth and final year of the initiative, school closures, reorganizations, and repurposing reduced the number of participating schools to eleven (11).

To select comparison schools for the TIF-LEAP evaluation study, the CMS Accountability Department employed a hierarchical cluster analysis to "cluster" similar elementary, middle, and high schools through the following demographic variables: school size, gender, ethnicity, free and/or reduced lunch, proficiency on the annual North Carolina EOG/EOC assessments, and designations of English proficiency, giftedness, and/or disabilities. Based on these cluster analyses, dendrograms were developed and coefficients assigned to schools based on their similarities to the original sixteen TIF-LEAP schools. Those with the smallest coefficient (greatest similarity) were selected.[1]

The method of selecting TIF-LEAP schools, and later the need to select comparison schools based on matches with TIF-LEAP schools, created some complications for the study of the impact of the intervention on student achievement. These issues are discussed in the analysis of student achievement in Chapter V. The TIF-LEAP and comparison schools are shown in *Table III.1*.

TABLE III.1

TIF-LEAP Schools and District-Selected Comparison Schools*

Elementary Schools	Year Entered	Comparison School One	Comparison School Two
Berryhill	09-10	Huntingtown	Pinewood
Billingsville	07-08	Allenbrook	Westerly Hills
Druid Hills	08-09	Bruns Academy	Walter G. Byers
Highland Renaissance	08-09	Sedgefield Elementary	Greenway
Lincoln Heights**	09-10	Winding Springs	Statesville Road
Reid Park	08-09	Ashley Park	First Ward
Shamrock Gardens	07-08	Sedgefield Elementary	Rama Road
Middle Schools	**Year Entered**	**Comparison School One**	**Comparison School Two**
Bishop Spaugh**	07-08	Ranson	Cochrane Collegiate
John Taylor Williams**	08-09	Coulwood	Ranson
J.W. Wilson**	07-08	Cochrane Collegiate	Eastway
Martin Luther King, Jr.	07-08	Albermarle Road Middle	Cochrane Collegiate
Sedgefield Middle	07-08	Cochrane Collegiate	Albermarle Road Middle
High Schools	**Year Entered**	**Comparison School One**	**Comparison School Two**
E.E. Waddell**	09-10	Vance	East Mecklenburg
Garinger Small School System	09-10	Olympic Small School System	Olympic Small School System
West Charlotte	09-10	Vance	Independence
West Mecklenburg	09-10	Vance	East Mecklenburg

* Due to an insufficient number of schools that were similar enough to the TIF-LEAP schools, some schools served as comparison to more than one TIF-LEAP school.

** These schools were closed at the beginning of the 2011-12 school year.

TABLE III.2

Evaluation Questions

Evaluation Question I

As a performance-based compensation approach in CMS, did Student Learning Objectives contribute to/measure teacher performance effectiveness?

Evaluation Question II

Did student achievement in TIF-LEAP schools improve compared to the designated match schools? As a performance-based compensation approach, how did Student Learning Objectives impact student performance in TIF-LEAP schools?

Evaluation Question III

How did the quality/fidelity of implementation of the performance-based compensation initiative in CMS impact outcomes for teachers and students?

Evaluation Question IV

What changes in stakeholder perspectives of performance-based compensation occurred over the life of the initiative? In the effectiveness of the TIF-LEAP initiative?

Evaluation Question V

Was the TIF-LEAP initiative able to realize its goals through the implementation of performance-based compensation?

Evaluation Study Questions

This evaluation is a study of the impact of the TIF-LEAP initiative on teacher performance and on student achievement. The broad questions developed for the study, and shown in *Table III.2*, embrace subsets of more focused questions in the search for understanding the impact on TIF-LEAP participants and students and the significance of the work carried out by Charlotte-Mecklenburg Schools in this compensation reform initiative.

Charlotte-Mecklenburg Schools Districtwide Demographic Data

To provide background, districtwide student and teacher data are highlighted below. Distributions of demographic data for TIF-LEAP and comparison schools can be found in Chapter V.

Districtwide Student Data

Charlotte-Mecklenburg Schools is a large district where more than 40% of the students

TABLE III.3

Student Characteristics by Year – Districtwide

	2007-08	2008-09	2009-10	2010-11	2011-12
Total Students	132,281	130,953	133,664	134,466	138,012
Ethnicity					
African American	55,678 (42%)	54,679 (42%)	55,121 (41%)	56,475 (42%)	56,585 (41%)
American Indian	673 (<1%)	612 (<1%)	592 (<1%)	1,345 (1%)	1,380 (1%)
Asian	5,828 (4%)	6,303 (5%)	6,488 (5%)	6,723 (5%)	6,901 (5%)
Hispanic/Latino	19,671 (15%)	19,977 (15%)	21,214 (16%)	21,515 (16%)	22,082 (16%)
Multi-Racial	4,445 (3%)	4,988 (4%)	4,988 (4%)	4,034 (3%)	5,520 (4%)
White	45,986 (35%)	44,394 (34%)	44,719 (34%)	44,374 (33%)	45,544 (33%)
Gender					
Female	65,611 (50%)	65,346 (50%)	66,164 (49%)	66,561 (49%)	67,625 (49%)
Male	66,670 (50%)	65,607 (50%)	67,500 (51%)	67,905 (51%)	70,387 (51%)
Socio-Economic Status					
Economically Disadvantaged	61,076 (47%)	62,974 (48%)	66,435 (50%)	69,608 (52%)	73,698 (53%)

are African American, one-third of the students are White, and 16% of the students are Hispanic/Latino. (See *Table III.3*.) Students from various ethnic groups are not distributed among the district's schools in direct proportion to the district averages. This also applies to the distribution of economically disadvantaged students.

Districtwide Teacher Data

Performance-based compensation programs frequently are intended to mitigate the effects of teacher assignment and transfer policies that place the newest and least experienced teachers in difficult-to-staff schools and allow transfers to higher performing schools after a probationary period. Districtwide data on teacher characteristics are shown in *Table III.4*. As with student characteristics (*Table III.3*), the distribution of teachers by characteristics among the district's schools is not proportional to the district averages.

Sources and Scope of Data Used in Evaluation

Student Achievement Data

The North Carolina Testing Program (www.dpi.state.nc.us/accountability/testing) has two parts: (1) North Carolina End-of-Grade (EOG) Tests, designed to measure student performance on the goals, objectives, and grade-level competencies specified in the *North Carolina Standard Course of Study*; and (2) North Carolina End-of-Course (EOC) Tests, used to sample a student's knowledge of subject-related concepts as specified in the *North Carolina Standard Course of Study* and to provide a global estimate of the student's mastery of the material in a particular content area. Over

TABLE III.4

Teacher Characteristics by Year – Districtwide

	2007-08	2008-09	2009-10	2010-11	2011-12
Total Classroom Teachers	9,116	9,095	8,781	8,429	8,876
Ethnicity					
African American	2,279 (25%)	2,364 (26%)	2,195 (25%)	2,107 (25%)	2,130 (24%)
American Indian	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Asian	91 (1%)	91 (1%)	88 (1%)	84 (1%)	89 (1%)
Hispanic	365 (4%)	273 (3%)	263 (3%)	253 (3%)	266 (3%)
White	6,381 (70%)	6,367 (70%)	6,235 (71%)	5,985 (71%)	6,391 (72%)
Gender					
Female	7,302 (80%)	7,312 (80%)	7,025 (80%)	6,768 (80%)	7,115 (80%)
Male	1,814 (20%)	1,783 (20%)	1,756 (20%)	1,661 (20%)	1,761 (20%)
Years of Experience					
0 to 3 Years	2,766 (30%)	2,547 (28%)	2,090 (24%)	1,849 (22%)	2,184 (25%)
4 to 10 Years	2,906 (32%)	3,021 (33%)	3,111 (35%)	3,005 (36%)	3,124 (35%)
11+ Years	3,444 (38%)	3,527 (39%)	3,579 (41%)	3,575 (42%)	3,568 (40%)
Selected Education Characteristics					
Advanced Degrees	2,604 (29%)	2,757 (32%)	2,793 (32%)	2,850 (34%)	3,547 (40%)
National Board for Professional Teaching Standards (NBPTS)	1,002 (11%)	1,102 (12%)	1,358 (16%)	1,494 (18%)	1,377 (16%)

TABLE III.5

North Carolina End-of-Grade (EOG) and End-of-Course (EOC) Student Achievement Data Collected During the TIF-LEAP Initiative, 2007-12

Assessment	Grade									
	3	4	5	6	7	8	9	10	11	12
Algebra I EOC							1-5			
Algebra II EOC									1-5	
Biology EOC								1-5		
Chemistry EOC									1-2	
Civics & Economics EOC								1-4		
English I EOC							1-5			
Geometry EOC								1-4		
Mathematics EOG	1-5	1-5	1-5	1-5	1-5	1-5				
Physical Science EOC							1-4			
Physics EOC										1-2
Reading EOG	1-5	1-5	1-5	1-5	1-5	1-5				
Science EOG			1-5			1-5				
US History EOC									1-5	
Writing-Argumentative EOG					1-2					
Writing-Informational EOG								1-5		
Writing-Narrative EOG		1-2								

Year 1 (2007-08); Year 2 (2008-09); Year 3 (2009-10); Year 4 (2010-11); Year 5 (2011-12)

the course of the initiative, some EOC assessments were dropped from the program by the State. *Table III.5* shows the EOG and EOC assessments and their corresponding grade levels that were administered during all or parts of the five years of the initiative.

Student achievement data from the North Carolina EOG/EOC assessments were collected and analyzed for initiative years 2007-08 to 2010-11 plus year 2006-07 for comparison purposes. These mathematics and reading/language arts assessments are administered in the spring of each year in grades 3-8 and for designated high school courses.

A careful examination of the year five (2011-12) database found that achievement data were unavailable for those students who had attended a TIF-LEAP school that was closed or reconfigured, which accounted for approximately one-third of the TIF-LEAP student population. In addition, several other gaps in the database resulted in a highly unstable and substantially unrepresentative sample. Together, these gaps precluded the ability to satisfactorily estimate effects of SLOs on student achievement in the initiative's final year (2011-12).

Other Student Achievement Measures

In the course of reporting annually to the Board of Education, the TIF-LEAP team compiled TIF-LEAP school averages for the following: (1) the Adequate Yearly Progress (AYP) reports, as calculated under No Child Left Behind; (2) the percent of students proficient, based on the North Carolina assessments itemized above; (3) the ABC Growth, a North Carolina value-added measure; and (4) the district value-added measure for the

applicable years. Where these data were available to the CTAC evaluators, they were included in analyses of the impact of the initiative and are discussed in Chapters IV and V. Dropout and English language learner data were not available to the evaluators.

District and TIF-LEAP Documents and Artifacts, 2007-2012

Relevant documents and artifacts of TIF-LEAP and district processes were collected over the course of the initiative. The primary artifact of the initiative is the SLO developed and implemented by the teacher. Beginning in the second year of the initiative (2008-09), most teachers in the participant schools developed two objectives. Nearly 4,000 SLOs were submitted by teachers in the participating TIF-LEAP schools: 707 in 2008-09, 1,904 in 2009-10 and 1,353 in 2010-11.

The CTAC evaluation team rated all of these SLOs over the course of the study using a four-point, validated rubric, where Level 4 is the highest level of quality. Over the three years of SLO implementation under study, 1,708 (43.1%) SLOs were at Level 4 Excellent, 1,964 (49.5%) were at Level 3 Acceptable, 283 (7.1%) at Level 2 Needs Improvement, and 9 were rated as Level 1 Too Little to Evaluate.

District and TIF-LEAP Participant Interviews, 2007-2012

The CTAC evaluation team conducted formal interviews and focus groups with TIF-LEAP leaders, participants, and stakeholders primarily during the winter of each of the five years of the initiative. *Table III.6* provides a summary of the interviewees by role.

TIF-LEAP teachers were invited for interviews or focus groups through a random selection process with the goal of including participants from all TIF-LEAP schools; non TIF-LEAP principals and teachers were selected randomly from diverse areas of the district. This latter group included principals and teachers from comparison

TABLE III.6

Number of Interview and Focus Group Participants,* 2007-12

Role or Role Group	2007-08	2008-09	2009-10	2010-11	2011-12
Board of Education	9	7	8	9	8
Superintendent	1	1	1	1	1
External/Business Partners		5	4	3	
Central Office Administrators	17	13	10	13	17
Area/Zone Superintendents		7	7	5	2
Principals/Assistant Principals: TIF-LEAP	6	8	21	20	10
Principals/Assistant Principals: Non TIF-LEAP	9	18	10	11	
Teacher Association Leaders	3	3	4	3	2
Teachers: TIF-LEAP	33	22	93	33	28
Teachers: Non TIF-LEAP	7	19	60	18	
Parents		13	40	16	25
Students		56	90	93	
TIF-LEAP Team			3	4	3
Total Participants	**86**	**172**	**351**	**229**	**96**

* Some participants served in multiples roles, i.e., member of district administration and Steering Committee; however, this is an unduplicated count.

and other district schools. In addition, members of particular role groups, such as the Board of Education, central administrators, teacher leaders, and TIF-LEAP principals and/or assistant principals were interviewed. A range of district schools was selected to sponsor parent and student focus groups. Parent focus groups were conducted in English and Spanish.

Interview questions for district and school personnel, TIF-LEAP participants, non TIF-LEAP principals and teachers and other role groups were developed to probe areas similar to those included on the educator survey (discussed below). Within the scope of these questions, interviewees and focus group members were provided latitude in expressing their opinions and concerns, offering suggestions for improvement, and generally sharing their thinking on issues of teacher compensation and, if a participant, also on the progress of the initiative.

District and TIF-LEAP Educator Survey Administrations, 2007-12

In cooperation with the TIF-LEAP office and Charlotte-Mecklenburg Schools, the CTAC team conducted an annual web-based survey of all district principals and teachers on district conditions and performance-based compensation, including questions specifically targeted to the views of TIF-LEAP participants related to the implementation of SLOs. More than twenty thousand surveys were analyzed across the five years.

Using a five-point Likert scale of Strongly Agree to Strongly Disagree, the first section of the survey invited principals and teachers from all of the district's schools to respond to queries about (1) educational beliefs and general knowledge of the goals of the TIF-LEAP initiative; (2) school conditions and supports relative to the initiative; and (3) beliefs and preferences related to performance-based compensation reform.

A summary of the educator survey responses, number of distributions or invitations (delivery to e-mail boxes), completions, and percentage of responses over the life of the initiative are shown in *Table III.7*.

For reasons of confidentiality, survey respondents were not asked to provide their names nor were e-mail addresses associated with educator responses. However, respondents were asked to identify their school affiliation and role (principal, assistant principal, teacher) and to answer questions about their years of experience and ethnicity. These self-reported data were used to verify the representativeness of the responses and are shown in *Table III.8*.

A second section was added to the educator survey in year two of the initiative and was directed only to TIF-LEAP school participants. It asked respondents to assess the substance, quality, and implementation of SLOs. From year to year, some items were added, modified, or removed to respond to specific priorities and changes in the district and TIF-LEAP schools, as well as to the developmental stages of the initiative. A final section of the survey provided respondents with the opportunity to offer open-ended comments.

TABLE III.7

TIF-LEAP Educator Survey Dates and Response Rates, 2007-12

Year	Start Date	End Date	Invites	Deliveries	Completed Responses	Response Rate
2007-08	11/21/2007	1/31/2008	10,486	10,427	6,134	58.8%
2008-09	12/03/2008	2/11/2009	9,839	9,839	5,887	59.8%
2009-10	12/21/2009	2/24/2010	9,618	9,618	5,563	57.8%
2010-11	02/03/2011	3/15/2011	8,268	8,140	3,830	47.1%
2011-12	12/11/2011	2/02/2012	8,847	8,847	2,295	25.9%
Total	**N/A**	**N/A**	**47,058**	**46,871**	**23,707**	**50.6%**

TABLE III.8

Educator Survey Respondents by Group, Role and Years of Experience in CMS

	2007-08	2008-09	2009-10	2010-11	2011-12
Total Surveys*	**6,131**	**5,887**	**5,561**	**3,830**	**2,272**
Respondents by School Group					
TIF-LEAP Schools	N=698	N=674	N=571	N=375	N=204
Principals	18 (2.6%)	18 (2.7%)	16 (2.8%)	14 (3.7%)	5 (2.5%)
Asst. Principals	27 (3.9%)	25 (3.7%)	16 (2.8%)	8 (2.1%)	3 (1.5%)
Teachers	653 (93.6%)	631 (93.6%)	539 (94.4%)	353 (94.1%)	196 (96.1%)
Non TIF-LEAP Schools	N=5,433	N=5,213	N=4,990	N=3,455	N=2,068
Principals	128 (2.4%)	122 (2.3%)	121 (2.4%)	89 (2.6%)	63 (3.0%)
Asst. Principals	183 (3.4%)	181 (3.5%)	134 (2.7%)	77 (2.2%)	36 (1.7%)
Teachers	5,122 (94.3%)	4,910 (94.2%)	4,735 (94.9%)	3,289 (95.2%)	1,969 (95.2%)
Years of Experience in CMS					
TIF-LEAP Schools	N=698	N=674	N=571	N=375	N=204
0 to 4 Years	377 (54.0%)	339 (50.3%)	232 (40.5%)	138 (36.8%)	70 (34.3%)
5 to 10 Years	173 (24.8%)	186 (27.6%)	189 (33.2%)	129 (34.4%)	72 (35.3%)
10+ Years	148 (21.2%)	149 (22.1%)	150 (26.3%)	108 (28.8%)	62 (30.4%)
Non TIF-LEAP Schools	N=5,433	N=5,213	N=4,985	N=3,455	N=2,068
0 to 4 Years	2,361 (43.5%)	2,183 (41.9%)	1,743 (35.0%)	1,002 (29.0%)	604 (29.0%)
5 to 10 Years	1,471 (27.1%)	1,485 (28.4%)	1,647 (33.0%)	1,253 (36.3%)	748 (36.2%)
10+ Years	1,601 (29.5%)	1,545 (29.6%)	1,595 (32.0%)	1,200 (34.7%)	716 (34.6%)

* Some respondents do not indicate school, role, years of experience resulting in differences in total responses between Tables III.7 and III.8.

Parent and Community Member Survey, 2007-12

The district conducts an annual telephone survey with parents of students of the Charlotte-Mecklenburg Schools and members of the community. Surveys were conducted in English and Spanish. A series of eight CTAC-developed items related to performance-based compensation for teachers and principals were included in the annual CMS parent and community survey, beginning in 2007-08. A ninth item was added the following year and continued in subsequent years. Over the five years of the initiative, the telephone survey was conducted annually with more than 400 parents and 400 community members.[2]

Parent Respondents. A total of 2,026 parent phone surveys were conducted across five academic years: 418 in 2007-08, 400 in 2008-09, 403 in 2009-10, 402 in 2010-11, and 403 in 2011-12. On average, primary parent respondents annually were African American, 40.3%; White, 36.8%; and Hispanic/Latino 15.5%. An average of 10.0% of the parent respondents chose to be interviewed in Spanish.

Community Member Respondents. A total of 2,007 phone surveys were conducted with community members across five academic years: 399 in 2007-08, 404 in 2008-09, 402 in 2009-10,

402 in 2010-11, and 400 in 2011-12. On average, community member respondents annually were African American, 26.5%; White, 58.1%; and Hispanic/Latino 10.0%. An average of 7.4% of the community member respondents chose to be interviewed in Spanish.

Parent and community member respondents share the following in common: most respondents are married and have lived in Mecklenburg County for 15 or more years. They have graduated from a four-year college, are between 35 and 44 years old, and are employed full time.

Other Data

CTAC collected other data for the evaluation from district sources: artifacts of the process;

TABLE III.9

Methods of Analysis

Quantitative Methods		
Method	**Purpose**	**Data**
Longitudinal Hierarchical Linear Modeling	Estimate the impact of the TIF-LEAP initiative on student achievement in TIF-LEAP and comparison schools	Student results on EOG and EOC student assessments, student demographic factors
Cross-sectional Hierarchical Linear Modeling	Analyze the relationships between the SLOs and student achievement in the TIF-LEAP schools	Student results on EOG and EOC student assessments, SLO ratings and attainment, school, teacher/ classroom and student factors
Chi-square Test of Association	Analyze the SLO ratings across the years, their attainment and their relationship to student achievement and receipt of a VAM bonus	Individual SLO ratings and their attainment; VAM scores and bonus payouts
Analysis of Variance/Independent Samples *t*-Tests	Analyze the differences between VAM scores and SLO ratings and their attainment; Analyze the difference between various groups and their responses to the educator and parent/ community member surveys	Individual SLO ratings and their attainment; VAM scores and bonus payouts; Responses to survey questions and their Rasch conversions
Principal Axis Factor Analysis	Analyze survey responses to identify items which cluster together to form scales	Educator survey responses
Cronbach's Alpha Reliability	Analyze the reliability of the six scales resulting from the factor analysis	Educator survey responses
Rasch Modeling	Analyze the person and item characteristics of the scales of the educator survey	Educator survey responses

Qualitative Methods		
Method	**Purpose**	**Data**
Grounded Theory	Analyze the responses from individual interviews and focus groups for common beliefs, assumptions, propositions, and understandings	Responses from individual interviews and focus groups across the five years
Constant Comparative Method	Analyze the responses from individual interviews and focus groups across the years by multiple readers	Responses from individual interviews and focus groups across the five years

training materials; communications; student demographic data; data or indicators related to student achievement in addition to the EOC/EOG data; human resource, payroll, and financial data related to the goals and conduct of the initiative and federal reporting requirements of the Teacher Incentive Fund.

Additionally, routine observations on the status and progress of the initiative through regular CTAC technical staff visits; Steering Committee meeting minutes and agenda; Working Group agenda and observations; and conversations and meetings with TIF-LEAP and district staff, school participants, and members of the Board of Education augmented other data.

Given the goals and design of this TIF-LEAP evaluation, the availability of teacher retention data would have benefited the study. However, only data related to teachers' date of hire into the district and outward mobility were available. No record was available regarding teacher mobility within the district. Teacher turnover rate was provided in the North Carolina school report cards. These data were gathered for the TIF-LEAP and comparison schools.

Methods of Analysis

A mixed-methods approach was used to analyze the wide variety of data collected over the five years of the initiative ranging from such quantitative methods as hierarchical linear modeling, analysis of variance and *t*-tests to chi-square test of association as well as qualitative methodologies such as the constant comparative method. For the purpose of validating the educator and parent/community surveys conducted each year, factor analysis and Rasch modeling methodologies were employed. The study's methods of analysis and their purposes are outlined in *Table III.9* and discussed below.

Quantitative Research Methods

Analysis of Impact of TIF-LEAP on Student Achievement

The impact of the initiative on student achievement results was analyzed using two hierarchical linear modeling methodologies. Longitudinal and cross-sectional hierarchical linear models (HLM) were created to estimate the impact of the TIF-LEAP initiative on student achievement. These methodologies are used to account for the nested nature of the data. HLM estimates relationships at the individual and group (class, school, or program) levels and allows for relationships at the student level to vary across group levels, as well as for modeling cross-level interactions.

- Longitudinal hierarchical linear modeling was used to compare student achievement over time for students in the TIF-LEAP schools to those in the comparison schools.
- Cross-sectional hierarchical linear modeling was used to analyze the relationships between the SLOs and student achievement in the TIF-LEAP schools. Detailed information regarding these two methodologies is provided in Chapter V.

Analysis of Quality and Attainment of Student Learning Objectives

Each SLO is individually rated using a validated four-level holistic scoring rubric based on the four criteria of Learning Content, Completeness, Cohesion, and Expectations. The scoring rubric integrates these four criteria and the four levels of quality described in *Table III.10*.

TABLE III.10

Student Learning Objective Quality Levels

Level 4 Excellent
The Student Learning Objective meets all criteria.
Level 3 Acceptable
The Student Learning Objective meets most criteria with some lack of clarity in expectations and/or cohesion among components.
Level 2 Needs Improvement
The Student Learning Objective meets some criteria but may be incomplete, lack a thoughtful use of baseline and assessment data and/or lack cohesiveness.
Level 1 Too Little to Evaluate
The Student Learning Objective meets none of the criteria and/or may show a lack of understanding of the process or intention to complete.

The CTAC team of six educators/researchers reads and rates all SLOs. Each SLO is rated separately as Level 1, 2, 3, or 4 by two raters; additional analyses are used to resolve differences between ratings. The percent of exact agreement was 83.5% while the percent of adjacent agreement was 100.0%.[3] Inter-rater reliability of the SLOs was assessed using Cohen's Kappa; kappa = 0.53; 95% confidence interval: 0.387 to 0.663; statistically significant at $p < .001$.

- Chi-square Test of Association is used to assess the quality of the SLOs across the three years (i.e., 2008-09, 2009-10 and 2010-11), the extent to which the SLOs are met or not met across the three years, and the relationship between SLO quality and the attainment classification (i.e., met/not met). This non-parametric statistic is used to answer questions with variables measured with nominal or ordinal data such as the SLO ratings which are on a 4-point scale from Excellent to Too Little to Evaluate or the dichotomous met or not met designation. The null hypothesis is that there is no relationship between the two variables or groups; therefore, when the chi-square is statistically significant it indicates that there is an association between/among the variables in the analysis.
- Analysis of Variance (ANOVA) is employed to determine the relationship between the VAM score/ranking and the quality of the SLOs and whether they were attained. This parametric statistic is used to answer questions regarding the difference between two or more groups on a continuous variable. The null hypothesis is that there is no difference between the two groups, such as the question of the difference in the ratings for those SLOs that were met compared to those that were not met. The homogeneity of variance or the degree to which the variances are different across the groups is tested using Levene's test; if significant, Welch's robust test of equality of means is used to confirm the finding of the ANOVA. When there were only two independent groups, independent samples *t*-tests were used to determine the relationship between the VAM bonus or ranking and SLO attainment.

Analysis of Perceptions of Educator Survey Respondents

The educator survey was administered in all of the CMS schools (i.e., TIF-LEAP and non TIF-LEAP) over five years starting with 2007-08. The survey administered in the fifth year (2011-12) was used to anchor all the other years' surveys.

- Principal Axis Factor Analysis with varimax rotation is conducted to confirm the factor structure of the survey and as a data-reduction technique so that the 51 common questions can be described as the six scales they measure. The aim of factor analysis is to seek parsimony in the description of the survey results.[4] Reliability analyses using Cronbach's alpha confirmed the cohesiveness of the scales for further analyses.
- Rasch Modeling, a form of Item Response Theory modeling, is used to convert the ordinal responses on the survey (i.e., strongly agree to strongly disagree) to a measure of agreement with the scale items that have interval properties.[5,6] In Rasch modeling a scale is established with a standard mean and standard deviation (i.e., with an item mean of zero and a standard deviation of one). For the purpose of this survey, a scale was selected with a mean of 500 and a standard deviation of 100. The higher the mean level of agreement of the respondents, the more strongly the respondents endorse or agree with the items. The lower the mean level of agreement, the less the respondents endorse or agree with the items. The outcome of the modeling process is two estimates: one of person "ability" (i.e., level of agreement of that person with the survey scale) and one of item "difficulty" (i.e., the measure of how easy or difficult it is to agree with the item/question).
- Analysis of Variance, as described previously, is conducted to determine the relationship between several group variables and the scores on the six scales as well as the resulting interaction effects. When appropriate, *post hoc* analyses are conducted.

Qualitative Research Methods

Analysis of Interview and Focus Group Responses

Two qualitative analysis methodologies were employed to analyze the interview and focus group responses across the years.

- Grounded Theory is employed to analyze the responses to interviewees and focus group participant comments. The defining characteristic of grounded theory is that the hypotheses or assumptions are derived from the data rather than in response to propositions stated *a priori*. In other words, the generalization or theory emerges from the data themselves.[7] In the case of the present evaluation, responses from interviewees and focus group participants were reviewed and explored for beliefs and understandings which were shared across participant groups and which may be consistent from one year to the next or which may change or evolve from one year to the next.
- The Constant Comparative Method complements grounded theory as a method for developing theory based upon the data one has gathered. Initially data is compared with other data gathered such as survey data to find similarities and differences. Responses are compared across groups and over time in conjunction with events that have occurred over the five years of the initiative to find changes in opinions, observations, and perspectives. Multiple readers participate in this analysis.[8, 9]

Summary

The evaluation of the TIF-LEAP initiative examined the implementation and impact of performance-based compensation on student achievement in the participant schools, using a variety of data sources and methodologies. CTAC evaluators regularly gathered and studied the ideas and perspectives of TIF-LEAP participants, district and initiative leaders, students, parents and community members on the progress, quality, and significance of the TIF-LEAP initiative. Furthermore, the evaluation examined school level and broader institutional factors and their level of influence on the initiative.

Endnotes

1 Report from the Center for Research and Evaluation (2007), *CMS Research Brief* at www.cms.k12.nc.us.

2 The sample of parents and community members surveyed is not the same each year.

3 Penny, J., Johnson, R. L., & Gordon, B. (2000). The effect of rating augmentation on inter-rater reliability: An empirical study of a holistic rubric. *Assessing Writing,* 7(2), 143-164.

4 Gable, R. K, & Wolf, M. B. (1993). *Instrument development in the affective domain: Measuring attitudes and values in corporate and school settings.* 2nd ed. Boston: Kluwer Academic Publishers.

5 Rasch, Georg. (1960). *Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests.* Oxford, England: Nielsen & Lydiche. xiii 184 pp.

6 Wright, B. D., Stone, M. H. (1979). Best Test Design. Rasch Measurement. MESA Press, 5835 S. Kimbark Avenue, Chicago, IL.

7 Mertens, D. M. (2010). *Research and Evaluation in Education and Psychology: Integrating Diversity with Quantitative, Qualitative, and Mixed Methods,* (3rd ed.). Los Angeles: Sage Publications. 236 pp.

8 Charmaz, K. (2005). *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis.* Los Angeles: Sage Publications. 54 pp.

9 Mertens, D. M. (2010). 323 pp.

CHAPTER IV

Student Learning Objectives: Quality and Attainment

The Student Learning Objective process refers to the instructional intervention selected to improve student achievement in the TIF-LEAP initiative. The emblematic artifact of the SLO process is the objective developed by the teacher. It is a measure of a teacher's performance and its attainment is one basis of the bonus payout provided by the initiative. Nearly 4,000 SLOs were submitted by teachers in the participating TIF-LEAP schools during the years included in this evaluation: 707 in 2008-09, 1,904 in 2009-10 and 1,353 in 2010-11.[1]

SLOs were introduced into the TIF-LEAP schools to meet several goals of the initiative. These include building the capacity of teachers and principals to plan effectively for student growth, compensating teachers for student academic gains, and retaining qualified teachers and principals in hard-to-staff schools. For this reason, analyzing the quality of SLOs, the attainment of SLO growth targets, and the relationship of quality to attainment is important in evaluating the outcomes of the TIF-LEAP initiative and its impact on measuring and improving teacher performance.

The examination of the impact of the SLO process on teachers and students begins with rating each SLO on a four-level, four-criteria rubric. After rating the SLOs, the analysis continues with an examination of (1) the distribution of all SLO ratings for school years 2008-09, 2009-10, and 2010-11; (2) the attainment of those SLOs by teachers (referred to as "met" or "not met"); (3) the relationship between the SLO ratings and their attainment; (4) the relationship between teacher experience (number of years) in the initiative and the quality of their SLOs; (5) the linkages between earning a value-added measure bonus and the quality of SLOs; and (6) the relationship of teacher turnover in the TIF-LEAP schools to the comparison schools. The following is a summary of the major findings.[2]

SLO Quality and Attainment: Summary of Major Findings

- The quality of SLOs increases from the first year (2008-09) of implementation to the second year (2009-10). This occurs even as 10 additional schools join the initiative.
- The proportion of SLOs attained (met) increases from 2008-09 to 2009-10. This proportion decreases from 2009-10 to 2010-11 at a time of significant changes in district conditions.
- The overall relationship between the quality of SLOs and their attainment is positive. Year-by-year findings vary with the highest correlation found in 2009-10.
- The number of years a teacher participates in SLO implementation matters. Teachers in TIF-LEAP for three years of SLO implementation develop higher quality SLOs and have greater success in attaining their SLOs.
- Teachers in TIF-LEAP schools who receive a VAM bonus are more likely to have high quality SLOs. This finding applies to 2009-10 and 2010-11. It is statistically significant in 2009-10.

Scoring Methodology for Student Learning Objectives

The CTAC evaluation team collected, read, and rated all SLOs submitted by TIF-LEAP teachers. These ratings take place after the SLO is approved by the principal, and before knowing if the SLO has been attained and whether or not the teacher met the eligibility criteria for receiving a bonus. Once the SLOs are rated, the scores are compiled by school, teacher,[3] rubric level/rating, and attainment status (met or not met) and used to generate further analyses.

The rating rubric is based on (1) four criteria associated with effective SLOs—learning content, completeness, cohesion, and expectations—and (2) four levels of quality, with Level 4 being the highest. The four levels are: 4 Excellent; 3 Acceptable; 2 Needs Improvement; or 1 Too Little to Evaluate. The four criteria described together with the four levels of quality are the building blocks of the rating rubric used by the evaluators. The attainment of an SLO—whether the teacher met his or her growth target—is confirmed by the school principal and the associated data are compiled and reviewed by the TIF-LEAP team.

Types of Student Learning Objectives: Class and Target

Two types of Student Learning Objectives were implemented in the TIF-LEAP participant schools—Class and Target. The primary difference between a Class and a Target SLO centers on the selection of students the teacher plans to address in the SLO, called the population, and the teacher's baseline analysis of student needs. For example, a Class SLO is developed for the teacher's entire class; a Target SLO is developed for a *group of students in the class* targeted by the teacher based on evidence of specific learning needs.

SLOs can be developed by an individual teacher, or by a group of teachers such as grade level, subject, or interdisciplinary cluster, to address similar learning content with the same assessment. For the SLOs developed by a group of teachers, the growth expectations are still set by each individual teacher for his or her own students based on baseline data and pre-assessments. Further, SLOs may also be developed by a classroom teacher in collaboration with a non-classroom teacher, such as a resource specialist. These SLOs are still either Class or Target, based on the population addressed. They provide a vehicle for a non-classroom teacher to collaborate in developing and implementing SLOs.

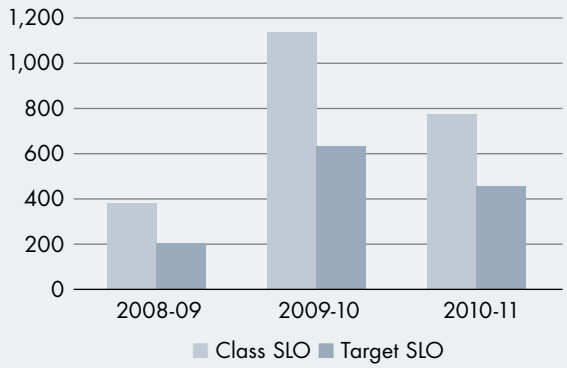
During years two, three, and four (2008-2011) of the initiative—the first three years of SLO implementation—Class SLOs (n=2,457) comprised 61.9% of all approved SLOs and Target SLOs (n=1,507) comprised 38.1%. *Figure IV.1* shows the distribution by type of SLO over the three-year period under study.[4]

Quality of SLOs: Excellent, Acceptable, Needs Improvement, or Too Little to Evaluate

The quality of SLOs changed over the life of the initiative, as shown in the rating distributions displayed in *Table IV.1*. In the first year of SLO implementation (2008-09), 707 SLOs are

FIGURE IV.1

Distribution of SLOs Submitted by Type, 2008-09, 2009-10, 2010-11



completed by ten schools. Nearly one-third of these SLOs (30.7%) are rated at Level 4 Excellent and 53.7% are rated at Level 3 Acceptable. The remaining 110 (15.6%) are rated Level 2 Needs Improvement. There are no SLOs identified as Level 1 Too Little to Evaluate in that year.

With the addition of ten more schools in 2009-10 (two elementary and eight high schools), the number of SLOs submitted more than doubled to 1,904. The percent of Level 4 Excellent SLOs increase to 50.4%, and the percent of Needs Improvement and Too Little to Evaluate SLOs drop to 6%, showing an increase of the quality of SLOs from 2008-09 to 2009-10.

Although the twenty schools are still in the initiative in 2010-11, the total number of SLOs submitted decrease to 1,353. There is a redistribution of the quality of SLOs. There are higher percentages of Levels 4 and 3 than in the first year of SLO implementation (2008-09), and there is shifting in the percentages of Levels 4 and 3 from the second year of SLO implementation (2009-10).

Chi-square test results confirm that the annual changes in the distribution of the quality of SLOs discussed above are statistically significant (chi-square = 51.67; $df = 2$; $p < .001$).

Attainment of SLOs: Student Growth Target Met or Not Met

As indicated met or not met refers to the question: *Do students meet the growth target in the SLO, based on the pre- and post-assessment analysis and using the measure identified in the SLO?*

During the three years under study, 3,964 SLOs were submitted and rated. A total of 2,711 SLOs (68.4%) met the stated growth targets and eligibility requirements for a bonus payout. *Table IV.2* shows the distribution of the attainment of growth targets. The highest attainment rate is in year 2009-10 (80.6%), increasing from a rate of 61.2% in 2008-09, but decreasing to 54.9% in year 2010-11.

Chi-square test results confirm that the change in the attainment of SLOs is statistically significant across the three years (chi-square = 262.07; $df = 2$; $p < .001$).

Relationship of the Quality of SLOs to their Attainment

Teachers with SLOs earning rubric ratings at Level 4 and 3 met their growth targets more frequently than those at Levels 2 and 1. The

TABLE IV.1

Level of Quality of Student Learning Objectives

Year	Level of Quality				Total SLOs
	4 Excellent	3 Acceptable	2 Needs Improvement	1 Too Little to Evaluate	
2008-09	217 (30.7%)	380 (53.7%)	110 (15.6%)	0 (0.0%)	707
2009-10	959 (50.4%)	830 (43.6%)	111 (5.8%)	4 (0.2%)	1,904
2010-11	532 (39.3%)	754 (55.7%)	62 (4.6%)	5 (0.4%)	1,353
Total SLOs	**1,708 (43.1%)**	**1,964 (49.5%)**	**283 (7.1%)**	**9 (0.2%)**	**3,964**

TABLE IV.2

Attainment of Growth Target

Year	Attainment of Growth Target		Total SLOs
	Met	Not Met	
2008-09	433 (61.2%)	274 (38.8%)	707
2009-10	1,535 (80.6%)	369 (19.4%)	1,904
2010-11	743 (54.9%)	610 (45.1%)	1,353
Total SLOs	**2,711 (68.4%)**	**1,253 (31.6%)**	**3,964**

results of a point bi-serial correlation confirm that this relationship is statistically significant at $p < .001$. *Table IV.3* shows these results.

The findings vary when the relationship is tested year by year. Please note: (1) positive relationships between SLO quality and attainment are found for 2009-10 and 2010-11, but not in 2008-09 which was the first year of implementation; and (2) 2009-10 has the highest SLO attainment across all SLO quality levels.

SLOs and Teacher's Years in TIF-LEAP

The study examined the relationship between the quality and attainment of a teacher's SLOs and a teacher's years of experience in the initiative. The distributions of the quality of SLOs and the attainment of SLOs relative to teachers' years of experience are shown in *Table IV.4*.

In terms of the quality of SLOs, out of all 733 teachers who submitted SLOs in 2010-11, 147 teachers were in their first year of participation, 351 teachers in their second year, and 235 teachers in their third year (the upper section of *Table IV.4*). A comparison of distributions in the upper section indicates that the longer the teachers participate in the SLO implementation (e.g., the number of years participating in the initiative), the higher the quality of their SLOs. Teachers with three years of experience have the highest percentage of Level 4 SLOs (47.5%), compared to those with one year and two years (32.9% and 36.2%, respectively). Chi-square tests reject the null hypothesis that the distributions of SLO quality are the same across teachers with different years of experience.

The middle section of *Table IV.4* shows the distributions of the quality of SLOs in 2009-10. It confirms that the positive association between the quality of SLOs and years of experience is higher for those teachers who were in the initiative for two years with 57.7% Level 4 SLOs. In 2008-09, the first year of SLO implementation, all teachers have one year of experience, thus providing a reference point or baseline for comparing the quality of SLOs across subsequent years.

TABLE IV.3

Comparison of the Quality of SLOs and their Attainment, 2008-11

Attainment of Growth Target	Level of Quality			
	4 Excellent	3 Acceptable	2 Needs Improvement	1 Too Little to Evaluate
Met	1,223 (30.9%)	1,310 (33.0%)	176 (4.4%)	2 (0.1%)
Not Met	485 (12.2%)	654 (16.5%)	107 (2.7%)	7 (0.2%)
Attainment Ratio	**2.52:1.00**	**2.00:1.00**	**1.64:1.00**	**0.29:1.00**

TABLE IV.4

Quality and Attainment of SLOs by the Teacher's Years of Experience in TIF-LEAP

Years of Experience	Number of Teachers	Level of Quality				Attainment of Growth Target	
		4 Excellent	3 Acceptable	2 Needs Improvement	1 Too Little to Evaluate	Met	Not Met
SLOs Submitted in 2010-11							
1 Year	147	83 (32.9%)	150 (59.5%)	17 (6.7%)	2 (0.8%)	123 (48.8%)	129 (51.2%)
2 Years	351	238 (36.2%)	385 (58.6%)	31 (4.7%)	3 (0.5%)	337 (51.3%)	320 (48.7%)
3 Years	235	211 (47.5%)	219 (49.3%)	14 (3.2%)	0 (0.0%)	283 (63.7%)	161 (36.3%)
SLOs Submitted in 2009-10							
1 Year	547	530 (45.7%)	545 (47.0%)	81 (7.0%)	4 (0.3%)	931 (80.3%)	229 (19.7%)
2 Years	328	429 (57.7%)	285 (38.3%)	30 (4.0%)	0 (0.0%)	604 (81.2%)	140 (18.8%)
SLOs Submitted in 2008-09							
1 Year	433	217 (30.7%)	380 (53.7%)	110(15.6%)	0 (0.0%)	433 (61.2%)	274 (38.8%)

The last two columns of *Table IV.4* show the distribution of SLO attainment for the same teacher groups in terms of years of experience. A comparison of the distributions within the same year indicates that the attainment rate increased as the teachers' years of experience in TIF-LEAP increased. Chi-square tests show the increase is statistically significant.

In sum, as teachers gained more experience in crafting SLOs, the quality of their SLOs improved and the higher their rate of attainment in both 2009-10 and 2010-11.

Linkages Between the VAM Bonus and Quality of SLOs

District leadership introduced the value-added measure for teachers and administrators in the TIF-LEAP schools in 2009-10. SLOs had been implemented in ten schools in the previous year, and ten additional schools joined the initiative in 2009-10. The VAM was continued in 2010-11 and 2011-12. The introduction of the VAM is detailed in Chapter II.

Like SLOs, the VAM is a measure of teacher performance and student achievement. However, they measure different aspects of performance and achievement: (1) the VAM measures performance of teachers only in tested grades and subject areas; and (2) SLOs measure the performance of teachers in all grades and subjects, while also engaging teachers in best instructional practices. The evaluation examined the relationship between these two approaches to performance-based compensation in the TIF-LEAP initiative.

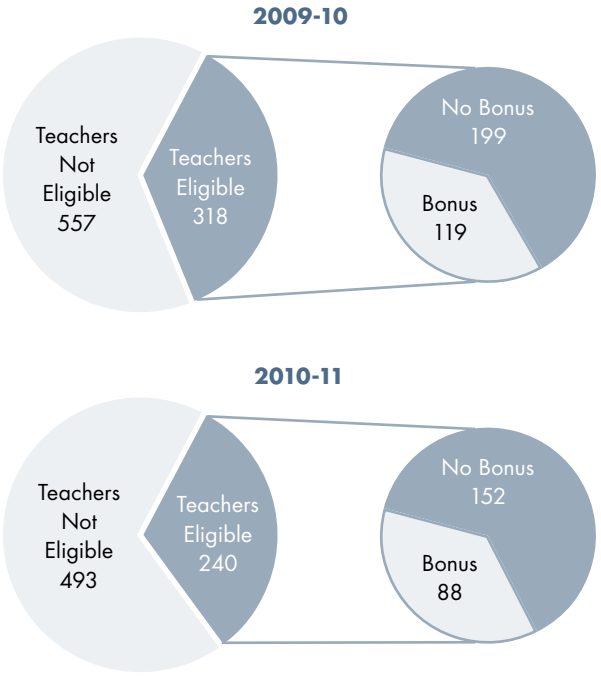
TIF-LEAP teachers in tested grades and subject areas were eligible to receive an individual VAM bonus; that is, only teachers whose students participated in the EOG/EOC assessments. In order to receive the VAM bonus, those eligible teachers had to have (1) a VAM score at or above the 70th percentile (be in the top 30% of teachers in the district); and (2) a rating of proficient or above on their performance evaluation.

Figure IV.2 shows the distribution of teachers' eligibility and achievement of VAM bonuses. Specifically, in 2009-10, 318 of the 875 teachers participating in TIF-LEAP taught in a state-tested grade or subject area. Of these eligible teachers, 34.7% (119 out of 318) received a VAM bonus. In 2010-11, 240 of the 733 teachers participating in TIF-LEAP taught in a state-tested grade or subject area. Of these eligible teachers, 36.6% (88 out of 240) received a VAM bonus.

The distributions of the quality of SLOs for VAM recipients and non-recipients in the TIF-LEAP schools are compared in *Table IV.5*. Teachers may create more than one SLO in a year, yet they receive only one VAM bonus. To account for this, the mean rating of each teacher's SLOs is used in this analysis. For school year 2009-10, the 119 teachers who receive a VAM bonus are more likely to have high quality SLOs than teachers who do not receive a VAM bonus. This result is statistically significant with chi-square = 10.72 and $p < .05$. A similar pattern is seen in 2010-11; however, the chi-square is not statistically significant in that year.

FIGURE IV.2

Distribution of Eligibility for and Achievement of VAM Bonuses, 2009-10 and 2010-11



Teacher Turnover

A thorough examination of teacher retention in TIF-LEAP and comparison schools is constrained by data limitations. Specifically, all classroom teachers employed in a school during March of the previous year but not employed as a classroom teacher *in the same school system/district* during March of the current year are included in the school's turnover rate. These data are reported each year as a percentage on the school's Report Card through the North Carolina "Education First NC School Report Cards" program and are compared to the state and district turnover rates.[5]

These data on outward mobility do not indicate if a teacher moves within the district or leaves the district, nor do they address such factors as: teacher layoffs or reduction in force; structural reorganization (i.e., changing a K-5 school to a K-8 school; dividing a high school into several small, theme-based schools or combining small schools into a comprehensive high school); or voluntary or involuntary re-assignment to another school.

In comparing the average turnover rates for TIF-LEAP schools to the average turnover rates for the comparison schools, the number of schools entering the initiative in different years (as either

TABLE IV.5

VAM Score and Mean SLO Ratings, 2009-10 and 2010-11

VAM Score	Mean SLO Rating, 2009-10				Mean SLO Rating, 2010-11			
	3.5-4.0	2.5-3.49	1.0-2.49	Total	3.5-4.0	2.5-3.49	1.0-2.49	Total
<70th percentile	89	83	8	180	56	73	6	135
>70th percentile	75	35	9	119	33	43	1	77
Total	**164**	**118**	**17**	**299**	**89**	**116**	**7**	**212**
	Chi-square = 8.70; *df* = 2; $p < .05$				Chi-square = 1.52; *df* = 2; $p = .47$			

TIF-LEAP or control schools) and remaining in the initiative for different numbers of years was taken into account. No clear trends emerged.

Research literature suggests several possible ways teacher turnover can impact student achievement and other school-related factors. For example, research indicates that high levels of turnover have a negative impact on student achievement when good teachers leave and their replacements may be less experienced; however, in some settings less effective teachers are more likely to leave. At the same time, there can be "disruptive" factors regardless of the effectiveness of the teachers, such as a loss of institutional knowledge or a decrease in staff collegiality and cohesion.[6]

Summary

The use of SLOs was the approach selected as a means to both improve and measure teacher performance. For that reason, the evaluation gives considerable attention to the quality of SLOs and how frequently the growth targets are met.

Individual year analyses show that the quality and attainment of SLOs are influenced by a number of factors, including teacher experience and changes in district conditions. The major findings are: SLO quality and attainment increased from 2008-09 to 2009-10; there is a positive relationship between the quality of SLOs and their attainment; SLO quality and attainment increase with a teacher's years of participation in the initiative; and teachers who receive a VAM bonus are more likely to have high quality SLOs.

Endnotes

1 SLOs were implemented beginning in 2008-09, the second year of the initiative. School year 2011-12 is the concluding year of the TIF-LEAP initiative. As indicated in Chapter III, gaps in the student achievement database for 2011-12 resulted in a highly unstable and substantially unrepresentative sample. These gaps precluded the ability to satisfactorily estimate effects of SLOs on student achievement in the initiative's final year.

2 The student achievement analyses are presented in Chapter V. The statistical methods used for all analyses are described in Chapter III.

3 Individual teacher rubric ratings are maintained confidentially and only used for evaluation of the CMS project. Individual ratings are not reported to the teacher or the district.

4 There were ten schools in the initiative in 2008-09 and twenty schools in years 2009-11.

5 School level turnover rates are derived from school report card data. See http://www.ncschoolreportcard.org/src/.

6 Ronfeldt, M., Lankford, H., Loeb, S., & Wyckoff, J. (2011). *How teacher turnover harms student achievement* (No. w17176). National Bureau of Economic Research.

CHAPTER V

Impact of TIF-LEAP on Student Achievement

The analysis of the relationship between the quality of Student Learning Objectives and the attainment of those objectives, as explained in Chapter IV, indicates that the higher the quality of the SLO, the greater the likelihood that the teacher meets or attains the student growth target. The relationship is statistically significant in all three years for all types of SLOs developed in TIF-LEAP schools. This finding merits further examination of the effect of the TIF-LEAP initiative, in general, and SLOs in particular, on student performance.

The analysis in this chapter examines the impact of TIF-LEAP and SLOs on student achievement using independent measures—the annual North Carolina End-of-Grade (EOG) mathematics and reading tests. Specifically, this analysis (1) compares the growth trajectory between TIF-LEAP and comparison schools; (2) evaluates the impact of TIF-LEAP on student achievement; and (3) evaluates the relationship between SLOs and EOG achievement in the TIF-LEAP elementary and middle schools.[1]

Different analytical techniques are employed to appraise the impact of the TIF-LEAP initiative on student achievement. First, descriptive statistics examine the growth trajectory between TIF-LEAP and comparison schools. Secondly, longitudinal hierarchical linear models (HLM) estimate the impact of the TIF-LEAP initiative on student achievement in relation to the achievement of comparison schools. Finally, cross-sectional HLMs analyze the relationships

between the SLOs and student achievement in the TIF-LEAP elementary and middle schools.

The following student achievement analysis includes: (1) an initial presentation of all of the major findings from these analyses; (2) a description of student demographic and achievement data and teacher and principal characteristics used in the analyses; (3) a presentation of descriptive plots of the relative performance between TIF-LEAP and comparison schools over time; and (4) a description and discussion of the longitudinal and cross-sectional hierarchical linear models and associated results.

TIF-LEAP and Student Achievement: Major Findings

Descriptive statistical results show that the growth rate of students in TIF-LEAP schools is greater than the growth rate of students in the comparison schools during the years under study.[2] Although the TIF-LEAP schools started with lower student performance, by the end of year four, the student test scores in both mathematics and reading are closely approaching those of the comparison schools.

The TIF-LEAP schools also show greater resilience to the negative shocks resulting from the economic recession, including teacher layoffs, and planning for school closures and restructuring[3] that occurred in 2010-11. Student test scores in the TIF-LEAP schools grew at a lower rate in that year than in the previous school year. However, they grew at a higher rate than the comparison schools that experienced the same disruptions.

The longitudinal HLM models provide the estimated effects of the initiative. They show that TIF-LEAP had a positive impact on the participating schools which is both statistically and practically significant. Specifically,

- In terms of mathematics achievement, students in TIF-LEAP schools on average have a growth rate 12% greater than students in the comparison schools.

  This growth difference is substantial and means that the TIF-LEAP students are growing 12% more than the 0.8% annual growth rate of the comparison school students. This growth translates into 0.34 points annual growth difference between TIF-LEAP and comparison students. As a result, at the end of year four of the initiative, the test scores of students in the TIF-LEAP schools improved, cumulatively, 1.4 points more than students in the comparison schools. This growth brings the TIF-LEAP schools close to par with the comparison schools (students in TIF-LEAP schools started 1.5 points lower than students in the comparison schools at the beginning of the initiative).

- In terms of reading achievement, students in TIF-LEAP schools on average have a growth rate 13% greater than students in the comparison schools.

  This growth difference is substantial and translates into 0.44 points annual growth difference between TIF-LEAP and comparison students. As a result, at the end of year four of the initiative, the test scores of students in the TIF-LEAP schools are only 0.7 points lower than those in the comparison schools. The initial test scores of the TIF-LEAP students started 2.5 points lower than students in the comparison schools.

Three cross-sectional HLM analyses were conducted over the course of the TIF-LEAP initiative. The findings of the cross-sectional HLM models vary by subject and year.

The first cross-sectional analysis is for 2008-09, the first year of SLO implementation. The full SLO effects on student achievement were expected to phase in over several years of implementation. The findings in the first year support this expectation:

- There are positive, statistically significant associations between the attainment of Target SLOs and student achievement both in mathematics and in reading.[4]

- There is no statistically significant association between the quality of SLOs (as indicated by the rubric rating)[5] and student achievement.

The second cross-sectional analysis is for 2009-10. In terms of achieving higher student performance, this is the peak year of SLO implementation. The key findings include the following:

- There are positive, statistically significant associations between the quality of SLOs and student achievement. This finding means that a teacher's SLO rating relates positively to student achievement in elementary school mathematics, elementary school reading, and middle school mathematics.
- There are positive, statistically significant associations between the attainment of SLOs and student achievement at the elementary school level. This finding means that the students whose teachers met their SLOs achieved higher scores in elementary school mathematics and reading.

The third cross-sectional analysis is for 2010-11. In this school year, as a result of the increase in the number of students and classrooms, the investigation is conducted at the individual grade level in grades 4-8 rather than combining grades into elementary and middle school analyses. The key findings include:

- There is a positive, statistically significant association between the quality of SLOs and student achievement in mathematics in grade five.
- There is a positive, statistically significant association between the attainment of SLOs and student achievement in reading in grade six.

The following discussion provides more technical detail on the databases, analytical methodologies, and findings. It may be useful to refer to Chapter III where the initiative design and data availability information are presented because information related to the design and timing of the initiative, including the selection of comparison schools, helps explain the analytical methodologies employed.

Student and Teacher Data

Whereas the overall data used in this report are described in detail in Chapter III, particularly relevant data are presented here. This analysis draws on student achievement data for school years 2006-07 to 2010-11. It is a rich data set, which includes information on student EOG scores, student demographics and other characteristics, as well as information on teachers and principals in 20 TIF-LEAP schools and 24 comparison schools. The data set is hierarchically structured at three levels: student, classroom and school.

Student Characteristics. *Table V.1* shows the student demographic data of TIF-LEAP and comparison schools over the four years of the evaluation. Student enrollment remained relatively stable over time, both in TIF-LEAP and comparison schools. In terms of race/ethnicity, both TIF-LEAP and comparison schools had a majority of African American students with Hispanic/Latino students being the second largest student group. The comparison schools had a larger proportion of White students than the TIF-LEAP schools.

The distribution of students by gender (51% male, 49% female) was consistent across both TIF-LEAP and comparison schools. However, the composition of schools by other demographic characteristics shows differences. Specifically, TIF-LEAP schools had a higher percentage of students receiving free or reduced price lunch (FRPL) and students with limited English proficiency (LEP) than the comparison schools.

Two additional student characteristics included in these analyses are incidents of disciplinary behaviors, including in-school and out-of-school suspensions (ISS, OSS), and absenteeism, including unexcused absences and tardiness, and excused absences and tardiness. Based on 2008-09 data, the average number of incidents of disciplinary behaviors was 14.0 and 8.2 for TIF-LEAP and comparison schools, respectively; the average number of incidents of excused absences and tardiness was 1.7 and 2.7, respectively.

Teacher Characteristics. *Table V.2* shows the teacher demographics for the TIF-LEAP and comparison schools. The average experience level and the proportion of fully licensed or advanced degree teachers in TIF-LEAP schools increased more than in the comparison schools.[6] These data support the TIF-LEAP goal of recruiting and retaining highly qualified teachers to high need schools.

Principal Characteristics. The characteristics in 2007-08 indicate that TIF-LEAP principals average more years of experience in the principalship overall, 11.1 years, than the principals in the comparison schools who average 7.8 years as a

TABLE V.1

Student Demographics, TIF-LEAP and Comparison Schools, 2007-11

	2007-08		2008-09		2009-10		2010-11	
	TIF-LEAP	Comp.	TIF-LEAP	Comp.	TIF-LEAP	Comp.	TIF-LEAP	Comp.
Enrollment	13,426	21,346	13,786	20,551	13,338	19,978	13,404	19,995
Race/Ethnicity								
African American	69.1%	58.5%	68.3%	59.0%	68.5%	58.2%	67.2%	56.9%
White	6.9%	14.8%	6.6%	13.4%	5.7%	12.8%	6.6%	14.6%
Asian	3.9%	4.3%	4.7%	4.4%	5.1%	4.5%	5.0%	4.4%
Hispanic/Latino	17.5%	18.7%	18.1%	19.3%	18.5%	20.3%	18.8%	20.8%
Other	2.6%	3.7%	2.3%	3.9%	2.2%	4.2%	2.4%	3.3%
Gender								
Male	49.9%	51.3%	51.4%	50.6%	51.6%	51.0%	51.3%	51.1%
Female	50.1%	48.7%	48.6%	49.4%	48.4%	49.0%	48.7%	48.9%
Other Demographic Characteristics								
FRPL	78.0%	64.7%	79.5%	66.9%	83.3%	70.3%	84.2%	72.9%
LEP/ELL	18.8%	19.0%	19.3%	20.3%	19.0%	19.9%	18.3%	20.3%
Students with Disabilities	13.2%	12.0%	11.7%	10.9%	13.7%	12.0%	13.9%	11.4%
Academically Gifted	2.4%	5.1%	3.0%	4.8%	2.6%	4.7%	2.4%	4.6%

TABLE V.2

Teacher Demographics, TIF-LEAP and Comparison Schools, 2007-11

	2007-08		2008-09		2009-10		2010-11	
	TIF-LEAP	Comp.	TIF-LEAP	Comp.	TIF-LEAP	Comp.	TIF-LEAP	Comp.
Classroom Teachers	997	1,409	1,003	1,401	968	1,281	880	1,127
Years of Experience								
0 to 3 Years	37.9%	29.5%	35.2%	28.8%	29.0%	25.1%	28.2%	21.4%
4 to 10 Years	31.2%	33.6%	33.5%	33.7%	35.3%	37.0%	34.2%	37.8%
11+ Years	30.9%	36.9%	31.3%	37.5%	35.6%	37.9%	38.2%	41.0%
Other Demographic Characteristics								
Fully Licensed	78.9%	87.4%	80.7%	88.1%	85.6%	89.5%	84.5%	89.7%
Advanced Degrees	26.2%	26.8%	27.5%	27.4%	30.5%	28.6%	32.6%	32.6%
NBPTS Certified	7.5%	7.8%	7.1%	9.1%	9.4%	12.3%	10.2%	14.9%

principal. Similarly, the TIF-LEAP principals are in their current school an average of 3.1 years, slightly longer than the comparison school principals who average 2.6 years in their current position.

All of the characteristics described above are included in the baseline models. However, only those that are statistically significant and/or contribute to the model fitness are included in the final models to maximize the efficiency of the model estimation.

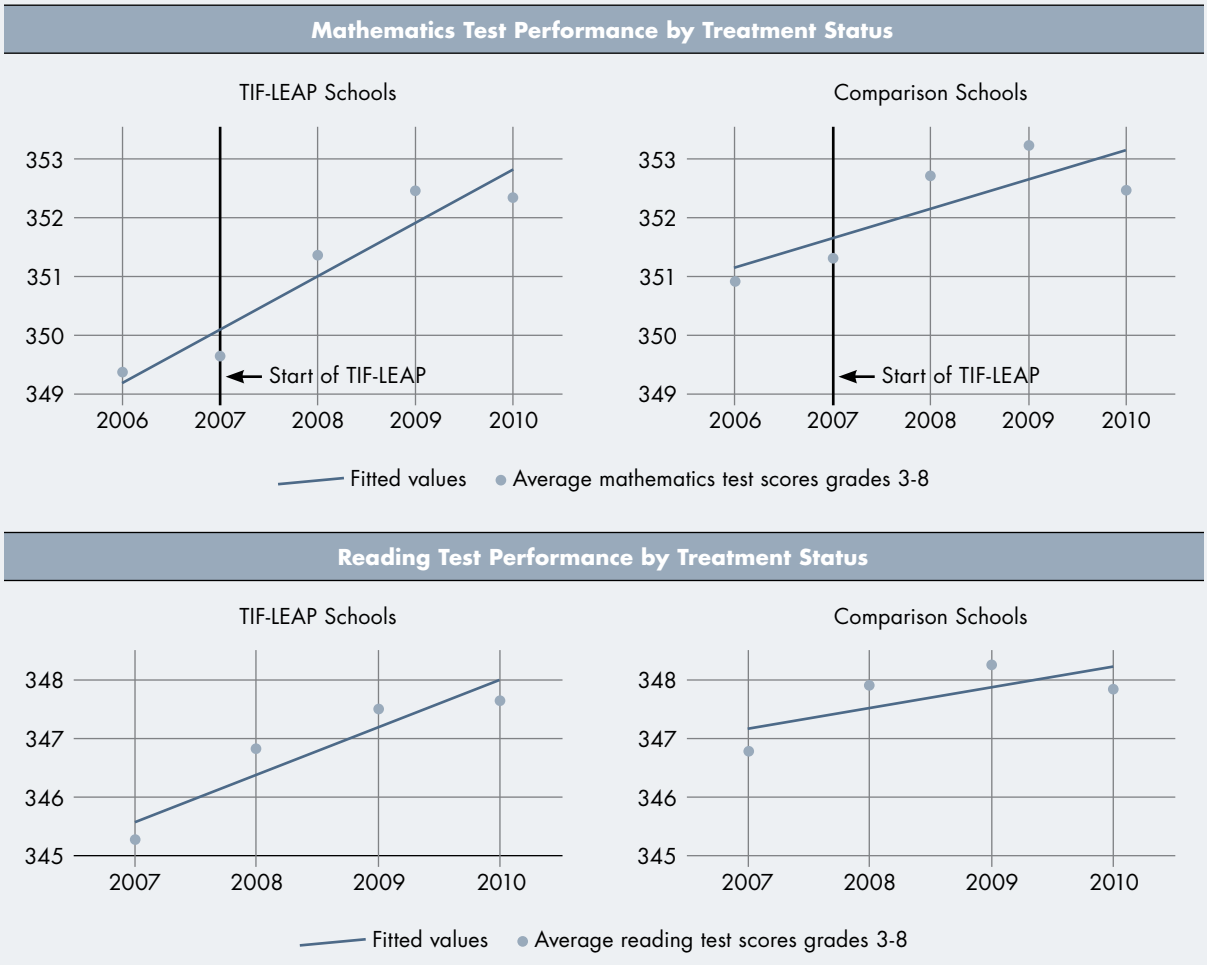
Descriptive Statistics

The simple descriptive statistics provide an initial view of the schools over the initiative. To examine if students in TIF-LEAP schools have a growth rate greater or less than those in the comparison schools, the analysis computed the average mathematics and reading scores for all students in the TIF-LEAP and comparison schools over the school years 2007-08 to 2010-11. Plots of these average test scores are shown in *Figure V.1*. Two main points from the plots of the average test scores by subject and by initiative status suggest that:

- Although the TIF-LEAP schools started with lower student performance, by the end of year four of the initiative, student test scores in both mathematics and reading are approaching those of the comparison schools. *Figure V.1* illustrates that students in the TIF-LEAP schools, during the years under study, are growing at a higher rate on the North Carolina EOG tests than those in the comparison schools.

FIGURE V.1

Mathematics and Reading Test Performance in TIF-LEAP Schools and Comparison Schools



- In 2010-11, the student test scores in TIF-LEAP schools are not growing at the rate expected from the linear projection of the previous years; however, their growth rate is higher than that of the comparison schools. It appears that, at the TIF-LEAP schools, some of the shocks from the district's changed economic circumstances are offset by the TIF-LEAP effects.

Figure V.1 on mathematics begins with 2006-07—a year before the launch of the TIF-LEAP initiative—to show that the TIF-LEAP schools were lagging behind the comparison schools before the implementation. However, due to changes in the reading EOG assessment, the 2006-07 reading data are not comparable to those in 2007-08 and subsequent years of the initiative.[7]

Because of the limitations of the simple descriptive statistics,[8] hierarchical linear models (HLM) are specified to estimate the various TIF-LEAP effects. HLM is used because of the nested nature of these data.[9] HLM estimates relationships at the individual and group (class, school, or TIF-LEAP status) levels and allows for relationships at the student level to vary across group levels, as well as for modeling cross-level interactions. A full discussion on the nested data is included later in this chapter because it is easier to illustrate in a specific context.

Two types of HLM models are specified. To evaluate the impact of TIF-LEAP on student achievement, a two-level longitudinal HLM is specified. To analyze the relationship between SLOs and student achievement in TIF-LEAP schools, a three-level cross-sectional HLM is specified. The model specifications are discussed in turn below.

Longitudinal Hierarchical Linear Models

Longitudinal HLM models are used to estimate TIF-LEAP treatment effects by exploring the differences in longitudinal achievement records across different students. In contrast to a cross-sectional analysis, which has between-students variation and therefore allows only for point-in-time estimates, longitudinal analysis explores a richer set of information of panel data, i.e., identifies both the temporal patterns (also called time-dimension or within-student variation) and cross-sectional patterns (also called spatial-dimension or between-students variation) in the data. In addition, longitudinal analysis allows for a higher level of "controls" so that the time invariant unobservable or unmeasurable factors (e.g., a student's innate ability), that cannot be controlled in a cross-sectional analysis, are contained in the error terms.

The time variant unobservable and unmeasurable factors (e.g., a student's FRPL status) cannot be controlled in the longitudinal HLM model. This can lead to selection bias in the estimation and constitute a main caveat in the interpretation of results.[10] For example, if the TIF-LEAP and comparison schools had different trend lines in the pre-initiative period, the estimate on the treatment effect is biased because it captures the "true" treatment effect plus a bias which is due to the difference in trend. Multiple years of pre-initiative data are not available for a direct test of the trends. However, artifacts point to similar trends showing stability and consistency across multiple years.

Because of the use of student level data to examine the initiative's effects, two-level longitudinal HLM models are specified. The first level estimates student achievement over time and the second level estimates the relationship between the achievement over time and initiative status. The cross-level interactions are captured by the covariance of first level and second level error terms. The analyses focus on mathematics and reading for students in grades 4-8. For each of these analyses, an unconditional mean model (UMM) is fitted first to partition the total variance in the student test scores. The UMM does not include any predictors. An analysis of the estimates of the UMM helps in guiding the approach to the final model specifications.

A summary of the main findings of the longitudinal HLM models is included first and then a detailed discussion of the models and associated results follows.

Summary of Findings of the Longitudinal Hierarchical Linear Models

In general, there are statistically significant TIF-LEAP treatment effects on students' achievement in both mathematics and reading. Specifically,

- In terms of mathematics achievement, TIF-LEAP students, on average, have a growth rate 12% greater than students at the comparison schools. The true initial test scores of the TIF-LEAP students are 1.5 points lower than students at the comparison schools. However, the annual rate of change in test scores for the average TIF-LEAP student is 0.34 points higher than his/her counterpart in comparison schools. This estimate is statistically significant at $p < .01$.
- In terms of reading achievement, TIF-LEAP students, on average, have a growth rate 13% greater than students at the comparison schools. The true initial test scores of the TIF-LEAP students are 2.5 points lower than students at the comparison schools. The annual rate of change in test scores for the average TIF-LEAP student is 0.44 points higher than his/her counterpart in comparison schools. This estimate is statistically significant at $p < .01$.

Description of Longitudinal Hierarchical Linear Models

The longitudinal HLM models, starting with the UMM model, are discussed below. With no predictor, the UMM model is as follows:

Level-1 Equation:

$$Y_{it} = \pi_{0i} + e_{it} \quad (1)$$

- Y_{it} is the test score of student i at time t;
- π_{0i} is the student i's mean test score;
- e_{it} is the residual with $E(e_{it}) = 0$ and $var(e_{it}) = \sigma_e^2$. It measures the deviation of i's test scores from π_{0i}, i.e., it measures within-student variation.

Level-2 Equation:

$$\pi_{0i} = \gamma_{00} + u_{0i} \quad (2)$$

- γ_{00} is the grand mean of the test scores of all students;
- u_{0i} is the Level-2 residual with $E(u_{0i}) = 0$ and $var(u_{0i}) = \sigma_{u0}^2$. It permits the Level-1 parameters to vary stochastically across students, i.e., measures the between-students variation.

The composite form of the longitudinal UMM:

$$Y_{it} = \gamma_{00} + u_{0i} + e_{it} \quad (3)$$

Equation (3) comes from combining equations (1) and (2).

The UMM is used to partition the proportion of variance within—and between—students, which is used as a baseline for evaluating the success of subsequent longitudinal models. The proportions of variance distributed across the two levels are described by the following formula:

Proportion of variance at Level-1:
$$\sigma_e^2/(\sigma_e^2 + \sigma_{u0}^2);$$
Proportion of variance at Level-2:
$$\sigma_{u0}^2/(\sigma_e^2 + \sigma_{u0}^2).$$

The fourth column of *Table V.3* shows the estimation results of the longitudinal UMM on mathematics EOG tests. The grand mean of the mathematics test scores in our total sample of 40,067 student observations is 353.87; 46.5% of the variation is from within-student variation (calculated using $\sigma_e^2/(\sigma_e^2 + \sigma_{u0}^2)$) and 53.5% is from between-students variation, so both within and between variations are significant and need to be explored.[11]

Next, the variations in student test scores are explored by looking at what role TIME plays, where TIME is defined according to school years. This is accomplished by specifying a second longitudinal HLM model, adding TIME into the UMM. The second model is called unconditional growth model (UGM), a model with TIME the only Level-1 predictor and no substantive predictors at Level-2, which helps evaluate the baseline amount of change. The composite form of the UGM model is as follows:

$$Y_{it} = \gamma_{00} + \gamma_{10} TIME_{it} + [u_{0i} + u_{1i} TIME_{it} + e_{it}] \quad (4)$$

- Y_{it} is the test score of student i at time t;
- $TIME_{it}$ is a generic variable with school year 2007-08 normalized to 0. It embodies our hypothesis about the growth of student test scores over time;

- γ_{00} is the population average of the Level-1 intercept π_{0i}, i.e., the population average true initial status of all students;
- γ_{10} is the population average of the Level-1 slope π_{1i} for students with a Level-2 predictor value of 0, i.e., the population average annual rate of true change for students;
- u_{0i} and u_{1i} are the Level-2 residuals with $E(u_{0i}) = E(u_{1i}) = 0$, $var(u_{0i}) = \sigma^2_{u0}$ and $var(u_{1i}) = \sigma^2_{u1}$; $cov(u_{0i}, u_{1i}) = \sigma_{01}$. They permit the Level-1 parameters to vary stochastically across students;
- e_{it} is the residual with $E(e_{it}) = 0$ and $var(e_{it}) = \sigma^2_e$. It now represents the deviations of i's test scores at time t from true change trajectory.

Equation (4) comes from combining the Level-1 and Level-2 equations of the UGM model.[12] The second column of *Table V.3* shows the estimation results of the UGM. The average true initial status of the student mathematics test scores in 2007-08 is 349.79. Student mathematics scores increase, on average, by 2.88 points annually for all schools. Comparing the σ^2_e between the UMM and UGM model, one can see that 18% of the within-student variation is associated with the linear TIME. The goodness-of-fit statistics all point to a better fit of the UGM than UMM.

The UGM model only looks at the average growth of all the schools. The next intermediate step is to examine the growth differential between the students in the TIF-LEAP schools and those in the comparison schools. This involves adding into

TABLE V.3

Estimation Results from Longitudinal HLM Models: Mathematics

		Parameter	Unconditional Mean Model	Unconditional Growth Model	Final Growth Model
Fixed Effects					
Initial Status π_{0i}	Intercept	γ_{00}	353.87 (5,391.24)**	349.79 (4,185.28)**	350.08 (4,040.67)**
Rate of Change π_{1i}	TIME	γ_{10}		2.88 (112.53)**	2.82 (89.01)**
TIF-LEAP Status π_{2i}	TIF	γ_{20}			–1.48 (9.60)**
	TIME*TIF	γ_{21}			0.34 (5.76)**
Random Effects					
Level-2	In Initial Status	σ^2_{u0}	2.01 (278.64)**	0.14 (1.96)*	0.12 (1.67)
	In Rate of Change	σ^2_{u1}		2.03 (290.79)**	1.53 (140.13)**
	Covariance	σ_{01}		–0.22 (6.69)**	–0.17 (1.49)
Level-1	Within-Student	σ^2_e	1.75 (376.59)**	1.44 (246.76)**	1.24 (123.24)**
Goodness-of-Fit					
Deviance			279,675.96	270,664.80	230,578.46
AIC			279,681.96	270,676.80	230,594.46
BIC			279,707.76	270,728.39	230,663.24
Number of Observations			**40,067**	**40,067**	**40,067**

*$p < .05$; **$p < .01$

equation (4) the variable TIF_{it}, a dummy variable, which equals 1 for students in TIF-LEAP schools and 0 for those in comparison schools. TIF_{it} is a time-varying variable because, for the same student, it equals 0 in the years before the start of the TIF-LEAP initiative and equals 1 thereafter. The composite form of this conditional growth model is as follows:

$$Y_{it} = \gamma_{00} + \gamma_{10} TIME_{it} + \gamma_{20} TIF_{it} \\ [u_{0i} + u_{1i} TIME_{it} + e_{it}] \quad (5)^{13}$$

The estimated coefficient on TIF_{it} is -0.89 in terms of mathematics test scores, which is statistically significant and indicates that the TIF-LEAP schools, on average, had lower test scores over the course of the initiative. This estimate refers only to the level of the test scores, rather than the growth of the test scores that captures the TIF-LEAP effects.[14]

In order to capture the TIF-LEAP effects, i.e., the growth of students' test scores as a result of TIF-LEAP, an interaction term, $TIME_{it} \times TIF_{it}$, is added into equation (5):

$$Y_{it} = \gamma_{00} + \gamma_{10} TIME_{it} + \gamma_{20} TIF_{it} + \gamma_{21} TIME_{it} \\ \times TIF_{it} + [u_{0i} + u_{1i} TIME_{it} + e_{it}] \quad (6)$$

- γ_{00} now is the population average true initial test scores for students in the pre-initiative status;
- γ_{10} now is the population average annual rate of true change for students in the comparison schools;
- γ_{20} now is the population difference in initial test scores between TIF-LEAP status;
- γ_{21} now is the population difference in rate of growth between TIF-LEAP status;

The definitions of Y_{it}, π_{0i}, π_{1i}, $TIME_{it}$, and TIF_{it} have not changed.

For the purpose of presentation, equation (6) is the condensed version of the final growth model (FGM) because it leaves out the control predictors.[15] The residual variances of the UGM, including both initial status and rate of change are substantial and statistically significant, an indication that one should add the control predictors.

The last column of *Table V.3* shows the estimation results of the FGM.[16] The main results include:

- The true initial test score for the average comparison student is 350.08. This is higher than the initial test scores of 349.79 in the UGM model.
- The true initial test score for the average TIF-LEAP student is 1.5 points lower. This is as expected because of the design of the initiative.
- The true annual rate of change in test scores for the average comparison student is 2.82. This is somewhat lower than the rate of change of 2.88 in the UGM model.
- The true annual rate of change in test scores for the average TIF-LEAP student is 0.34 points higher than that of the comparison student. This estimate is statistically significant at $p < .01$. It indicates that the mathematics test scores of TIF-LEAP students, on average, grow at a rate 12% greater than that of the comparison students.
- The within-student variance is substantially reduced to 1.14 by allowing for the TIF-LEAP effect and adding additional student characteristic control variables. The between-students variance is also significantly reduced. Moreover, the covariance between Level-2 equations is no longer statistically significant. All of these indicate that the FGM has a greater explanation power than the UMM and UGM models.
- The statistics of goodness-of-fit improved substantially in this final model specification. This is due to two reasons: (1) student characteristic controls were added, which significantly reduced the unexplained variation in the error terms, as can be seen from the previous point; and (2) the TIF and TIF*TIME variables both significantly predict the variation in student achievement and therefore improved the fitness of the model.

The same analyses of the TIF-LEAP effects on reading test scores have been conducted. The main findings are similar to those of mathematics, as the estimation results in *Table V.4* show.

- The true initial test score for the average comparison student is 344.86. This is a little higher than the initial test scores of 344.38 in the UGM model.
- Similar to the case of the mathematics test, the true initial test scores of the TIF-LEAP students is 2.5 points lower.
- The true annual rate of change in test scores for the average comparison student is 3.44. This is somewhat lower than the rate of change of 3.48 in the UGM model.
- The true annual rate of change in test scores for the average TIF-LEAP student is 0.44 points higher than that of the comparison students. This estimate is statistically significant at $p < .01$. It indicates that the reading test scores of TIF-LEAP students, on average, grow at a rate 13% greater than that of the comparison students.
- The within-student variance and the goodness-of-fit statistics follow the same pattern as mathematics.

TABLE V.4

Estimation Results from Longitudinal HLM Models: Reading

		Parameter	Unconditional Mean Model	Unconditional Growth Model	Final Growth Model
Fixed Effects					
Initial Status π_{0i}	Intercept	γ_{00}	349.45 (4,999.05)**	344.38 (3,637.41)**	344.86 (3,583.17)**
Rate of Change π_{1i}	TIME	γ_{10}		3.48 (140.16)**	3.44 (112.44)**
TIF-LEAP Status π_{2i}	TIF	γ_{20}			–2.48 (16.17)**
	TIME*TIF	γ_{21}			0.44 (7.83)**
Random Effects					
Level-2	In Initial Status	σ^2_{u0}	2.07 (283.34)**	0.06 (1.29)	0.05 (1.08)
	In Rate of Change	σ^2_{u1}		2.40 (329.76)**	2.37 (314.02)**
	Covariance	σ_{01}		–0.53 (17.22)**	–0.46 (10.14)**
Level-1	Within-Student	σ^2_e	1.79 (382.32)**	1.35 (232.63)**	1.15 (131.54)**
Goodness-of-Fit					
Deviance AIC BIC			280,301.1 280,307.1 280,332.9	267,396.0 267,408.0 267,459.5	227,138.8 227,154.8 227,223.5
Number of Observations			**39,571**	**39,571**	**39,571**

**$p < .01$

Cross-sectional Hierarchical Linear Models

Cross-sectional HLM models are used to analyze the relationship between SLOs and student achievement in mathematics and reading in TIF-LEAP schools.[17] As mentioned above, HLM models are needed due to the nested feature of the data, i.e., students nested within classrooms and classrooms nested within schools. First, because students in the same classroom or school frequently are more similar to each other than to students in another classroom or school, two of the assumptions of classical OLS regression are violated: the assumption of independence of observations and their associated residuals, and the assumption of homoscedasticity. Second and more importantly, when the nested nature of the data is not taken into account, the possible heterogeneity of regression slopes among groups is ignored.

By taking into account the nested nature of the data, cross-sectional HLM reduces the standard errors and the inflation of the Type I error is avoided. In HLM, separate regression equations are determined for each level of the data and the issue of correlated errors is avoided. HLM estimates the relationships at the individual and group levels and allows for relationships at the student level to vary across groups, as well as for modeling cross-level interactions.[18] The cross-sectional HLM models are fitted to determine two relationships: (1) the quality of SLOs and student achievement; and (2) the attainment of SLOs and student achievement.

The first cross-sectional analysis is conducted using 2008-09 data, the first year of SLO implementation. The analysis is conducted for the initial five TIF-LEAP elementary schools and the five TIF-LEAP middle schools. The elementary schools are Billingsville, Druid Hills, Highland Renaissance Academy, Reid Park, and Shamrock Gardens. The middle schools are Bishop Spaugh, John Taylor Williams, Martin Luther King, Jr., Sedgefield Middle, and J.W. Wilson.

The second and third cross-sectional HLM analyses are conducted in 2009-10 and 2010-11, respectively. These analyses included the same schools as the first model plus the two additional elementary schools, Berryhill and Lincoln Heights, that entered the initiative in 2009-10.

The four TIF-LEAP high schools (E. E. Waddell, Garinger, West Charlotte and West Mecklenburg) that entered the initiative could not be included due to data limitations. Three major issues limiting the ability to analyze high schools emerged during the study:

1. When students move from middle school to high school, they are difficult to track and match within the district's data system.
2. Even with those students who could be identified and tracked, one faces the problem of how to access pre-test controls. Algebra and English are the corresponding tested-subject classes to grade level mathematics and reading. An attempt to use grade 8 mathematics as the pre-test for Algebra generated statistical test results that indicate the correlation between the two is low, preventing its use as a pre-test control in the cross-sectional HLM model.
3. Students may take EOC courses in different grades. This causes further loss of observations and creates complications in terms of alignment of pre-tests and grades.

For each of these analyses, an unconditional mean model (UMM) is fitted first to determine the proportion of variance at the student, classroom and school levels. When the variance at the third level is not statistically significantly different from zero, a two-level HLM is employed. If the proportion of variance at all levels is different from zero, the next analysis proceeds to fit the three-level model. At the middle school level in both mathematics and reading the variance at Level-3 was virtually zero. In these cases the analysis proceeds to fit a two-level model.[19]

Because a large set of information is presented in this section, it is organized by a summary of the main findings of the cross-sectional HLM models; a description of the data sample used and key variables; and a detailed description of the models and empirical results.

Summary of Findings of the Cross-Sectional Hierarchical Linear Models

The first year of the implementation of SLOs is 2008-09. The full SLO effects on student achievement are expected to phase in over the first three years of the initiative. The findings in the first year are:

- There are positive, statistically significant associations between the attainment of Target SLOs and student achievement both in mathematics and in reading.[20]
- There is no statistically significant association between the quality of SLOs (as based on rubric ratings) and student achievement.

For the 2009-10 analysis, key findings are:

- There are positive, statistically significant associations between the quality of SLOs and student achievement. This finding means that a teacher's SLO rating relates positively to student achievement in elementary school mathematics, elementary school reading, and middle school mathematics.
- There are positive, statistically significant associations between the attainment of SLOs and student achievement at the elementary school level. This finding means that the students whose teachers met their SLOs achieved higher scores in elementary school mathematics and reading.

For the 2010-11, key findings are:

- There is a positive, statistically significant association between the quality of SLOs and student achievement in mathematics in grade five.
- There is a positive, statistically significant association between the attainment of SLOs and student achievement in reading in grade six.

Description of the Data Sample

Because of the similarity of the cross-sectional HLM models conducted on different school years, the 2009-10 analysis is presented here as a representation of the data sample and the model specification. Also presented in detail are the associated 2009-10 results.

The data sample in the cross-sectional analyses differs from that used in the longitudinal analyses for two reasons: (1) only TIF-LEAP schools are included in the analyses because comparison schools have not implemented SLOs; and (2) the number of students included in the TIF-LEAP schools is smaller due to data limitations (e.g., lack of data for students from previous years).[21]

The most distinct data reduction occurs at the sixth grade level (both mathematics and reading) and in seventh grade mathematics. The overall data size reduction is 43%. Although the reduction in data is substantial, all groups are affected in a similar manner. Data loss appears to be random.

There are two main reasons for the data loss. First, prior achievement data are not available for more than one-third of the students. Second, the analysis excludes those classrooms that are comprised of fewer than seven students with both pre-test (2008-09) and post-test (2009-10) information.

The TIF-LEAP sample discussed above pertains to the analysis of the association between the quality of SLOs and student achievement. The number of students and classrooms included in the analyses is presented in *Table V.5*.[22]

TABLE V.5

Data Sample, 2009-10

	Elementary School Mathematics	Elementary School Reading	Middle School Mathematics	Middle School Reading
Number of Students	603	664	931	1,561
Number of Classrooms	46	44	82	108

Student, Classroom, and School Variables

The outcome variable is a standardized student test score (z-score) in mathematics or reading on the 2009-10 North Carolina EOG General Test.[23] Since the North Carolina state tests are not vertically aligned, and because more than one grade is included in each analysis (grades 4 and 5 at the elementary school level and grades 6, 7, and 8 at the middle school level), the scaled scores are transformed into z-scores grade by grade. Specifically, z-score variables are created in mathematics (ZM0910) and reading (ZR0910), and include observations in z-scores on students from grade 4 to grade 8, respectively. Several student level, classroom level, and school level variables are collected and explored as potential predictors of student achievement. Note that information on free and reduced price lunch (FRPL) status is not available from CMS in 2009-10 for individual students.

Description of Cross-sectional Hierarchical Linear Models

Generally speaking, the cross-sectional HLM models are specified at three levels: student, classroom and school. However, depending on the proportion of variance at each of the three levels, the three-level model reduces into a two-level model in some cases. Specifically, a three-level unconditional mean model (UMM) is first specified and used to determine the levels by investigating the variance at each level.[24] After determining levels, the SLOs and other control variables are entered into the three-level (or two-level) UMM, initially allowing all the intercepts and slopes to vary.

The cross-sectional HLM model presented below is the most generous model that has three levels and allows all the intercepts and slopes to be random. In the actual final model, some of the intercepts and slopes are fixed due to the problem in estimating too many parameters, and some of the control variables are not included due to non-significance of the estimated coefficients on them.

Level-1 Equation (student level):

$$Y_{ijk} = \pi_{0jk} + \sum_{p=1}^{P} \pi_{pjk} a_{pijk} + e_{ijk} \qquad (7)$$

- Y_{ijk} is the test score for student i in classroom j of school k;
- π_{0jk} is the intercept for classroom j and school k;
- a_{pijk} with $p = 1, 2, \ldots, P$, is one of the student characteristics that predicts achievement;
- π_{pjk} is the corresponding Level-1 coefficient that indicates the direction and strength of association between students' characteristic and test scores in classroom j of school k; and
- e_{ijk} is the residual with $E(e_{ijk}) = 0$ and $var(e_{ijk}) = \sigma_e^2$. It measures the deviation i's test scores from π_{0jk}, controlling students' characteristics.

Level-2 Equations (classroom level):

For Level-1 intercept:

$$\pi_{0jk} = \beta_{00k} + \sum_{q=1}^{Q_0} \beta_{0qk} X_{qjk} + u_{0jk} \qquad (8)$$

For Level-1 slopes:

$$\pi_{pjk} = \beta_{p0k} + \sum_{q=1}^{Q_p} \beta_{pqk} X_{qjk} + u_{pjk} \qquad (9)$$

where $p = 1, 2, \ldots, P$; $q = 1, 2, \ldots, Q_p$.

- β_{00k} is the intercept for school k, after controlling the effects of X_{qjk}, in modeling the classroom intercept π_{0jk};
- X_{qjk} is a classroom characteristic used as a predictor of the classroom effects π_{0jk} and π_{pjk}; β_{0qk} and β_{pqk} are the associated coefficients, respectively; and
- u_{0jk} and u_{pjk} are the Level-2 residuals with $E(u_{0jk}) = E(u_{pjk}) = 0$, $var(u_{0jk}) = \sigma_{u0}^2$, $var(u_{pjk}) = \sigma_{up}^2$, $cov(u_{0jk}, u_{pjk}) = \sigma_{0p}$, $cov(u_{1jk}, u_{pjk}) = \sigma_{1p}$, etc. It permits the Level-1 parameters to vary stochastically across classrooms and schools.

Level-3 Equations (school level):

For Level-2 intercepts:

$$\beta_{00k} = \gamma_{000} + \sum_{s=1}^{S_0} \gamma_{00s} W_{sk} + \vartheta_{00k} \qquad (10)$$

$$\beta_{p0k} = \gamma_{p00} + \sum_{s=1}^{S_p} \gamma_{p0s} W_{sk} + \vartheta_{p0k} \qquad (11)$$

For Level-2 slopes:

$$\beta_{0qk} = \gamma_{0q0} + \sum_{s=1}^{S_q} \gamma_{0qs} W_{sk} + \vartheta_{0qk} \quad (12)$$

$$\beta_{pqk} = \gamma_{pq0} + \sum_{s=1}^{S_{pq}} \gamma_{pqs} W_{sk} + \vartheta_{p0k} \quad (13)$$

where $p = 1, 2, \ldots, P$; $q = 1, 2, \ldots, Q_p$; $s = 1, 2, \ldots, S_{pq}$.

- γ_{000} is the grand mean test score of all the students; and
- ϑ_{00k} is the residual with $E(\vartheta_{00k}) = 0$ and $var(\vartheta_{00k}) = \sigma^2_{\vartheta 0}$. It permits the Level-2 parameters to vary stochastically across schools.
- γ_{pq0} is the intercept term in the school level model for β_{pqk};
- W_{sk} is a school characteristic used as a predictor for the school effect; γ_{pqs} is the corresponding Level-3 coefficient that represents the direction and strength of association between school characteristics W_{sk} and β_{pqk}; and
- ϑ_{pqk} is a Level-3 random effect that represents the deviation of school k's Level-2 coefficient, β_{pqk}, from its predicted value based on the school level model.

Elementary School Mathematics: Analysis of the Quality of SLOs, 2009-10

A total of 603 students in 46 classrooms are included in the elementary school mathematics analyses of the association between teachers' SLO ratings and student achievement. Within the unconditional model, the proportion of variance at the student level is 79%, the proportion of variance at the classroom level is 14%, and the proportion of variance among schools is 7%. *Table V.6* shows the point estimates and standard errors of individual predictors.

As shown in the table, the pre-test (ZM0809) is a significant positive predictor of student achievement in 2009-10 ($p < .001$). FRPL and SWD status as well as the number of out-of-school suspensions (OSS) are statistically significant but negatively associated with student achievement ($p = .035$, $p = .010$, and $p = .005$, respectively). African American status is also a negative predictor of student achievement ($p = .040$). There are no significant predictors of student achievement at the individual school level.

There is a positive, statistically significant association between the quality of SLOs and student achievement in elementary school mathematics ($p = .033$). In terms of z-scores, a one level increase in SLO rating is associated with an increase in the z-score by 0.17 points. This improvement is substantial because it translates into a 1.44 point increase in the mathematics test scores, indicating that the quality of the SLOs makes a difference in student performance.

Elementary School Reading: Analysis of the Quality of SLOs, 2009-10

A total of 664 students in 44 classrooms are included in the elementary school reading analyses. The analyses focus on the association between teachers' SLO ratings and student achievement on the elementary reading EOG tests. Within the unconditional model, the proportion of variance at the student level is 76%, the proportion of variance at the classroom level is 21%, and the proportion of variance among schools is 3%.

Similar to the findings on elementary school mathematics, the pre-test is a significant positive predictor of student achievement in 2009-10 ($p < .001$); SWD status and the number of out-of-school suspensions (OSS) are negative and statistically significantly associated with student achievement. Gender is a significant predictor of reading with girls performing at higher levels than boys.

There is a positive, statistically significant association between the quality of SLOs and student achievement in elementary school reading. In terms of z-scores, a one level increase in SLO rating is associated with an increase in the z-score of 0.21 points, which translates into a 1.89 point increase in the reading test scores.

Middle School Mathematics: Analysis of the Quality of SLOs, 2009-10

A total of 931 students in 82 classrooms are included in the middle school mathematics analyses of the association between teachers' SLO ratings and student achievement. Within the unconditional model the proportion of variance

TABLE V.6

Elementary School Mathematics Quality of SLOs: Significant Effects, 2009-10

Fixed Effect	Coefficient	Standard Error	*t*-Ratio	*p*-Value
For INTRCPT1, π_{0jk}				
For INTRCPT2, β_{00k}				
INTRCPT3, γ_{000}	0.093	0.035	2.682	0.037
For ISS_L2, β_{01k}				
INTRCPT3, γ_{010}	1.778	0.811	2.191	0.034
For SLOs, β_{02k}				
INTRCPT3, γ_{020}	0.166	0.075	2.205	0.033
For ZM0809 slope, π_{1jk}				
For INTRCPT2, β_{10k}				
INTRCPT3, γ_{100}	0.746	0.025	29.500	0.000
For SWD slope, π_{2jk}				
For INTRCPT2, β_{20k}				
INTRCPT3, γ_{200}	–0.216	0.084	–2.581	0.010
For FRPL slope, π_{3jk}				
For INTRCPT2, β_{30k}				
INTRCPT3, γ_{300}	–0.180	0.086	–2.107	0.035
For AF_AM slope, π_{4jk}				
For INTRCPT2, β_{40k}				
INTRCPT3, γ_{400}	–0.207	0.100	–2.059	0.040
For AM_IND slope, π_{5jk}				
For INTRCPT2, β_{50k}				
INTRCPT3, γ_{500}	–0.022	0.522	–0.043	0.966
For ASIAN slope, π_{6jk}				
For INTRCPT2, β_{60k}				
INTRCPT3, γ_{600}	0.199	0.162	1.228	0.220
For HISP slope, π_{7jk}				
For INTRCPT2, β_{70k}				
INTRCPT3, γ_{700}	–0.017	0.105	–0.160	0.873
For M_RACIAL slope, π_{8jk}				
For INTRCPT2, β_{80k}				
INTRCPT3, γ_{800}	–0.059	0.136	–0.436	0.663
For OSS slope, π_{9jk}				
For INTRCPT2, β_{90k}				
INTRCPT3, γ_{900}	–0.052	0.018	–2.825	0.005

TABLE V.7

Middle School Mathematics Quality of SLOs: Significant Effects, 2009-10

Fixed Effect	Coefficient	Standard Error	*t*-Ratio	*p*-Value
For INTRCPT1, π_{0j}				
INTRCPT2, β_{00}	–0.006	0.031	–0.205	0.838
OSS_L2, β_{01}	–0.048	0.019	–2.490	0.015
SLOs, β_{02}	0.151	0.054	2.779	0.007
For ZM0809 slope, π_{1j}				
INTRCPT2, β_{10}	0.656	0.025	26.401	0.000
For ISS slope, π_{2j}				
INTRCPT2, β_{20}	–0.033	0.016	–2.057	0.040
For OSS slope, π_{3j}				
INTRCPT2, β_{30}	–0.021	0.006	–3.605	0.001

at the student level is 70% and the proportion of variance at the classroom level is 30%. There is no variance at the third level, school; consequently, a two-level hierarchical linear model has been fitted to the data. *Table V.7* provides the significant effects of individual predictors at the student and classroom levels.

The pre-test (student achievement in 2008-09) is a significant positive predictor of student achievement in 2009-10 ($p < .001$). The number of in-school suspensions (ISS) and out-of-school suspensions (OSS) are negative and statistically significantly associated with student achievement ($p = .040$ and $p = .001$, respectively). At the second level, the average number of out-of-school suspensions (OSS_L2) predicts student achievement negatively ($p = .015$).

There is a positive, statistically significant association between the quality of SLOs and student achievement ($p = .007$) in middle school mathematics. In terms of *z*-scores, a one level increase in SLO rating increases the *z*-scores by 0.15 points, indicating that the quality of the SLOs indeed makes a difference on student performance.

Middle School Reading: Analysis of the Quality of SLOs, 2009-10

A total of 1,561 students in 108 classrooms are included in the middle school reading analyses of the association between teachers' SLO ratings and student achievement. The proportion of variance among schools is approximately zero. Consequently, a two-level model has been fitted to these data. For the two-level model, the proportion of variance at the student level is 66% and the proportion of variance at the classroom level is 34%.

Similar to the previous analyses, there are significant associations between the pre-test, ISS and OSS and student achievement in middle school reading. There is no statistically significant association between the quality of SLOs and student achievement in middle school reading.

Analysis of SLO Attainment, 2009-10

Cross-sectional HLM analyses are conducted on the attainment of SLOs for elementary and middle school mathematics and reading. The findings are consistent with those on the quality of SLOs.

There is a positive, statistically significant association between attainment of SLOs and student achievement in elementary school mathematics and reading. In terms of *z*-scores, attaining the SLO growth target increases the *z*-scores in elementary reading by 0.15 points, and elementary mathematics by 0.11 points. The attainment of SLOs is not a statistically significant predictor of student achievement in either middle school mathematics or reading.

Summary

This evaluation employed both descriptive statistics and HLM models to examine the TIF-LEAP initiative's impact on student achievement. The descriptive statistics examined the growth trajectory between TIF-LEAP and comparison schools. Longitudinal hierarchical linear models estimated the impact of the initiative on student achievement in relation to the achievement of comparison schools. Finally, cross-sectional HLM analyzed the relationships between SLOs and student achievement in the TIF-LEAP schools.

The TIF-LEAP initiative made a positive impact on student achievement at the participating schools. Results show that (1) the growth rate of students in TIF-LEAP schools is greater than the growth rate of students at the comparison schools during the years under study; (2) there are positive relationships between both the quality and attainment of SLOs and student achievement; and (3) the TIF-LEAP schools also show greater resilience than the comparison schools to district changes that occurred in 2010-11.

Endnotes

1 The quantitative analysis in this chapter is applied only to elementary and middle schools. Grade 3 is not included in the elementary school analyses because a pre-test is not administered in Grade 2. The three issues limiting the ability to analyze high school EOC test scores are discussed in detail later in this chapter.

2 See section on Student Achievement Data regarding 2011-12 in Chapter III.

3 For further discussion of these changed circumstances, see Chapters I and VI.

4 In subsequent cross-sectional analyses no statistically significant differences were found between Class and Target SLOs, therefore results for both type of SLOs were merged.

5 For details on how the SLO ratings are defined and measured, please refer to Chapters III and IV.

6 Table V.2 includes teachers from all subject areas. In terms of the tested-subject teachers, the total numbers of teachers in TIF-LEAP and comparison schools in 2007-08 were 157 and 233, respectively. The patterns of teacher characteristics for tested-subject teachers are similar to those in Table V.2.

7 For details on the comparability issues on reading scores, please refer to the "Technical Report, North Carolina Reading Comprehension Tests", April 21, 2009. See http://www.dpi.state.nc.us/docs/accountability/testing/reports/eogreadingtechman3.pdf.

8 The magnitude of the impact cannot be identified from the descriptive statistics. Further, they do not control for factors such as the impact of the student and teacher characteristics that may account for the differences in student performance.

9 Traditional regression models, such as ordinary least squares (OLS) regression, are methods that analyze data by focusing on one level. In many situations, including educational settings, the observations are clustered into groups, which in turn can be clustered into larger groups. This characteristic is called "nesting."

10 Through controlling the observables and time-invariant unobservables, together with the quasi-experimental design, estimates from longitudinal HLM models are interpreted as estimates on causal treatment effects, with the caveats discussed in the main text.

11 The results of the UMM not only provide a baseline for evaluating the success of subsequent longitudinal models, but also tell us that there are sufficient variations for meaningful cross-sectional analyses, where only between-student variation is explored.

12 An illustration of the corresponding Level-1 and Level-2 specification is as follows: Level 1 Equation: $Y_{it} = \pi_{0i} + \pi_{1i} TIME + e_{it}$, where π_{0i} is now the intercept of i's change trajectory – i's "initial status", i.e., his/her true value of test score at the first test year; π_{1i} is the slope of i's change trajectory – i's "annual rate of change", i.e., his/her yearly rate of change in test scores; Level-2 Equations: $\pi_{0i} = \gamma_{00} + u_{0i}$ and $\pi_{1i} = \gamma_{10} + u_{1i}$.

13 The coefficient of TIF_{it} is fixed. It could be set to vary randomly across students by adding a residual term in equation (5). However, this was not done because no significant variations were found in the process of exploring the data, and adding the random effect term which prevents the model estimation to converge uses up degrees of freedom.

14 It is for this reason that we did not report the estimation results of the CGM in Tables V.3 and V.4.

15 The uncondensed form of FGM is as follows: $Y_{it} = \gamma_{00} + \gamma_{10} TIME_{it} + \gamma_{20} TIF_{it} + \gamma_{21} TIME_{it} \times TIF_{it} + \sum_{p=3}^{P} \pi_p a_{pit} + \sum_{s=2}^{S} \gamma_{0s} \theta_{si} + \sum_{s=2}^{S} \gamma_{1s} \theta_{si} \times TIME_{it} + [u_{0i} + u_{1i} TIME_{it} + e_{it}]$ where $p = 3, 4, \ldots, P$; s = 2, 3, …, S. a_{pit} is a control which captures other time-varying factors, such as student's number of in-school suspensions (ISS) or total absences, that influence student test scores; π_p is the associated coefficient. Since it does not seem that the effects of a_{pit} would systematically differ by TIF-LEAP status, the coefficient is fixed. θ_{si} captures time-invariant factors,such as student's gender or disabilities, that influence student test scores; γ_{0s} and γ_{1s} are the associated coefficient.

16 The evaluation does not report the estimates on a_{pit}, θ_{si}, and their interaction terms.

17 Cross-sectional, instead of longitudinal, analysis is used here because the SLOs are not vertically aligned to be directly comparable.

18 In comparison to cross-sectional HLM, longitudinal HLM models do not have separate levels for classrooms and schools.

19 The three-level HLM models are presented. The two-level models are conceptually identical to the three-level models, but they do not include the third level and the notation is simpler.

20 See endnote 4.

21 Missing previous year's test scores was not a problem for the longitudinal analysis because the regression technique used was sufficiently robust to handle missing observations.

22 Although the presentation here is from the 2009-10 analysis, the same problem and strategy applies for the 2010-11 analysis.

23 z-scores were calculated using the mean and standard deviation of the state EOG test scores for a given subject, test year, and grade. This is not needed in the longitudinal analysis because the longitudinal HLM model pools information on the test scores from all grades in a certain test year and explores the variation year after year to achieve the estimate, i.e., to explore if the average on all grades' test scores grow or not, so the lack of vertical alignment is not an issue.

24 The analyses of variance are similar to the ones in longitudinal HLM and the results are omitted. In general, the variance decreases from the UMM to the final model.

CHAPTER VI

Perspectives on the TIF-LEAP Initiative and Performance-Based Compensation

The subject of performance-based compensation in education is nuanced and complex. It may bring up questions about how teacher and principal performance can be measured in fair, transparent, and inclusive ways. It may challenge deeply held personal beliefs about why and how one teaches. It may cause suspicions about the motives and intentions of those for whom and with whom one works. If these issues are not disquieting enough, there is also the somewhat mottled history of performance-based compensation in education that suggests connecting teaching, learning, and compensation is challenging.

However, as pointed out earlier, TIF-LEAP teachers and principals, together with CMS leaders, were experienced consumers of pay incentives from earlier district and state programs. Additionally, they were, during the design phase of the initiative, beneficiaries of the learnings from newer and more successful performance-based compensation approaches, such as ProComp in the Denver Public Schools. All in all, including the federal funding support, CMS was in an opportune position to initiate a successful new performance-based compensation initiative.

Survey and interview responses show that, at the outset, district and community stakeholders held favorable views of the possibilities of connecting additional compensation to student outcomes. Some in the district welcomed the opportunity presented by the TIF-LEAP initiative. A central administrator observed: "The [school] community is ready for pay for performance." Another thought that the initiative would "increase respect from the community." A school board member said, "The time and commitment necessary to make teaching a career should be compensated."

Over the course of the initiative, the participant and stakeholder outlook on the potential of performance-based compensation remained favorable, even during several unusual and complicating district events. However, their viewpoints were neither uniform nor uncritical; and though experienced and primed at the outset, TIF-LEAP participants, CMS leaders, and other stakeholders still had more to learn about performance-based compensation.

Perceptual Data Sources and Respondents

TIF-LEAP schools joined the initiative in phases, increasing not only the number of teacher and principal participants over time, but also the amount of survey and interview feedback on the progress of the initiative. The TIF-LEAP team, the Steering Committee, and many members of the central administration provided additional perspectives on the initiative's progress and impact. Other viewpoints came from Board members, students, parents and community members.

The educator survey, administered annually between November and January by the Community Training and Assistance Center, contained items intended to tap not only the views of teacher and principal participants in the TIF-LEAP schools but also those of teachers and principals in non TIF-LEAP schools across the district. Individual and focus group interview instruments addressed topics similar to those on the survey with teachers and principals from both TIF-LEAP and non TIF-LEAP schools, as well as with Board members and other district leaders. Parents and students were interviewed in focus groups and business and community leaders in individual interviews. The numbers of participants and the methods of surveying and interviewing are presented in Chapter III.

Surveys and interviews focused on: (1) the general knowledge and beliefs held by respondents about performance-based compensation; (2) the efficacy of and preference for various types of incentives and performance-based compensation approaches; and (3) the quality and maintenance of district support services, including curriculum and instruction, accountability, technology, professional development, teacher evaluation, and instructional leadership.

As the SLO approach to performance-based compensation and the VAM approach respectively were underway in the TIF-LEAP schools, survey and interview questions were expanded to learn TIF-LEAP participant perspectives on the helpfulness of the implementation supports, the fidelity to the SLO process, the understanding of the VAM, and the effectiveness of communication. Participant views on changes in district priorities relating to the initiative also were sought through interviews and surveys. In the final year (2011-12), interviews and focus group protocols centered on lessons learned through the eyes of the participants and other stakeholders.

Participant and Stakeholder Perspectives: Five Observations

Analyses of the responses from the surveys and interviews, augmented by informal observations and by evaluation of artifacts, reveal consistent and telling patterns in the perspectives of TIF-LEAP participants and other stakeholders on the goals, value, implementation, and impact of the initiative. Some viewpoints were apparent from the outset of the initiative, others became evident as the initiative progressed, and still others changed over time. Besides weighing strengths and challenges of two performance-based compensation approaches, participants and stakeholders identified the effects on teaching and the impact of changes in the initiative, some of which were mid-course corrections and others which represented changes in district priorities and conditions.

Over the five years, participant and stakeholder perspectives highlight (1) what was of consistent

and enduring significance in the TIF-LEAP schools; (2) how the district culture, beliefs, and systems impacted the initiative; and conversely, (3) how the initiative impacted the district over a five-year period. The experiences and interpretations gleaned from interviews and surveys were examined and summarized, and are organized into the following five observations:

1. TIF-LEAP stakeholders supported the concept of performance-based compensation initially and throughout the initiative.
2. Teachers and principals in TIF-LEAP schools used the Student Learning Objectives process to improve student learning.
3. Leaders in Charlotte-Mecklenburg Schools underscored the potential of performance-based compensation as a pathway to systemic reform.
4. The quality of the support from many district systems proved critical to the outcomes of the initiative.
5. Parents and community members supported linking student achievement and compensation for teachers and principals.

Actual words from stakeholders, together with survey responses, explain both the common themes and the diversity of views and responses to the initiative and, perhaps more than any other data, demonstrate why *compensation reform is about more than money*.

Stakeholder Support of Performance-Based Compensation

Observation One: TIF-LEAP stakeholders supported the concept of performance-based compensation initially and throughout the initiative.

Baseline Year: 2007-2008. Baseline survey and interview responses[1] point to readiness in the CMS community to implement performance-based compensation, in general, and to actualize the TIF-LEAP initiative, in particular. Overall, the first year survey and interview data indicate that the community, both internally and externally, engaged positively with the fundamental concept of teacher performance-based compensation—that a link can be made between what teachers earn and students learn—but only if, as one district leader indicated, "the initiative heeds lessons of the past, [the link] is developed collaboratively, and provides supports to improve instruction."

Within schools, teachers and principals are generally amenable to an array of approaches to earning performance-based compensation. An examination of *Table VI.1* reveals that survey responses from all district schools in the first year of the initiative (2007-08) and all three role groups of school site educators—principals, assistant principals, and teachers—are largely in agreement with performance-based compensation as a legitimate approach to improving student achievement, although teachers are slightly less likely to strongly agree or agree than principals, especially in regard to monetary rewards for principals.

It is of particular note that teachers are nearly as likely to strongly agree or agree with awarding pay based on student improvement by *classroom*, that is, by individual teacher, as they are with awarding additional compensation based on student improvement by *school*. In both cases, teacher agreement in CMS approached 75% at a time when teachers in other states and districts are skeptical of the concept of connecting student outcomes to teacher compensation and/or performance appraisal.

The results of this baseline survey (*Table VI.1*) also reveal that a plurality of site educator responses is in agreement with the potential of various types of performance-based awards. These findings indicate openness (strongly agree/agree) among site administrators for rewarding teachers who accept assignments in, or remain in, high need or difficult-to-staff schools. More than two-thirds of teachers, principals, and assistant principals support additional compensation for teachers who serve as mentors or teachers who improve their instruction through professional development activities. In addition, site educators indicate solid support for awarding extra compensation to teachers for *outstanding* performance evaluations, but less support for awarding extra compensation for *satisfactory* teacher performance ratings, particularly among principals.

TABLE VI.1

Districtwide Perceptions of Potential Teacher and Administrator Performance Rewards, Baseline Year 2007-08

Performance Rewards	Respondent Group	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
Reward teachers who improve student learning in their schools.	Principals	92 (63.0%)	44 (30.1%)	2 (1.4%)	6 (4.1%)	2 (1.4%)
	Asst. Principals	132 (62.9%)	52 (24.8%)	10 (4.8%)	13 (6.2%)	3 (1.4%)
	Teachers	2,955 (51.2%)	1,964 (34.0%)	263 (4.6%)	427 (7.4%)	166 (2.9%)
Reward teachers who improve student learning in their classes.	Principals	98 (67.1%)	40 (27.4%)	1 (0.7%)	5 (3.4%)	2 (1.4%)
	Asst. Principals	131 (62.4%)	45 (21.4%)	10 (4.8%)	20 (9.5%)	4 (1.9%)
	Teachers	3,034 (52.5%)	1,706 (29.5%)	292 (5.1%)	519 (9.0%)	224 (3.9%)
Reward principals who improve student learning in their schools.	Principals	93 (63.7%)	43 (29.5%)	2 (1.4%)	5 (3.4%)	3 (2.1%)
	Asst. Principals	126 (60.0%)	58 (27.6%)	10 (4.8%)	12 (5.7%)	4 (1.9%)
	Teachers	2,076 (35.9%)	2,315 (40.1%)	421 (7.3%)	670 (11.6%)	293 (5.1%)
Reward teachers who accept assignments to high-need or difficult-to-staff schools.	Principals	90 (61.6%)	38 (26.0%)	3 (2.1%)	11 (7.5%)	4 (2.7%)
	Asst. Principals	136 (64.8%)	48 (22.9%)	10 (4.8%)	14 (6.7%)	2 (1.0%)
	Teachers	3,172 (54.9%)	1,799 (31.2%)	223 (3.9%)	370 (6.4%)	211 (3.7%)
Reward teachers for remaining in high-need or difficult-to-staff schools.	Principals	94 (64.4%)	37 (25.3%)	4 (2.7%)	7 (4.8%)	4 (2.7%)
	Asst. Principals	141 (67.1%)	43 (20.5%)	9 (4.3%)	16 (7.6%)	1 (0.5%)
	Teachers	3,339 (57.8%)	1,643 (28.5%)	232 (4.0%)	352 (6.1%)	209 (3.6%)
Reward teachers for serving in a mentoring capacity to other teachers.	Principals	85 (58.2%)	55 (37.7%)	0 (0.0%)	6 (4.1%)	0 (0.0%)
	Asst. Principals	130 (61.9%)	67 (31.9%)	5 (2.4%)	7 (3.3%)	1 (0.5%)
	Teachers	3,196 (55.3%)	2,092 (36.2%)	155 (2.7%)	247 (4.3%)	85 (1.5%)
Reward teachers for improving their teaching through professional development.	Principals	75 (51.4%)	57 (39.0%)	2 (1.4%)	10 (6.8%)	2 (1.4%)
	Asst. Principals	121 (57.6%)	57 (27.1%)	8 (3.8%)	19 (9.0%)	5 (2.4%)
	Teachers	3,084 (53.4%)	1,909 (33.1%)	231 (4.0%)	409 (7.1%)	142 (2.5%)
Reward teachers for receiving outstanding performance evaluations.	Principals	68 (46.6%)	50 (34.2%)	6 (4.1%)	21 (14.4%)	1 (0.7%)
	Asst. Principals	109 (51.9%)	55 (26.2%)	15 (7.1%)	27 (12.9%)	4 (1.9%)
	Teachers	3,151 (54.6%)	1,587 (27.5%)	287 (5.0%)	517 (9.0%)	233 (4.0%)
Reward teachers for receiving satisfactory performance evaluations.	Principals	29 (19.9%)	41 (28.1%)	4 (2.7%)	54 (37.0%)	18 (12.3%)
	Asst. Principals	62 (29.5%)	51 (24.3%)	12 (5.7%)	70 (33.3%)	15 (7.1%)
	Teachers	1,666 (28.8%)	1,937 (33.5%)	424 (7.3%)	1,411 (24.4%)	337 (5.8%)

Furthermore, parents and community members—critical external constituencies—express support for performance-based compensation at the outset. They were asked directly whether differences in compensation should be separated from student learning or linked to increases in student learning. Approximately two-thirds of the respondents in each group indicate that differences in compensation should be linked to increases in student learning, and when linked to principal and teacher performances in the school and classroom, respondent agreement jumps to more than 85%. They remain on board for the duration of the initiative, as further described in Observation Five.

All in all, views collected in the first year of the initiative indicate the readiness among nearly all district teachers and principals, together with district leaders, parents, and community members, to engage in performance-based compensation and to consider a variety of approaches for doing so.

Over Time. Longitudinal analyses[2] of teacher and principal responses, drawn from multi-year survey data, found that teachers and principals continued to be supportive of performance-based compensation, but there were changes over time in the level of support. As shown in *Figure VI.1*, TIF-LEAP and non TIF-LEAP schools started from the same point in 2007-08 in terms of their general views on teacher and principal compensation; however, the momentum for support of performance-based compensation held stronger in TIF-LEAP schools than in comparison schools, albeit the level of support did decrease over time even in TIF-LEAP schools.[3]

These longitudinal results attest to the relatively consistent commitment of teachers and principals to the tenets of TIF-LEAP, a factor that, most certainly, helped the initiative weather a storm of fiscal difficulty and change in district circumstances and leadership in the later years of the initiative.

SLOs and Student Learning

Observation Two: Teachers and principals in TIF-LEAP schools used the Student Learning Objectives process to improve student learning.

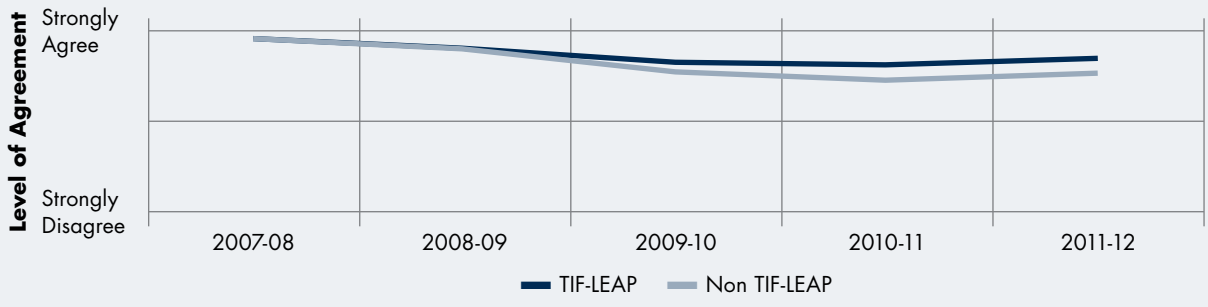
The goals of the TIF-LEAP initiative, as indicated in Chapter I, are always *about more than money*. These goals are particularly reflected in teacher and principal perceptions of SLOs.

In interview and survey responses, teachers and principals demonstrate (1) value in the student achievement features of SLOs, and (2) resolve to focus on the real target of the initiative—student achievement—through five years of implementation and more than just a few distractions. In order to improve learning through SLOs, according to one teacher, "The school administration has to make it important for the staff because SLO results can make a difference for students."

Some principals place SLOs in the spotlight and on center stage in their schools and recognize the importance of their own contributions to teacher planning, support, and success, such as in fostering discussions about high expectations for students or promoting rigor in the measurement(s) of SLOs. Several principals stress the importance of

FIGURE VI.1

Trend Analysis of Teacher and Principal Agreement with Educator Survey Items Related to Teacher and Principal Compensation



integrating the SLO process into other goals and practices in their schools and using it to improve collaboration. Teachers and principals in TIF-LEAP schools also discuss key parts of the SLO process that resonated with them and their colleagues.

TIF-LEAP Teachers

"The [SLO] process helped us pull the pedagogy together."

"The process of writing an SLO is like writing a very thorough lesson plan."

"Colleagues help each other [with SLOs] and [help] the induction of new teachers [into the process]."

"We expect to meet our targets, but the challenge is rigor [in setting and measuring targets]."

TIF-LEAP Principals

"The way of doing business [in the school] is the SLO process. It's the process of how you should be doing the business anyway without bonuses. It recognizes teachers for what they already should be doing. It adds to the spirit of cooperation and includes the support staff to plan with a grade level team. We will always keep the basics of what we learned from the SLOs."

"Developing SLOs forces teachers to target low performing kids and then work harder to pull those kids up."

"[The SLO process] has helped teachers plan effectively for students that are behind. Also, it has helped all the teachers look at student data differently to differentiate instruction."

"The SLO process is what we build other things at the school around. It has gone from an outlier to common practice."

"We have moved from a priority school to a high growth school, and I give a lot of credit for this change to the SLO process."

A preponderance of TIF-LEAP teachers were never wary of accepting individual accountability for student learning outcomes.

Not all teachers are as successful in reaching their growth targets as they hoped, but in their explanations, they frequently speak of getting better next time, such as one who declares, "[This year] I am determined to meet my objectives." In comparison, many teachers who find that they have a low VAM ranking are concerned because they cannot find out where they missed the mark or what they should or could do to improve their performance.

Five years of TIF-LEAP teacher and principal responses to survey items that ask about performance-based compensation preferences show little change in a generally held conviction that teacher performance and additional compensation can be linked to individual student achievement outcomes. These responses provide evidence that a preponderance of TIF-LEAP teachers were never wary of accepting individual accountability for student learning outcomes. Beginning in year one, 89.4% of TIF-LEAP participants strongly agree/agree with performance-based compensation being based on student learning in individual classrooms. Subsequent survey data show agreement remaining high at 88.1% and 80.5% in years two and three and dropping slightly to 77.6% and 75.9% in the last two years of the initiative.[4]

Comparatively, teachers and principals working in non TIF-LEAP schools also start off with high levels of agreement to the concept of tying compensation to student performance, but percentages fall off more distinctly (63.3% and 62.8%) in years four and five than do levels of agreement in TIF-LEAP schools in these final two years.

In the last years of the initiative, the value-added measure and the implementation of a new North Carolina teacher evaluation system led to more teacher caution about how information that connects teacher performance and student outcomes might be used. However, three-fourths of TIF-LEAP teachers and principals remain supportive of the concept of performance-based compensation throughout the initiative.

Interviews with TIF-LEAP teachers and principals over the course of the initiative speak

to how SLOs sharpen their focus on student learning and help them become more effective with their students. Interviewees, both individually and in focus groups, discuss their efforts to improve student learning through the SLO process, describing what matters most to them and what is most distracting.

What Mattered Most in the SLO Process

Several aspects of developing and working with SLOs stand out consistently as being of the highest importance, or making the greatest impact, to a preponderance of TIF-LEAP principal and teacher interviewees—even in the later years, as more secondary teachers enter the initiative and district conditions change. In terms of their work with students, it matters most to teachers and principals that the SLO process provides TIF-LEAP participants with the opportunity to:

1. analyze student baseline data, including pre-tests, for use in their planning and teaching;
2. set individual student growth targets, for both more informative and accurate assessment of student learning; and
3. participate in collegial collaboration in the development of SLOs.

The analysis of baseline data and the setting of student growth targets are clearly delineated steps of the SLO process. Collaboration with colleagues is not a required step in the process, though it is recommended in training materials and presentations. Two of the steps in the SLO process appear to foster and nourish the third one.

Baseline Data. In the SLO process, baseline data refer to (1) longitudinal student achievement data in the district system, most frequently annual state assessments and possibly the results of other formative assessments; and/or (2) the results of pre-assessments administered by the teacher. Interviewees consistently remark on the SLO baseline data analysis step as one that was informative, beneficial, and frequently enlightening, in the conduct of their instructional planning:

> *"The pre-assessment was good to see. Students had skills that I didn't know they had, particularly in writing."*

> *"I like knowing where they start from, via the pretest."*

> *"I use data more effectively and have higher expectations for student growth."*

The effective use of baseline data is facilitated, first of all, through timely access to longitudinal student data—teachers see data before an instructional path already has been mapped out—and secondly, through access to or support in developing useful and telling measures for the pre/post assessment of students. TIF-LEAP principals and teachers indicate that both of these were facilitated through the SLO implementation.

Student Growth. The student growth target, as the term is used in the SLO process, refers to the aggregation of individual student growth targets that the teacher expects to meet at the end of the teaching unit. In either the pre or summative assessment, students might be ranked with all of their classmates by (1) the number of items correct; (2) a proficiency scale; (3) a standardized percentile ranking; or (4) a rubric level. But in all cases, the three questions of most interest are the following: where did the student start? how far did he/she go? and was the distance or growth adequate?

TIF-LEAP teachers feel that setting individual student growth targets is thought provoking, and leads to identifying and planning for each student's needs early in the teaching process. Setting targets also becomes an expression of self-expectation. A *specific goal over a general goal in a classroom is more likely to be achieved*, much like a personal goal that is specific rather than general, such as "I will walk two miles a day" over "I will exercise more."

Collaboration. As noted, collaboration is not a required element of the SLO process, and survey and interview questions did not specifically ask about its role in the process, but TIF-LEAP interviewees frequently bring up the value of working on the process with their peers. Structures for professional learning communities [PLC] are prominent in some TIF-LEAP schools, particularly at the elementary level, but less evident in other schools. An elementary teacher observes: "The [SLO] process clarifies and helps in [PLC] planning meetings as we work to intervene for students." A high school English department chair

tells about how SLOs guide discussions in department meetings. A principal says: "The quality of grade level planning has improved as a direct result of designing and implementing SLOs." When collaboration time is built into the teacher week, participants indicate SLOs provide thought-provoking content for teachers to discuss, and they believe that students benefit.

What Distracted Most from the SLO Process

TIF-LEAP teachers and principals recognize that they were participating in a pilot initiative. They are generally patient with the progress of the implementation and always generous about the efforts of hardworking project staff even as aggravations sprang up. But by year four of the initiative (year three of SLO implementation and year two of the VAM implementation), three areas emerge as particularly distracting to participants:

1. inadequate communication about and/or lack of opportunity to provide input into significant changes to the initiative;
2. misgivings about the perceived intent and accuracy of the district-developed VAM; and
3. issues with software developed to document SLOs, including three versions in as many years.

Communication. Responses from 2007-08, 2008-09, and 2009-10 surveys point to a downward trend in the percentage of principals and teachers who thought that the district communicated effectively with teachers. A plurality of TIF-LEAP principal and teacher respondents believe that the district communicated effectively with principals; however, fewer than half of principals and teachers think that communication from the district to teachers is effective. In both cases, principals agree with these two survey items at a slightly higher level than teachers. Communication in large organizations always requires vigilance, but particularly so in times of change.

Issues related to communication became paramount, according to TIF-LEAP teachers and principals, as the value-added approach was added to the SLO approach in their schools. The VAM required a revision of the payouts, resulting in decreases in the amounts of SLO bonuses for both teachers and principals in order to generate funding for VAM bonuses. A number of participants view the changes as a departure from what they had agreed to do. According to one principal, "the amount of time spent by administrators working on the SLO process doesn't correlate with the [reduced] compensation provided principals [that resulted from paying out the VAM bonuses from the same funding pot]."

Further, teachers and principals feel that they had not been asked for input or provided with information before the new plan went forward. Some observers interpret this oversight as a "fragmentation" between divisions of the district. A principal explains this viewpoint:

> *" We see a significant disconnect between TIF-LEAP, under Curriculum and Instruction division and [value-added] pay for performance, under the accountability division…The result of this fragmentation is that now there are two different sets of numbers from two different offices about teacher effectiveness."*

TIF-LEAP teachers and principals feel a perceived disconnect which they believe was heightened by the need for more two-way communication between district decision-makers and the school sites.

Value-Added Measure. Teachers, in particular, are vocal about the character of the VAM implementation during the last two years of the initiative. Besides misunderstandings about and disagreements with the VAM calculation and its application, there is distrust about its purpose.

> *" Some of us do not like the idea of VAM. You are only testing kids one way. For kids who are struggling, other methods of testing should be implemented."*

> *" [VAM] is an urban thing. They have found a new way to play with the less fortunate; they always lose out."*

> *" The VAM is another number to label teachers."*

> *" How can I [special education teacher] measure my impact on the students I serve? The data folks [VAM administrator] won't let me see the list of students I impact. I see 90 students a week, but don't know which ones I'm credited with."*

On the other hand, there are those teachers who are open to VAM possibilities.

> *"It's positive to be using value-added [methodology] in order to level the playing field, but the definitions [referring to variables included in the model] have to be right."*

> *"The jury is still out on the impact [of VAM] and revisions are still coming."*

> *"The haves and have-nots are [talking about] the things that make them alike and different [on the VAM]."*

> *"I like VAM because [our] teachers will be ranked with all of the other teachers in CMS. The perception out there is that [our] teachers aren't effective, but in fact, look at the value [some of them] added."*

These voices, however, were obscured amidst the disquiet caused by participant reactions to what they perceived as the dismissive nature of the VAM implementation.

While teachers and principals in the TIF-LEAP schools express numerous concerns about the VAM, they are comparatively more open to the concept of using growth measures than are other teachers and principals in the district. Teachers and principals in the non TIF-LEAP schools are far less supportive of using VAMs to reward teachers than those in TIF-LEAP schools with only 34.8% and 36.9% strongly agree/agree in years 2011 and 2012, respectively, as shown in *Table VI.2*.

Further, responses show a plurality of TIF-LEAP teachers and principals strongly agree/agree (50.0% in 2011 and 54.7% in 2012) with rewarding principals whose schools show valued-added growth on the state-mandated assessment. By comparison, only a third of the non TIF-LEAP respondents strongly agree/agree that principals should be awarded when their schools show value-added growth on these assessments.

These survey results indicate that while the VAM implementation is problematic and distracting, it does not completely dissuade TIF-LEAP teachers and principals from their interest in performance-based compensation in the last two years, including the use of the value-added approach.

Software. The software glitches that frustrated participants are not unusual in new technology implementation, but there were changes to the software each year, which teachers said cost them valuable time and made more than a few teachers suggest that the level of effort was not

TABLE VI.2

Use of Value-Added Measures to Reward Teachers, 2011 and 2012

		Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
Use value-added measures (VAM) to reward eligible teachers who show gains on state-mandated assessments.						
2011	TIF-LEAP	65 (17.4%)	141 (37.7%)	78 (20.9%)	51 (13.6%)	39 (10.4%)
	Non TIF-LEAP	367 (10.7%)	828 (24.1%)	867 (25.3%)	744 (21.7%)	625 (18.2%)
2012	TIF-LEAP	47 (23.2%)	59 (29.1%)	51 (25.1%)	22 (10.8%)	24 (11.8%)
	Non TIF-LEAP	266 (13.0%)	490 (23.9%)	485 (23.7%)	431 (21.0%)	377 (18.4%)
Use value-added measures (VAM) to reward eligible principals whose schools show gains on state-mandated assessments.						
2011	TIF-LEAP	62 (16.6%)	125 (33.4%)	94 (25.1%)	46 (12.3%)	47 (12.6%)
	Non TIF-LEAP	319 (9.3%)	784 (22.8%)	932 (27.1%)	773 (22.5%)	626 (18.2%)
2012	TIF-LEAP	40 (19.7%)	71 (35.0%)	49 (24.1%)	19 (9.4%)	24 (11.8%)
	Non TIF-LEAP	233 (11.4%)	468 (22.8%)	525 (25.6%)	430 (21.0%)	395 (19.3%)

commensurate with the level of the bonus. Furthermore, according to interviewees, there was a delay in getting the third year SLO web-based platform live for teachers to use, leading to a late start on SLOs; the system was perceived as unreliable about saving data, resulting in lost work; and in the third year, it could not be accessed from home.

" In a way, [this year] the process got in the [way] of the product. It took a long time for teachers to get acclimated to use of the [new software] platform."

" The process is good but the technology is not. The techs should get it right before they give it to teachers. When teachers have to work out kinks, the techs lose credibility."

Ironically, software changes represented the TIF-LEAP team's intent to make the input platform easier—with pull-down menus, questions, and prompts designed to assist teachers and principals in being thorough and to prevent omissions that might result in an SLO being returned to the teacher for revision. By the final year of the project, most issues had been resolved.

Performance-Based Compensation as Systemic Reform

Observation Three: Leaders in Charlotte-Mecklenburg Schools underscored the potential of performance-based compensation as a pathway to systemic reform.

The TIF-LEAP initiative got underway with the advantage of district leaders who believed in the potential of performance-based compensation as a catalyst for reform and as a vehicle for improving student achievement. Over the five years of the initiative, the Board of Education and the district leadership team became more knowledgeable and most remained supportive of the initiative and its possibilities, even as changes in district circumstances hindered the capacity to take performance compensation to scale in the district or sustain it in TIF-LEAP schools. A concerted effort to interview Board members and district leaders each year meant a significant level of district policymakers and decision-makers routinely added their perspective on performance-based compensation and the progress of the initiative (see a summary of interviews in Chapter III).

Board of Education. The original Board members (those in office at the beginning of the initiative) became more fluent over time with the potential of performance-based compensation. Several predicted that the impact of TIF-LEAP in the district would be long-term and systemic, even if not exactly as originally planned. Greater awareness of the nuances and complexity of performance-based compensation—that it goes beyond allocating money, begins with a strong vision of effective teaching, and extends to sound human capital practices and a commitment to leadership and teaching force development—is increasingly evident among Board member responses in each succeeding year of the initiative.

" Successful compensation reform will cost more than any current compensation schemes. But the return on investment includes more effective teaching, sustained over time; retention of effective teachers in the classroom; bench strength in the teaching and principal ranks; and sustainable improvements in student achievement. Ultimately, successful compensation reform has the potential to pay for itself by generating more net producers graduating from our schools."

" Reformation [of public education] must be particularly focused on how we measure and reform [teacher] performance. While there is nothing inherently wrong with honoring advanced degrees and longevity, they, in and of themselves, must not be the only measures we reward."

" We need to identify the key competencies of an effective teacher, the prime motivations of an effective teacher, and the personality traits of an effective teacher. Then we should create multiple profiles and use them to inform our targeted recruiting efforts."

" Curriculum and instruction, in particular, needs to be more responsive to the various developmental needs of teachers as they try to improve their performances. Value-added does nothing unless there's an instructive piece to go with it."

" To be meaningful, compensation reform must be sustainable over the long term, and insulated from the shifting political and economic sands that affect the district's annual operating budget."

Board members had always expected support from the community on performance-based compensation, believing the collective culture of the Charlotte-Mecklenburg area to be strongly influenced by the norms of the financial sector industry, a notion that the annual community survey validates.

> *"The [Charlotte] community lives in a world of performance-based measures and pay so they will support [compensation reform], and they are owed an explanation of what (and why) we intend to do."*

In the third and fourth years (2009-11) of the TIF-LEAP initiative, the Board and district leadership faced a sobering financial picture, much like the rest of the country, and were involved in teacher layoffs and school closures and reorganizations. Though certain that the district would be unable to sustain the current TIF-LEAP initiative at a districtwide scale, they expressed their commitment to performance accountability as a focal point of the new strategic plan.

The fifth and final year (2011-12) of the initiative brought in a largely new slate of Board members who, while facing an opening in the Superintendency and more fiscal austerity, committed to the goal of having highly effective teachers in every classroom. The following observations reflect their initial understanding and opinions regarding the TIF-LEAP initiative.

> *"I talked to teachers and they found the SLO process was helpful and gave the teacher much more insight."*

> *"Most teachers did not have a problem with the SLO process, but the administrators did not have enough knowledge of the process because of the constant changes in school leadership."*

> *"CMS should use the state measure instead of our own [VAM] because it confuses teachers and sets them up for failure."*

The new Board of Education expresses interest in the links between educator performance and compensation, while being mindful of the changed economic circumstances facing the district and the country.

Central Administration. For the duration, the level of knowledge and sustained interest in the TIF-LEAP initiative among central administration was uneven, with some executive and department leaders more participative than others. However, they all express support of the concept of reforming how teachers are evaluated and paid.

Many central leaders affirmed their belief in the core tenets of the SLO process, such as the use of baseline student data in planning instruction, the recognition of effective teaching, the emphasis on accountability for individual student growth, and the value of the site-based program implementation and support. Several central administrators faithfully served as members of the TIF-LEAP Steering Committee. Later, they struggled with the difficulty, as leaders, of supporting compensation reform while carrying out staff layoffs and cutting school programs.

Some central administrators consider the TIF-LEAP design, and approach to supporting schools on SLOs, as a model for further improvements to teacher effectiveness:

> *"TIF-LEAP has made educators pay more individual attention to students' unique sub-group needs, made them more goal-oriented, and [made it] easier to tie student achievement to goals and finite objectives."*

> *"The TIF-LEAP impact has been in asking the right questions, thinking about how we approach the work. It has allowed us to think about our theories and try them out, to question, to test. It has shaped our thinking."*

> *"TIF-LEAP could be a very critical project: how [do we] identify effective teachers by using non-standardized ways to measure their impact on students? We are learning to understand the impact of the SLO process on teaching.... The project is fraught with challenges. Over time [TIF-LEAP] is clarifying it for all of us. It requires enormous amounts of training. We are asking teachers to do something very different. For instance, to judge their own work."*

The potential of using Student Learning Objectives as a component in teacher performance evaluation and human capital management, including teacher recognition, is on the minds of many central administrator interviewees.

"The [teacher evaluation] instrument has five strands of what matters in teaching. I can see close ties to the SLO process: content, leadership, and data. The strands do not operate separately."

"CMS must figure out how to adapt the [TIF-LEAP team] approach for non-EOC/EOG areas to incorporate performance measures."

"I like the SLO concept as one component of a multi-faceted approach [together with] teacher evaluations, academic growth measures (VAM), state measures, classroom observations."

"Teachers are getting deserved recognition through the TIF-LEAP department. It has sent us toward seeking out data."

Central administrator interviewees express cogently the role of the TIF-LEAP initiative in directing district focus to the use of individual student growth as the measure of student achievement over student proficiency levels alone. They also acknowledge lessons learned about the two approaches to performance-based compensation.

"The best outcome of TIF-LEAP is in the expectation of a gradual change from student proficiency toward more attention to student growth, especially in instances of closing the gap."

"Ultimately, student growth is not only aimed at the reduction of the ravages of poverty attributed to performance gaps, but [at] expected measurable gains for all students."

"Value-added threw us a curve and made public the anxiety over pay for performance. There are trust issues about 'how will it play out?'"

"The pilot [has] provided some lessons learned: SLOs are a labor intensive process and the district has not yet achieved a uniform understanding of how to use them effectively; much more communication of the 'why' of compensation reform and performance pay is needed districtwide; one cannot isolate things like SLOs and VAMs when contemplating changes in salary and compensation by simply calling them initiatives. Rather, they need to be integrated into how the district principals and teachers will perform their duties in the future."

These observations from CMS central administrators capture the breadth of thinking from district leaders about the meaning and impact of Student Learning Objectives and the value-added measure in the district. Overall, it mattered to the TIF-LEAP initiative that a host of district leaders remained committed to the initiative and learned that measuring and compensating teacher effectiveness requires thinking about teaching and learning priorities and planning systemically.

Teacher Association Leadership. Teacher association leaders focus on a range of issues related to performance-based compensation and teachers during both the design and implementation phases. While mostly positive about the effects of the SLO process, concerns from their constituents about the inclusion of value-added in the initiative, and potentially in the district moving forward, give them pause. Issues include: (1) teachers losing the opportunity to provide input in performance-based compensation system design when the district established the VAM without teacher participation; (2) fear of teachers losing salary so that the VAM could be funded; and (3) teachers losing trust in the district as a result.

Teacher leaders pondered what could be learned from the TIF-LEAP initiative by watching changes in the direction of the district toward value-added, and considering what lay ahead for CMS teachers in performance evaluation and compensation:

"Year two [2009-10] was the best year [of TIF-LEAP initiative] because SLOs were well-written and were research-based; in year three, the entire process was changed by adding the growth measure [VAM], which then devalued the non-tested teachers in the schools."

"We had not learned the lessons from TIF [when implementing VAM]; for example, we 'implemented TIF with teachers and not to teachers.' [VAM] is 'to and not with'."

"I have no problems [with performance pay] as long as no teacher takes a pay cut of any kind. It must be a bonus only."

"[The VAM] is gearing up to be suspect. I support a system of opting in and opting out of the value-added model."

Teacher association leaders want performance-based compensation programs where teachers do not take cuts in pay, where teachers participate as partners in the development and implementation of the programs, and where options are provided for teachers during transitions to new pay structures.

Overall. It is notable that a significant number of leaders from all levels of a very large district remained open and thoughtful about the benefits of a performance-based compensation effort throughout the life of the initiative. They also were willing to work through and learn from the implementation complexities of the TIF-LEAP initiative and maintain the fiduciary commitments of a partner in a federal grant.

Several district leaders were reflective about missed opportunities, lessons learned, and how the concepts and practices of the TIF-LEAP initiative might be sustained and maximized through a new strategic plan even with the more austere fiscal landscape ahead.

Critical Role of District Systems

Observation Four: The quality of the support from many district systems proved critical to the outcomes of the initiative.

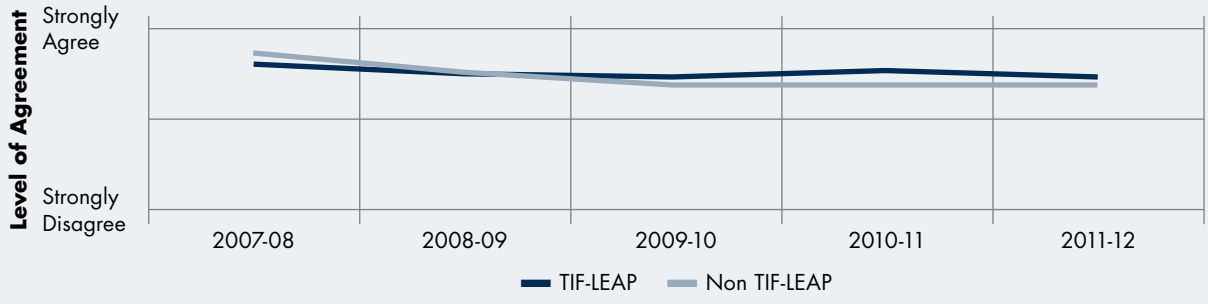
Developing a reliable network of supports from an array of school district systems was critical to successful implementation of the TIF-LEAP initiative. Recognizing that district systems must be aligned and operate fluidly during new program implementation, the SLO Design Team with CTAC technical counsel identified the district supports essential to the success of the initiative. These included (1) integration of SLOs with the core curriculum and instruction components (i.e., alignment, instructional materials); (2) analysis of the availability and suitability of assessments in wide use in the district and the identification of assessment gaps; (3) timely access for the classroom teacher to the most up-to-date student data; (4) high quality principal instructional leadership and supervision; (5) connection of the SLO process to teacher and principal professional development and evaluation; (6) support from information technology services; and (7) transparent interfaces with human resources and payroll.

The evaluation study followed participant perceptions of the presence and effectiveness of these fundamental system supports through the annual educator survey and interviews, as well as through informal observations and artifact analyses. An overview of responses from TIF-LEAP teachers and principals to survey items[5] related to system supports are grouped under the subheadings of Classroom Conditions and Instructional Support, and Professional Growth and Evaluation.

Classroom Conditions and Instructional Support. Seven items related to classroom conditions and instructional support from the educator survey are included in the trend analysis (*Figure VI.2*), such as: "The curriculum is well articulated between and among the grades;" "Students have access to quality learning materials;" and "Student achievement data are used by teachers to plan and adjust instruction."

FIGURE VI.2

Trend Analysis of Teacher and Principal Agreement with Educator Survey Items Related to Classroom Conditions and Instructional Support



The results from the Rasch analysis of survey responses indicate that, over the course of the initiative, the fundamentals required for successful teaching and learning—articulated curriculum, aligned teaching materials and assessments, and the relevant student data for those decisions—were in place and worked to the benefit of the TIF-LEAP teachers and principals.

In comparing TIF-LEAP and non TIF-LEAP school survey responses, these results show that the non TIF-LEAP school-based educators are somewhat more likely than their counterparts at the TIF-LEAP schools, in the baseline year, to rate classroom support highly. However, the TIF-LEAP participants' positive views of classroom conditions and instructional support are more stable through the five years of the initiative.

Professional Growth and Evaluation. The trend analysis of survey responses to statements about teacher professional growth and evaluation (*Figure VI.3*) includes items, such as the following: "Teachers participate in professional development that addresses the needs of students in the school/classrooms;" "The principal provides helpful feedback to teachers on their performance;" and "Teacher evaluations conducted by the principal or others are fair."

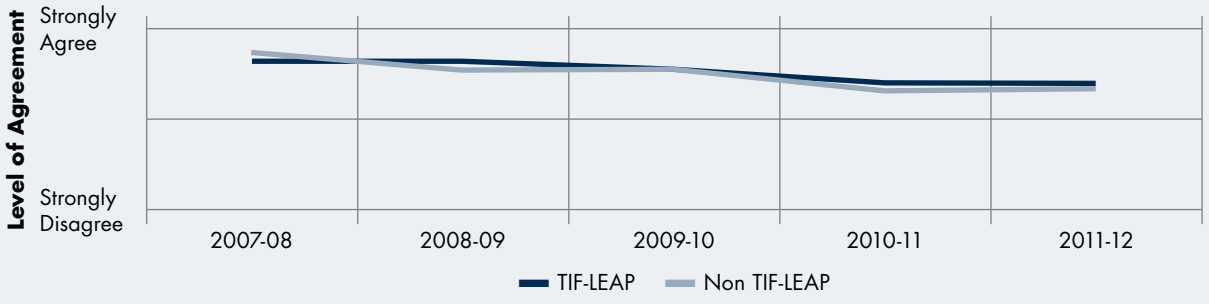
The results from survey responses find most TIF-LEAP principals and teachers strongly agree/agree with statements relating to support and feedback for professional growth and evaluation in the schools and district. However, as shown in *Figure VI.3*, the level of positive responses to these items decline during the final years of the initiative, co-terminus with concern over changes in the initiative and in district conditions, as well as concerns over how the changes impacted teachers. It is worth noting that the decline in positive responses in TIF-LEAP schools is less than in non TIF-LEAP schools.

Performance evaluation is often a point of dissatisfaction for teachers, and expectations for a good experience may be low. During the fourth year of the initiative, amid other district changes, principals and teachers were working through a new state teacher evaluation process that emphasizes the use of evidence, including the collection of artifacts. TIF-LEAP principals rated the outcome of the new process very favorably (more so than the TIF-LEAP teachers who were being evaluated), and many saw helpful overlaps between the new evaluation process and the SLO process.

Two new evaluation issues surfaced during interviews with school level educators and district leaders in the fourth year of the initiative: (1) the complexity of the principal's role in the new teacher evaluation system; and (2) the need for more professional development for principals at a time when there was less funding to support it. While implementation of a new evaluation tool can be expected to create some disequilibrium initially among users, the change came at a difficult time. Layoffs were on the horizon and the new teacher evaluation system required teachers and principals to depart from past evaluation practices and place more emphasis on the use of artifacts and evidence. For teachers in the TIF-LEAP schools, evaluation stakes were higher than for teachers in non TIF-LEAP schools, since the receipt of an earned merit-based

FIGURE VI.3

Trend Analysis of Teacher and Principal Agreement with Educator Survey Items Related to Teacher Professional Growth and Evaluation



bonus for meeting an SLO or reaching the VAM threshold was contingent on receiving a satisfactory performance evaluation.

Survey items related to teacher quality and compensation tracked the impact of the district leadership's interest in and support for teacher performance accountability. Responses show that (1) TIF-LEAP teacher trust in the central administration's recognition of teacher quality and effective administration of compensation plans was tepid at the outset; but that (2) in the latter years of the initiative, a plurality of teachers believed expectations for teacher effectiveness from the district were clear; and that (3) approximately 80% of survey respondents agree that teacher accountability for student success should be a component in evaluation.

Overall. District department leaders provided a range of systemic supports to the TIF-LEAP initiative in the interest of improving teaching and student learning. From the perspective of participants, they mostly succeeded, particularly in the area of classroom and instructional support.

Parents and Community Members Support Linking Student Achievement and Educator Compensation

Observation Five: Parents and community members supported linking student achievement and compensation for teachers and principals.

Parents and community members have a significant stake in the quality of a school district's teachers and principals because of (1) the strength of the connection between outstanding teaching and high levels of student achievement, and (2) the importance of a highly educated workforce and low dropout rate to the financial stability of the community. Parents in particular understand the important contribution of outstanding teachers to excellence in student achievement; and regardless of income, language, or ethnicity, parents seek the best teachers for their youngsters.

How teachers and principals are compensated for their work and how compensation relates to teacher performance, placement, and professional growth are less transparent subjects for parent and community members. Even so, school districts need to know how teacher compensation is understood or viewed in the community. So asking for the community and parent perspective on performance-based compensation and tracking any changes over time is worthwhile information for a district implementing a compensation initiative.

In years one through four[6] of the TIF-LEAP initiative, the evaluation sought the views of parents and community members through annual telephone surveys. The survey is sponsored by CMS annually, and conducted by an external vendor,[7] to garner opinions on a range of issues important to the district. To avoid an additional survey going to parents and community members, items related to teacher and principal performance-based compensation were developed by CTAC and included in the CMS survey. The survey (1) asks question about linking pay to increases in student learning; and subsequently (2) asks for opinions about a range of potential performance bases upon which additional or reward pay might be structured.

TABLE VI.3

Parent and Community Member Responses to Linking Pay to Increases in Student Learning, 2007-2010

	2007	2008	2009	2010	Total
Pay should be linked to increases in student learning	533 (65.2%)	568 (70.6%)	535 (66.5%)	539 (67.0%)	2,175 (67.3%)
Pay should be kept separate from increases in student learning	224 (27.4%)	200 (24.9%)	219 (27.2%)	212 (26.4%)	855 (26.5%)
Don't Know	60 (7.3%)	36 (4.5%)	51 (6.3%)	53 (6.6%)	200 (6.2%)

Pay Linked to Increases in Student Learning

Initial questions ask whether respondents thought that pay should be linked to increases in student learning or kept separate from increases in student learning. Two-thirds of respondents in each of four years indicate that pay should be linked to increases in student learning, as shown in *Table VI.3*.

Performance-Based Compensation for Improving Student Achievement

For the purposes of finding out the areas of teacher and principal performance that parents and community members believe to be feasible bases of bonus compensation, survey respondents are asked to respond to a series of statements using a five-point Likert rating scale. Questions include whether a compensation system should reward teachers for improving student learning in their classrooms and schools, and principals for improvements in the schools *(Table VI.4)*. Responses of strongly agree/agree average nearly 95% for teachers over four years and more than 87% for principals on the respective items.

Performance-Based Compensation for Difficult Assignments and Professional Contributions

An average of 87% of parent and community member respondents strongly agree/agree with rewarding teachers who accept and remain in difficult school assignments and an average of more than 83% agree with rewarding principals who accept difficult assignments.

The level of parent and community member agreement with rewarding teachers for special professional contributions, such as accepting critical, hard-to-staff positions and improving teaching skills through professional development, follow the pattern of the baseline survey items with strongly agree/agree averaging 86 to 90%.

TABLE VI.4

Parent and Community Member Views on the Bases of Teacher and Principal Rewards for Improving Student Learning

Year	Strongly Agree	Agree	Disagree	Strongly Disagree	Don't Know
A CMS compensation system should *reward teachers who improve student learning in their individual classrooms.*					
2007-08	320 (39.2%)	435 (53.2%)	39 (4.8%)	13 (1.6%)	10 (1.2%)
2008-09	358 (44.5%)	393 (48.9%)	38 (4.7%)	6 (0.7%)	9 (1.1%)
2009-10	341 (42.4%)	420 (52.2%)	33 (4.1%)	5 (0.6%)	6 (0.7%)
2010-11	347 (43.2%)	385 (47.9%)	43 (5.3%)	11 (1.4%)	18 (2.2%)
A CMS compensation system should *reward teachers who improve student learning overall in their schools.*					
2007-08	289 (35.4%)	475 (58.1%)	26 (3.2%)	13 (1.6%)	14 (1.7%)
2008-09	350 (43.5%)	407 (50.6%)	35 (4.5%)	3 (0.4%)	8 (1.0%)
2009-10	340 (42.2%)	425 (52.8%)	28 (3.5%)	4 (0.5%)	8 (1.0%)
2010-11	327 (40.7%)	420 (52.2%)	38 (4.7%)	7 (0.9%)	12 (1.5%)
A CMS compensation system should *reward principals who improve student learning in their schools.*					
2007-08	218 (26.7%)	481 (58.9%)	88 (10.8%)	14 (1.7%)	16 (2.0%)
2008-09	264 (32.8%)	448 (55.7%)	71 (8.8%)	11 (1.4%)	10 (1.2%)
2009-10	243 (30.2%)	468 (58.1%)	72 (8.9%)	7 (0.9%)	15 (1.9%)
2010-11	231 (28.7%)	465 (57.8%)	71 (8.8%)	14 (1.7%)	23 (2.9%)

Use of More Than One Measure

A survey item asks if more than one measure of student achievement should be used where teacher and principal compensation is linked. The average percent of agreement is highest with this item as shown in *Table VI.5*.

These survey results are underscored by parents in focus groups who spoke of concerns about the impact of the overuse of and emphasis on one annual high stakes assessment.

The parent and community member survey results provide evidence of the stability in parent views about the use of performance compensation as a means of improving teaching and learning in the district. It shows that they were open to a range of approaches to performance-based compensation for teachers and principals, giving the district and the initiative options with which to work. Even given the highly public fiscal difficulties of the district that characterized the final years of the initiative, parents and community members remain constant in seeing performance-based compensation as deserving of resources.

Summary

There can be little doubt that TIF-LEAP participants, district leadership, parents, and other stakeholders in the CMS community all wanted higher achievement for students and believed that an effective teacher is the portal to better outcomes. The ethos of the district and community consistently support connecting teacher and principal compensation to student academic growth as a measure of teacher performance and accountability. These priorities gave the TIF-LEAP initiative forward movement at the outset.

A second impetus comes from the work of teachers and principals who, as described in their own words, came to understand and use the SLO process to learn more about their students, refine their focus and teaching strategies, set and assess growth targets, and in many cases earn a bonus based on student academic growth. The VAM is more troublesome and confusing from the perspective of most interviewees, but there are those principals and teachers who see the purpose of a valued-added measure and potential of the approach.

TABLE VI.5

Parent and Community Member Views on Using More Than One Measure in Compensation Awards

Year	Strongly Agree	Agree	Disagree	Strongly Disagree	Don't Know
A CMS compensation system should *use more than one measure of student achievement as the basis for awarding extra compensation to teachers.*					
2007-08	Item not asked				
2008-09	311 (38.7%)	415 (51.7%)	6 (0.7%)	49 (6.1%)	22 (2.7%)
2009-10	317 (39.4%)	419 (52.0%)	7 (0.9%)	42 (5.2%)	20 (2.5%)
2010-11	338 (42.0%)	395 (49.1%)	10 (1.2%)	34 (4.2%)	27 (3.4%)

Endnotes

1 In 2007-08, more than 100 individual and focus group interviews were conducted and survey responses from almost 7,000 educators and community members were compiled and analyzed; also, a wide range of artifacts and documents were reviewed. Information drawn from reviews of artifacts and documents (i.e., school and district plans) contributed to interview and survey analyses.

2 The analyses were conducted using a Rasch model of the Likert scale on teacher and principal responses in TIF-LEAP and non TIF-LEAP schools, with the mean of 500 and standard deviation of 100.

3 There are nine items on teacher and principal compensation from the educator survey (see the items in Table VI.1) in the trend analysis.

4 The percentages reflect responses from all TIF-LEAP schools, though the schools were phased in over three years; the number of survey respondents varied over five years.

5 The survey was administered in each of the five years of the initiative between late November and the end of January.

6 Survey questions related to compensation on the CMS parent and community survey were modified in year five.

7 Conducted by Market Wise with about 400 parents and 400 community members each, the survey does not identify which participants, if any, have been interviewed in the past, so responses are treated as cross-sectional rather than time series for this analysis.

CHAPTER VII

National Implications for Performance-Based Systems

Recognizing effective teaching and measuring it are two very different and complex things. Connecting measures of teacher performance to monetary incentives or performance evaluation ratings ramps up the complexity even further. Finally, creating and sustaining the systemic conditions that make great teaching and learning possible is the most exacting challenge of all. The TIF-LEAP initiative in CMS was an ambitious and complex multi-year effort, rooted in a school district's aspirations for higher student achievement in schools most needful of great teachers.

TIF-LEAP was implemented as a systemic initiative. The initiative is a testament to the CMS pursuit of effective teaching in its highest need schools and to its willingness to take on the challenges of performance-based compensation in the interest of reaching that goal. This evaluation examined the genesis, development, and implementation of two different approaches to performance-based compensation, Student Learning Objectives (SLOs) and a value-added measure (VAM). In this context, the evaluation confirms that embarking on a journey of performance-based compensation is about *more than money.*

The evaluation found (1) a positive impact on student achievement at the TIF-LEAP schools relative to the comparison schools, (2) a learning curve for the implementation of SLOs, and (3) a critical role for building supportive conditions in the district.

Among key findings, those related to student achievement are particularly noteworthy. For example, students in the TIF-LEAP schools have a growth rate 12% greater in mathematics and 13% greater in reading than students in the comparison schools. These are both statistically and practically significant. Further, the TIF-LEAP schools show greater resilience to the negative shocks due to the economic recession than the comparison schools.

Teacher effectiveness is a function of management effectiveness.

The relationship between the quality of an SLO and its attainment (meeting or exceeding the stated growth target) is positive and statistically significant, showing that the higher the quality of the SLO, the greater the likelihood it will be attained. In addition, eligible teachers who receive a VAM bonus are more likely to have high quality SLOs. Equally as important, the Curriculum and Instruction Department identified SLOs as an instructional best practice.

The following discussion draws on TIF-LEAP accomplishments and their implications in the larger national school reform context.

National Implications

The national educational landscape is changing dramatically. There is an over-riding focus on creating an effective and credible link between effective teaching, student growth, and compensation.

In particular, the introduction of Student Learning Objectives that began with Denver, and expanded to Charlotte-Mecklenburg, now comprises a major movement extending to districts and states, nationwide. It draws on thirteen years of field-tested practice and research.

The TIF-LEAP initiative in Charlotte-Mecklenburg Schools was part of the second cohort of Teacher Incentive Fund recipients. Since that time, many other districts and states across the country have qualified for funding under this federal program and Race to the Top. They are developing and implementing approaches that link teacher performance and student achievement for the purposes of employee evaluation, professional development, monetary incentives, and other human capital decisions.

Those who are moving down the path of measuring teacher performance will find that the TIF-LEAP initiative advances the formal knowledge about how SLOs impact teacher thinking and get results. Findings also highlight key considerations when introducing and implementing a value-added measure.

The TIF-LEAP experience contributes to a broader, research-based, and practical understanding of what is required to effectively implement performance-based systems for purposes of compensation and evaluation. Specifically, the national implications drawn from the TIF-LEAP initiative are highlighted below.

An effective performance-based system requires a dual emphasis on support and accountability.

Support and accountability are the twin pillars of performance-based reforms. Initiatives to increase student learning by improving teacher performance require significantly more systemic thinking and resources than the usual piecemeal programmatic efforts. They also require leadership with sustained focus.

Findings from the evaluation demonstrate that it is not only the lone star teacher rising to teaching excellence but also the well-oiled machinery of district systems that creates a foundation and the building blocks for successful teaching and learning. Strong systemic support is a critical anchor of an effective performance-based system.

This is an essential point: teacher effectiveness is a function of management effectiveness. Strengthening teacher effectiveness, and assessing it for evaluation and compensation purposes, is serious work. It only thrives through deliberative and cooperative capacity building. This work can be truly effective when there is an oversight structure—with all key decision makers at the table—that is empowered to cut through issues of turf and jurisdiction to ensure that implementation efforts are supported with accountability by relevant district departments and participant groups.

Simply put, the goal of performance-based systems is to help more teachers be more effective with more students. These are, at root, instructional reforms. Therefore, districts should avoid locating the performance-based system in a human resources or accountability department. Instead, they should be located in and owned by the curriculum and instructional arm of the district. Doing so helps ensure that supporting and recognizing effective instruction is the primary purpose of the reform.

The goal of performance-based systems is to help more teachers be more effective with more students.

In practical terms, collaboration matters. As indicated in the Six Cornerstones of Performance-Based Compensation, this is a reform "best done with teachers and not to them."[1] Working with teachers is not just "getting some input." It also means bringing participants into critical decisions like the selection or design of the approach (measure and/or intervention), planning the implementation, and identifying necessary system supports so that the approach is understood inside and out by all who are involved. Teacher/management collaboration is also a means to ensuring that the approach is vetted for equity and opportunity and that it is conducted in a manner that contributes to teacher professionalism.

SLOs provide a measure of student growth and a measure of teacher practice—and quality matters with both.

SLOs serve as a measure of student growth in both tested and non-tested grades and subject areas. They also use a range of different assessments. The concomitant results in advancing student achievement through the use of SLOs are statistically and practically significant as evidenced in both Denver and Charlotte-Mecklenburg.

Equally as important, SLOs also provide teachers and administrators with a research-based methodology to help strengthen instructional practice. SLOs provide a vehicle to ensure that high quality and rigorous student assessments are matched by high quality and rigorous instructional practices.

SLOs provide a measure of teacher practice. When effectively implemented, research results to date indicate that SLOs develop teacher accountability for providing classroom instruction grounded in the use of data, research, and content knowledge as well as in evidence of student learning and growth. In developing and implementing SLOs, teachers demonstrate their level of ability to:

- Analyze and use student data to focus instruction through a deeper understanding of the academic needs of all their students.
- Align classroom instruction with state standards, goals and improvement plans.
- Employ research-based practices within their instructional program.
- Demonstrate knowledge of their discipline and how to use that knowledge to create effective lesson plans with meaningful content and appropriate instructional strategies.
- Use assessments to ascertain the degree of student learning and adjust instructional strategies for students in need of additional support.
- Establish and meet challenging student growth expectations for all of their students.
- Reflect on their practice by understanding where they have been successful, where changes will need to be made to improve student learning, and what type of additional professional development will assist them in their practice.

Measuring teacher performance rigorously through the use of student results is not, in and of itself, an improvement strategy. This point mirrors the dilemma that districts face when trying to improve student performance by testing students more frequently or to increase rigor by raising the thresholds for mastery. As an improvement strategy, SLOs provide a more rigorous, data- based planning process and foster teacher thinking about and efforts to increase student academic growth. They also can be integrated with other approaches, such

as teacher observations, as part of comprehensive performance-based systems.

When used for purposes of measuring student growth or teacher practice, the quality of SLOs matters. In Denver, students whose teachers had high quality objectives outperformed their peers at elementary, middle and high schools levels over a multi-year period as demonstrated by two independent measures.[2] In Charlotte-Mecklenburg, results are similarly positive.

SLOs provide a vehicle to ensure that high quality and rigorous student assessments are matched by high quality and rigorous instructional practices.

The whole process counts when implementing Student Learning Objectives.

The SLO framework is composed of a set of key components. Research findings indicate that, in following the framework to develop SLOs, teachers are thinking differently.

The SLO process fosters planning for instruction with more science through the use of data and systematic analysis. Further, it motivates performance through the setting of growth targets to be reached by both teachers and students. The process of thinking this through is what matters most. Therefore, it is critical to develop this thinking process rather than distribute "model" SLOs or use boilerplate, one-size-fits-all samples.

Teachers and principals identify the interrelated features of the SLO process as being critically important to helping teachers improve student learning:

- The thorough baseline analysis of student data, using both archival and pre-assessment data.
- The focus on key learning content and instructional strategies as detailed below.
- The use of an individual growth measure—the difference between where a student starts and ends—as a more informative and exacting measure for teacher use than the typical proficiency scale.
- The collaboration among teachers and principals on data analysis and instructional planning.
- The alignment of SLO components with an evidence-based teacher appraisal system.

The integration of these features is a core requirement of the SLO process.

Learning Content and Instructional Strategies are key to effective SLOs.

Teachers, principals and central administrators emphasize that identifying student needs, determining the learning content critical for students to master, and then planning and delivering instructional strategies go hand-in-hand. More is involved than just determining whether an SLO growth target is met or unmet.

In tandem with baseline data analysis, many teachers speak of how best to approach student needs through their teaching strategies. Practical knowledge and research indicate that teachers with in-depth knowledge of a range of instructional strategies for addressing the subject matter content and different needs see more student success. This is not to suggest variety for variety's sake. Rather, it is to stress that there are strategies which are research-based and more pedagogically sound for teaching certain content and meeting specific learning needs. Teachers and principals indicate that strategic thinking about instructional strategy is prompted in the SLO process. Absent an emphasis on instructional best practices, performance-based systems run the risk of simply labeling educators with a number while not improving practice—the "gotcha" so rejected by teachers and parents.

In Denver, the emphasis on learning content and instructional strategies catalyzed marked improvements in student academic growth. Charlotte-Mecklenburg built TIF-LEAP on a similar foundation and also had positive results.

Effective SLO implementation requires distinguishing between training, professional development, and leadership development.

Successful SLO implementation depends, in significant measure, on understanding the differences between these three types of support and the need for their implementation to be mutually reinforcing. *Training* refers to helping teachers and principals to understand the components of the SLO framework, the rating rubric, and the relationship of the SLOs to the new evaluation and/or compensation system. *Professional development* focuses on the capacities teachers need to plan and deliver instruction differently and more effectively. *Leadership development* refers to the support needed by principals, assistant principals and site-based instructional leaders to advance the work of all the educators in the school building.

When districts and states under conceptualize the role of these three supports and simply merge them together under a professional development heading, they characteristically fall short of providing the assistance needed by frontline practitioners to implement the reform effectively. By contrast, the TIF-LEAP team in Charlotte-Mecklenburg developed a quick response capability to address anticipated and emerging needs from the schools in all three areas.

The four pivotal considerations when introducing a value-added measure are role, understanding, fairness, and application.

With respect to *role*, there needs to be clarity on whether the VAM is the only measure of student academic growth or one of several measures. To build *understanding,* there needs to be a straight-talk explanation of what the VAM is. The straight-talk should be a true dialogue between district leadership and teachers. The need for both the reality and perception of *fairness* means that growth has to be measured with teachers in both tested and non-tested grades and subject areas. In terms of *application,* a district needs to show how the VAM is going to be used to help improve instruction.

A high level of teacher and principal accountability benefits both schools and districts when it is the impetus for ongoing problem solving and continuous improvement in practice. This may explain why the value-added measure, as implemented in the TIF-LEAP schools, failed to connect with many teachers—the implementation failed to map out the route from a VAM percentile rank to improving one's planning and delivery of instruction.

These four pivotal considerations undergird a VAM approach, particularly if it is introduced by educators who are perceived as credible by teachers and site administrators. Abstract and complicated to communicate, misunderstood VAMs easily become a distraction for the initiative participants, particularly in tandem with other changes, such as layoffs and school closures. Thoughtful and precise implementation is always important, but program fidelity, together with open communication with participants, is critical in performance-based initiatives because the stakes are so high for teachers and principals.

Performance-based systems must meet three standards of validity—statistical, educational, and political.

Addressing these three kinds of validity is important for purposes of program measurement and constituency building.

First, there is *statistical validity.* Whatever measures are reported or actions taken should be the result of assessments that are measured using statistically valid methods. There are various approaches to addressing this problem, including using multiple measures of achievement and/or multiple years of a teacher's results. While these methods add some complexities to the process, they can be used to increase the statistical validity of an assessment, making it both fairer and more useful.

Second, there is *educational validity.* It is possible for statistical results to support practices that are not educationally valid. It also is possible for educationally sound practices to be difficult to measure or prove statistically. Any initiative put into place must also satisfy what is known about

how students learn: it must have educational validity. Further, if the performance-based system does not make sense to teachers and principals in educational terms, it will not be implemented with fidelity or buy-in.

Third, there is *political validity.* This becomes extremely important if comparing scores on standardized tests is one of the methods being used to gauge teacher effectiveness. Even where results are significant statistically, they may not be perceived as legitimate. If teachers perceive that measures being used to partially determine their compensation or evaluation are not legitimate, no amount of statistical validation will be of value. Therefore, political validity—the perception that the system is fair—is critically important at every step of implementation of performance-based systems.

Summary

Increasing student achievement means identifying and fostering an outstanding teacher for every classroom. The Board of Education of Denver Public Schools and the Denver Classroom Teachers Association, during contract negotiations in 1999, took up the challenge of connecting compensation to effective teaching. Starting with a pilot of SLOs, they reformed the teacher compensation, evaluation, and professional development system and won a tax measure to fund it. More than a decade later, they continue to monitor and improve it.

Charlotte-Mecklenburg Schools is an accomplished district. While implementing the TIF-LEAP initiative, it was in the midst of qualifying for and then winning the Broad Prize in 2011. In the design, structure, and staffing of the TIF-LEAP initiative, the strengths of the district were evident. In its laboratory of performance-based approaches, the district experienced what worked and what did not. In navigating through a fiscal crisis, the district managed to keep the initiative going.

This initiative benefitted students, teachers and administrators—thereby demonstrating that, when it comes to effective performance-based systems, more is involved than money alone.

Endnotes

[1] Slotnik, W. (2009, July 15). Get performance pay right: Six cornerstones of successful compensation reform. *Education Week*, *28*(36), 26-32.

[2] Slotnik, W., Smith, M. *et al.* (2004). *Catalyst for change: Pay for performance in Denver.* Boston, MA: Community Training and Assistance Center.

ctac | COMMUNITY TRAINING
AND ASSISTANCE CENTER

30 WINTER STREET • BOSTON, MA 02108
TEL: 617.423.1444 • E-MAIL: ctac@ctacusa.com
www.ctacusa.com