2-16-2015

Illinois Teacher Evaluation Student Growth Component: A Predecision-Making Policy Analysis

Joshua Robert Olsen
Illinois State University, josh.olsen@olympia.org

ILLINOIS TEACHER EVALUATION STUDENT GROWTH COMPONENT:

A PREDECISION-MAKING POLICY ANALYSIS

Joshua R. Olsen

124 Pages May 2015

Public school districts in Illinois are going to face massive changes to their evaluation systems in the coming years. Thanks to the adoption of the Performance Evaluation Reform Act of 2010 and Senate Bill 7 of 2011, teacher evaluation must contain student growth as a significant factor, beginning with the 2016-2017 school year. Not only must student performance on tests become part of the evaluation of teachers, but, moreover, teacher evaluation will play a greater role in the hiring, retention, and release of teachers.

This study was organized as a predecision-making policy analysis. Its purpose was to provide a framework for evaluating teachers to maximize teacher performance and student achievement. This was accomplished by reviewing all of the major legal and legislative enactments dating back to the release of *A Nation at Risk*. The study also reviewed the relevant research on best practice in teacher evaluation to outline the possible elements that can be used to evaluate teachers.

From these contexts, a process for building a local evaluation tool that meets the legal requirements while maximizing teacher performance was outlined. The teacher evaluation system in a school district affects many more areas within the organization

than the personnel file of the teacher being evaluated. It has a broad-reaching effect on hiring practices, culture and climate, budgetary considerations, and many other human resource practices. These implications were addressed directly to provide the local school board member, administrator, and teacher a thorough understanding of what must happen to successfully navigate these critical changes in school law.

ILLINOIS TEACHER EVALUATION STUDENT GROWTH COMPONENT:

A PREDECISION-MAKING POLICY ANALYSIS

JOSHUA R. OLSEN

A Dissertation Submitted in Partial
Fulfillment of the Requirements
for the Degree of

DOCTOR OF EDUCATION

Department of Educational Administration and Foundations

ILLINOIS STATE UNIVERSITY

2015

ILLINOIS TEACHER EVALUATION STUDENT GROWTH COMPONENT:

A PREDECISION-MAKING POLICY ANALYSIS

JOSHUA R. OLSEN

COMMMITTEE MEMBERS:

Elizabeth Lugg, Chair

Neil Sappington

Guy Banicki

Lydia Kyei-Blankson

ACKNOWLEDGMENTS

I owe the highest of thanks and praise to God for equipping me to navigate this incredible journey toward a doctorate. His gifts have come in many forms, the greatest of which is my wife, Cora. Without her love, support, and encouragement throughout our marriage, I would never have been able to dedicate the time, energy, or effort necessary to complete this program. I love you, Cora!

My children, Nathan, Alex, and Kayleigh are the world to me. I don't know what I did to deserve any of you, but God has made me a rich man three times through the blessing of each of you. I love watching you grow, play, and learn each day, and I thank you for allowing Dad the time to get this done. I know you may not have chosen it, and you may not understand it, but someday you will. And I pray that you will see it as something within your grasp, should you choose this path.

Growing up as a farm kid, I learned hard work from my parents, Bob and Paula. They taught me the value of a hard day's work, and to always strive to achieve your greatest potential in whatever endeavor you choose. I have tried sincerely to do that throughout my life, and while I have fallen short on many occasions, I can only hope that I have given them a sense of pride in my efforts and accomplishments.

My in-laws, Paul and Mary, have given me great trust and support for 14 years. You have treated me as though I was your own son, and helped our young family through all of the trials of starting out on our own. Thank you so much for your love and support.

To Dr. Elizabeth Lugg, my committee chair for this study, I owe a tremendous debt of gratitude. I am certain that without your persistence and encouragement, I would not have begun this journey, let alone completed it. You have been a tremendous resource to my educational and professional career since my first day in your Principal's Law class. You have made every learning experience authentic, seeking only to prepare me and my colleagues for the challenges that we would face as school administrators. Thank you, Dr. Lugg, for your faith and support over these many years.

I owe tremendous thanks to the members of my dissertation committee: Dr. Guy Banicki, Dr. Lydia Kyei-Blankson, and Dr. Neil Sappington. I have been fortunate to learn from the best at Illinois State. Each of you has brought a thought-provoking learning experience that has allowed me to maximize my potential. I appreciate how each of you has challenged me to reach farther in my thinking, learning and understanding of education, and what we do for children. You are remarkable professors, and I appreciate your special efforts on my behalf.

Finally, I would like to thank each staff member that I have had the privilege of teaching, coaching, or working with during my time in education. Whether we were next to each other on the game field, in the classroom, or around a conference table, I have learned from all of you. Every single one of you has shared something with me that has shaped me into the educator and person that I am today. I appreciate all of it, and I greatly appreciate the time we spent together doing everything we could for the betterment of our students.

J. R. O.

CONTENTS

FIGURES

CHAPTER I

THE BACKGROUND AND DESIGN OF THE STUDY

Introduction

Parents are frequently asked about the effectiveness of their local school as well as public schools across America. Since 1992, the number of Americans that would rate their public school with a letter grade of "A" or "B" has steadily increased and stands at 77% as of 2012 (Phi Delta Kappan, 2012). However, when those same parents have been asked to grade public schools across the nation as whole, the trend has remained nearly identical over the last decade. Only 19% would give the nation's schools a grade of "A" or "B" (Bushaw & Lopez, 2012).

Since the release of *A Nation at Risk* (NCEE, 1983), the United States Government, and many state governments, have been determined to reform the education system. Reports such as this have been used to blame the educational system for students' lack of top-tier performance on standardized tests of achievement. School districts continue to produce top-quality students who excel at the top universities in the country, all the while being told that, overall, not enough of their students are demonstrating success on standardized tests.

A proposal to address this deficiency was launched in 2001 under Secretary of Education, Roderick Paige, and President George W. Bush. The *No Child Left Behind Act* (NCLB, 2002) was designed to hold school districts accountable for student

achievement of all students, specifically those of specific groups of students, such as minorities, low income, and special education. Since the adoption of NCLB, the state and federal government have been working to hold schools, and ultimately teachers, accountable for their students' progress as measured by standardized tests of achievement.

President Obama continued this mission with the implementation of "Race to the Top" (RT3) as part of the *American Recovery and Reinvestment Act of 2009* (H.R. 1, 2009). With the passage of this bill, $4.35 billion in incentive money became available to State Education Agencies (SEA) and local education agencies (LEA) to adopt drastic changes to the way schools operate. Almost immediately, states began scrambling to get their systems in place to access the federal stimulus money.

The Illinois laws on teacher evaluation, tenure, etc., in place at the time of passage of RT3, were rather contradictory to the RT3 proposed structures. For example, the application for RT3 funding included a rubric with a score for a number of factors determined to be critical in the success of a school system. The single highest point value on the rubric was for teacher and principal accountability measures (USDE, 2009, p. 3).

On January 15, 2010, Governor Pat Quinn signed into law the Performance Evaluation Reform Act (PERA). Reform measures immediately began sweeping across the state under the new law. Professional development, licensure, preparation, pension, and, most notably, the evaluation of educators were all changed in this legislative action. The next wave of reform movement in Illinois was now ushered in.

The most notable evaluation activity now required under PERA was to include student growth in the assessment of teacher evaluation. For decades, legislators and other

reform activists had worked to hold teachers accountable for their students' performance on standardized tests. For the first time in Illinois, this was now required by law. While there is some opportunity to ramp up to the final requirement, ultimately 30% of every teacher evaluation must be comprised of measurements of student performance on assessments. The new law provides specific requirements for mandated collective bargaining of the new evaluation tool.

Included in the law is the requirement to include standardized assessment data (Type 1), when available. Unfortunately, standardized test data is not available for all subject areas and grade levels. Nor are there sufficient curriculum-based assessments that are scored outside the district (Type 2), nor scored inside the district, but administered to all students in the subject area/grade level population (Type 3).

Another issue created by the student growth requirement is the relinquishment of control of the learning environment. Teachers' evaluations, and ultimately their jobs, are now dependent upon, to a degree, the level of success of their students on assessments. The utilization of student teachers, volunteers, special education teachers and aides, or any other person who may be part of the learning environment, represents a loss of control by the regular classroom teacher over her students' performance.

Statement of Purpose

The purpose of this study was to provide a framework for meeting the requirements of PERA and Senate Bill 7 in local school districts. This analysis will serve as a primer for local Boards of Education, administration, and teachers as they navigate the process of changing the evaluation system in their districts to provide for maximum professional growth of teachers, while meeting the requirements of the law.

Delimitations

The historical legal analysis of this study dates back to 1985 to capture all legal and political developments since *A Nation at Risk*. Certainly, there are policies and procedures for teacher evaluation that pre-date this era. For the purpose of this study and the relevance of adding student growth to teacher evaluation in Illinois, the legal and political issues of significance began with *A Nation at Risk*.

This dissertation was formed as a predecision-making policy analysis due to the timing of its production in relation to the implementation of the law. For most Illinois School Districts, student growth is not a required component of teacher evaluation until the 2016-17 school year. It is important for school districts to begin looking at its evaluation policies and procedures and developing the new system in advance of that deadline.

Research Questions

1. Based on the legal requirements of the Performance Evaluation Reform Act (PERA), and the research on best practices in teacher evaluation, what processes, procedures and considerations should a school district navigate when developing their teacher evaluation system?

2. How will the new requirements for teacher evaluation effect the operation of school districts as it relates to finances, human resources, culture and climate, and collective bargaining?

Significance of the Study

Every school district in the State of Illinois must undergo a process of re-design of their teacher evaluation program. PERA requires a plethora of items that must be contained in the teacher evaluation system, least of which is the student growth component. However, simply incorporating all of the requirements in a manner that

meets the legal obligation is only one part of developing a teacher evaluation system. Much more importantly, the teacher evaluation system must provide an effective framework for promoting growth of the teachers evaluated. School districts and teachers alike should endeavor to develop this system and must be well versed in the legal requirements and best practices in teacher evaluation.

Definitions and Technical Terminology

The following is a compilation of definitions of technical terms that were utilized in this study:

A Nation at Risk—a document released in 1985 by President Reagan's Committee for Excellence in Education that is credited for bolstering the educational reform movement (Guthrie & Springer, 2004).

Omnibus School Reform Bill of 1985—effective August 1, 1985, called for 169 specific initiatives that ranged from the Illinois Math and Science Academy to new dismissal procedures for teachers with "unsatisfactory" performance.

Performance Evaluation Reform Act of 2010—effective January 15, 2010, required the use of student growth data to be used as significant factor in the evaluation of teachers.

Race to the Top Educational Reform Initiative of 2009—effective February 17, 2009, as a portion of the American Recovery and Reinvestment Act, was designed to encourage and reward States that were creating conditions for education innovation and reform; achieving significant improvement in student outcomes, including making substantial gains in student achievement, closing achievement gaps, improving high school graduation rates, and ensuring student preparation in college and careers. The

American Recovery and Reinvestment Act provided $4.35 billion for the Race to the Top Fund (U.S. Department of Education, 2009).

Technical Change—change processes that involve a prescribed method that is somewhat mechanical in nature in which issues and solutions are arrived at and repaired. For the purposes of this study, Technical Change is also referred to as First Order Change (Heifetz, 1994).

Adaptive Change—change processes are definable, but no clear-cut solution is readily available for incorporation. Teaming efforts and innovative thinking become essential tasks for addressing issues related to this type of change. For the purposes of this study, Adaptive Change is also referred to as Second Order Change (Heifetz, 1994).

Organization of the Study

Chapter I provides the introductory statements, including identification of the problem, research questions, methodology and significance of the study. Also included in Chapter I are the definitions of technical terminology and delimitations of the study. Chapter II provides an overview of policy analysis as a methodology for research studies. It explains the different type of policy analyses and the methodology chosen for this study, with an explanation of why this is the methodology of choice. Chapter III provides the historical account of teacher evaluation in Illinois and the United States, dating back to 1985. The best practices in teacher evaluation that promote teacher growth and development are outlined in Chapter IV. Chapter V provides a framework for developing a teacher evaluation system that meets the requirements of the Performance Evaluation Reform Act (PERA) and best practices in teacher evaluation.

Chapter VI focuses on implications for school districts as they seek to implement the new system as required under PERA.

Assumptions of the Study

1. School districts will be required and held to the standard presented in PERA to use student growth data as a least 30% of the total teacher evaluation rating beginning in the 2016-17 school year.

2. School districts will endeavor to build evaluation systems that not only meet the minimum requirements of the law, but that also have a positive effect on professional growth of educators.

CHAPTER II

METHODOLOGY

This dissertation was organized as a predecision-making policy analysis. The goal of this type of analysis is, "a means of synthesizing information to draw from it policy alternatives and preferences stated in comparable, predicted, quantitative and qualitative terms as a basis or guide for policy decisions" (Dunn, 1981, p. 51). Policy analysis can be used in a prospective manner (prior to adoption of the policy), in a retrospective manner (after the policy has been adopted), or in an integrated format, where the policy is studied prior to adoption and then monitored post-adoption to determine its effect. With the student growth requirement of the law set to take effect during the 2016-17 school year, it was appropriate for this study to be prospective. However, it would be beneficial for any reader employing its use to continue to study effectiveness of the end product over several years beyond the initial adoption.

According to Dunn (1981), four specific elements of policy problems lend themselves to policy study: interdependence, subjectivity, artificiality, and dynamics. Interdependence suggests that policy problems in one area are inherently linked to policy problems in other areas. In this study, the primary policy problem examined was the addition of student growth to an effective evaluation system. However, that problem was linked with problems in human resources, climate and culture, and other aspects of the school district.

There is also a great deal of subjectivity to the problem being studied. While student growth data can be objectively scored, the interpretation of such data is highly subjective and can be interpreted in very different manners.

The artificiality, as Dunn refers to it, comes from the fact that lawmakers see a need to make this change to the way teachers are evaluated. There is no inherent damage or impending disaster that will strike if student growth is not incorporated into the teacher evaluation system. However, the "powers that be" see this as a means for improving or "fixing" public schools.

Finally, this problem is highly dynamic in the way it can be solved. There is no one perfect template for teacher evaluation. A great deal of history and research should be examined, along with the conditions in the local school district, to develop a tool that will ultimately be legal and encourage student growth. For a local district to implement the requirements of PERA in the proper way, a significant amount of time will need to be invested by a number of stakeholders. Given the nature of the task, it is predictable, that many districts will look for solutions that guide them to the necessary outcome, rather than filtering through the research at nauseam. While this analysis does not seek to provide a finished evaluation tool or product, it does propose to provide the reader with a guide or "how to" on achieving the desired outcome: a tool that meets the legal standard and provides a strong opportunity for teacher growth.

According to Bardach (2012), there are six steps in conducting policy analysis research: define the problem, assemble the evidence, construct the alternatives/ synthesize the information, project the outcome, and communicate the results. This dissertation follows the Bardach formula as a means for providing guidance on the policy

being addressed.

Define the Problem

The problem examined in this study is a problem for every school district in Illinois: student growth and performance must be included in the teacher evaluation instrument and account for at least 30% of the summative rating. While research examined in Chapter IV suggests a correlation between teacher effectiveness and student achievement, this is a new requirement and a significant departure from the traditional evaluation methods. This will, undoubtedly, generate a great deal of angst among teachers who will be reliant on their students for part of their evaluation results. To add even further to the pressure, under PERA and Senate Bill 7, the evaluation results have an even bigger impact on job retention decisions than ever before.

Another aspect of the problem of incorporating student growth into the evaluation tool is that it will eventually account for 30% of the summative rating. So, what about the other 70%? Is there a meaningful tool in place that is promoting teacher development and, ultimately, student growth? This study endeavors to provide a guide for effective teacher evaluation. It offers a framework for the complete process that should be undertaken by Illinois School Districts to evaluate, adopt, and articulate an effective teacher evaluation program.

Assemble the Evidence

The evidence for this policy analysis included the history of legal and political movements dating back to *A Nation at Risk* in 1983. This account, which is shared in Chapter III, uncovers the major developments that have driven the reform movements, and ultimately have resulted in the requirement to add student growth to the teacher

evaluation process. The evidence also included the research on best practices in teacher evaluation. Chapter IV looks at what processes and procedures are proven to be most effective and reliable when evaluating teachers. The key search terms used for research include: teacher evaluation, teacher appraisal, teacher observation, teacher growth, student growth, value-added model, Nation at Risk, No Child Left Behind, Race to the Top, Performance Evaluation Reform Act and Senate Bill 7.

Construct the Alternatives/Synthesize the Information

Under PERA/SB 7, the only change that school districts will be required to adopt is the addition of student growth and performance in the teacher evaluation plan. However, it is reasonable to expect that in any revision of the evaluation plan, a detailed and thoughtful analysis of its components is undertaken by stakeholders.

This study examined the processes that should be engaged in and which stakeholders should be a part of those processes. With the ultimate goal of the evaluation system being to improve teacher performance, which ultimately leads to increased student performance, it is essential that districts have an intentional analysis of all parts of the tool to ensure a high level of effectiveness.

The teacher evaluation laws allow for many options for districts to meet their basic objectives. The required components are outlined to ensure that legal compliance is achieved in the production of the evaluation tool. Best practices, however, provide a much more specific narrative of the components that research shows as having a strong correlation with identifying high performing teachers. This study examined those practices and provides and outline for the creation of a plan that will allow for incorporation of each aspect.

Bardach (2012) suggests that analytical, evaluative, and practical criteria be incorporated in the analysis of policy alternatives to determine their effectiveness and ultimate incorporation in the policy. This study followed this prescribed formula by using statistical evidence to identify the best practices that were identified as being a valuable component to the teacher evaluation plan. Evaluative judgments were used to determine the efficiency, equity, and objectivity aspects of the policy alternatives. Practical judgments were made to ensure that the suggested policy alternatives met the legal litmus test.

Project the Outcome

Everything that happens in schools should be geared toward improving student achievement. This study will provide for policy formulation that is no exception. The goal of the evaluation plan is to improve teaching performance that results in increased student achievement, year after year of implementation.

Horn and Sanders (1997) found, "The most important factor affecting student learning is the teacher" (p. 63). With this evidence, it is expected that the evaluation plan that results from a local school district's implementation of these policy suggestions could increase student achievement by 10-20% annually, based on continuous growth and improvement in teacher performance.

Communicate the Results

Chapter V provides a detailed report of the findings. This includes the legal aspects required in the evaluation plan creation, as well as the plan itself. It also includes the tools and processes that should become a part of an effective evaluation plan, based upon research of best practices.

Unfortunately, all teachers will not improve their craft each year. All students will not achieve expected and predicted levels of growth. It will be critical to develop a system that differentiates the teachers who are not causing increased student achievement from those teachers who are the victim of factors beyond their control. The implications section of this study, Chapter VI, discusses issues that resulted from the evaluation plan. A number of aspects of the evaluation plan have far-reaching, complicated effects on human resources, finance, legal liability, collective bargaining, culture/climate and ultimately, student learning.

CHAPTER III

THE LEGAL AND POLICY BACKGROUND OF TEACHER EVALUATION

A Nation at Risk

> Our Nation is at risk. Our once unchallenged preeminence in commerce, industry, science, and technological innovation is being overtaken by competitors throughout the world. This report is concerned with only one of the many causes and dimensions of the problem, but it is the one that undergirds American prosperity, security, and civility. We report to the American people that while we can take justifiable pride in what our schools and colleges have historically accomplished and contributed to the United States and the well-being of its people, the educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a Nation and a people. What was unimaginable a generation ago has begun to occur—others are matching and surpassing our educational attainments. (National Commission on Excellence in Education, p. 1)

The statement above was written in 1983, when I was a young elementary student. Yet today, as a doctoral student, I am terrifyingly familiar with this opening paragraph to President Reagan's National Commission on Excellence in Education report, *A Nation at Risk.* This guillotine was dropped on the heads of educators across the country some 29 years ago, yet its effects still ripple throughout the educational landscape today.

A Nation at Risk did not cause the first education reform movement in our country's history, but it spawned nearly three decades, to date, of reforms aimed at launching our country to the top of the international directory on standardized tests of achievement. While the intention of many of the crafting members of the committee who

wrote *A Nation at Risk* may not have been to cause a national debate on the effectiveness of education, President Ronald Reagan and his staff certainly set out to do just that. According to Guthrie and Springer (2004),

> When the report was presented to the President, Reagan, who had apparently not read it, called a press conference in which he praised the NCEE for endorsing school vouchers, prayer in public schools, and the abolition of the Department of Education. However, the text of the report had mentioned none of these things. (p. 11)

The President's agenda worked quickly. So quickly, in fact, that one year later the Department of Education Task Force on Education for Economic Growth (1984) noted that 46 states were working on comprehensive state action plans to reform public education.

One of the main targets set forth in *A Nation at Risk* was to improve the performance of teachers. To accomplish this mission, it suggested that, "Salary, promotion, tenure, and retention decisions should be tied to an effective evaluation system that includes peer review, so that superior teachers can be rewarded, average ones encouraged, and poor ones either improved or terminated" (p. 30). Reforms in teacher evaluation systems began to spawn, state by state, across the country, almost immediately. According to Furtwengler (1995), this first wave of the reform era began with 20 states enacting their first requirements for the evaluation of school personnel, mostly between 1983 and 1985, and would be characterized by an increased need for accountability (p. 3). The momentum of state legislation for teacher evaluation continued across the country as 38 states enacted 67 pieces of legislation prescribing specific evaluation procedures by 1992 (Veir & Dagley, 2002, p. 3).

Many teachers and other school personnel were alienated by *A Nation at Risk*. However, the President of the American Federation of Teachers, Albert Shanker, embraced the report, acknowledging the need to improve the quality of American Education (Kahlenberg, 2007).

Illinois Reform Package of 1985

Illinois entered the school reform era officially on July 18, 1985, with the passage of comprehensive education reform bills: Senate Bill 730 and House Bill 1070. These bills outlined 169 specific initiatives that would be the driving force in the Illinois education agenda for years to come (Baker, Ashby, & Rau, 1997, p. 1). According to the Illinois State Board of Education's review, *Legislative Package 1985*, the Illinois General Assembly established a study group, the Illinois Commission on the Improvement of Elementary and Secondary Education, which began its work in 1983 and completed it in 1985 (p. 5). This working group was the driving force behind the comprehensive reform legislation. The Illinois State Board of Education believed that 1985 would be forever known as the "The Year of Education" (p. 5).

A Nation at Risk was not the sole, nor even the beginning, force in Illinois' dedication to school reform. In fact, the Illinois State Board of Education had begun a review of state education mandates as early as 1981. Following *A Nation at Risk*, the national landscape was littered with reports and studies that provided context for a national debate on the effectiveness of public schools. Illinois Speaker of the House, Michael Madigan, still the current Illinois Speaker of the House, added to the furor by holding a series of statewide debates on educational issues.

The Illinois State Board of Education (1985) reported that all of the debates, meetings, and discussions about education over the years leading up to passage of legislation had focused on several key topics that were eventually addressed in the omnibus bill:

- Overall student achievement, production of the school systems;
- What should be contained in the educational program, what outcomes should be expected;
- How to assess and measure student learning and respond to underachieving students;
- The quality and professionalism of educational staff, including the growth and measurement of performance;
- Special needs students, including at-risk, gifted, handicapped, limited-English proficient;
- The organization of Illinois school districts and other regional delivery systems;
- Funding for all programs and services. (p. 7)

These broad issues were addressed in literally hundreds of pieces of legislation that surfaced in the General Assembly before the two landmark bills were signed into law by Governor Thompson. Although the outcome was a tremendous change in the educational system in Illinois, it took a tremendous process in both houses, facilitated by the leadership, to whittle the bills down and cover all of the sacred cows.

The first topic addressed in the reform bill was to put into law the purpose of education. The bill stated:

> The State of Illinois, having the responsibility for defining requirements for elementary and secondary education, establishes the primary purpose of schooling as the transmission of knowledge and culture through which children learn in areas necessary to their continuing development. These areas include: language arts; math; biological sciences; physical and social sciences; the fine arts; and physical development and health. Each school district is required to give priority in the allocation of resources, including funds, time allocation, personnel, and facilities, to fulfilling the primary purposes of schooling. (p. 20)

The bill also granted powers to the Illinois State Board of Education and the local boards of education to outline goals, knowledge, and skills that should be mastered by grade level (p. 22). The bill did not grant these powers and forget the accountability measures to ensure they were enacted properly. The 1985 reform created statewide standardized tests of achievement and the Illinois School Report Card to provide the public with evidence of the effectiveness of their local school district, and the state as a whole, on an annual basis beginning in 1986 (p. 31). The Illinois Goal Assessment Program, or IGAP, was the first step into the icy waters of high-stakes testing for Illinois. According to Vogel et al. (2006), districts were required to develop learning objectives aligned to the 34 goals identified in the 1985 bill (p. 43). This assessment data would be the foundation of the level playing field that districts would report in the State Report Card.

The practice of social promotion was still in place in 1985, until the reform package passed, which disallowed such practice, and mandated interventions for students who were determined to be one or more grade levels behind (ISBE, p. 28). Many curricular changes were addressed in the package, including more rigorous physical education requirements, driver's education fee increases, consumer education proficiency exams, parent education, a more fully developed health education program, and expansion of the American History Curriculum (p. 9). Further landmark changes to Illinois education program included the creation of the Illinois Math and Science Academy, full-day kindergarten, and pre-school programs for 3- and 4-year olds who were identified as "at-risk for academic failure when they enter school" (p. 55). Also included was the creation of "Educational Service Centers" (ESC) to provide for regional

coordination of services and programs in the areas of "gifted, computer technology, math, science, and reading" (p. 134).

The 1985 reform package produced many outcomes that would change how teachers were licensed, hired, evaluated, and retained. Educators would now have to demonstrate proficiency on tests of basic skills, as well as on content-area exams, prior to being certified. Prospective teachers were then given designations on their certificates, based on the program that they had completed and the assessments they had passed, that identified specific areas the certificate holder was qualified to teach. The landmark legislation was the first to require trained evaluators to evaluate the effectiveness of teachers. The Illinois State Board of Education noted,

> School boards are required to establish and implement programs of certified employee evaluation approved by the State Board of Education. It requires the State Board of Education to train school district administrators in evaluation techniques. The Act also eliminates the current requirement of a due process dismissal hearing unless requested by the teacher and reduces various time requirements in the due process hearing procedure. If a school district has not evaluated all of its teachers by the end of the 1987-88 school year, or fails to evaluate such teachers within every two years thereafter, the State Board of Education must evaluate the teachers. The General Assembly urges the Illinois State Board of Education to provide a comprehensive program throughout Illinois to train administrators to evaluate personnel effectively, including documentation, remediate, and implementation of the process. (p. 107)

While evaluation was near the middle of the reform package, item number 58 out of 169, it is clear from the language that it was not an afterthought. The prescriptiveness of the requirements for evaluating local personnel was unprecedented. The plan, as a change in working conditions, would need to be bargained with the local education association, where applicable. Upon approval by both the local agency and district, the plan had to be submitted and approved by the Illinois State Board of Education.

Even more surprising might be the consequence of failing to meet the requirement. Rather than reducing financial support, disciplining the administrator, or in some other way punishing the district, the Illinois State Board of Education was required to complete the evaluation themselves. The legislature made it crystal clear that Illinois teachers were to be evaluated going forward, one way or another. Additionally, the plan had to be bargained with the local association (union) and then submitted to the State Board. This was in direct conflict with the notion that ineffective teachers could be dismissed for poor performance and a root cause for the extremely low number of teachers that have been released in Illinois over the past 4 decades.

Principals and superintendents were not forgotten in the package, as the Illinois Administrator Academy system was born from this legislation as a requirement for continuing recertification. Principals were now defined in school code as "instructional leaders" and districts were required to provide professional development programs for their staff. It is interesting to note that the legislature was progressive enough at this time to call them instructional leaders, yet their statement on the purpose of education was the transmission of knowledge and culture. Current trends in education focus administrators on instructional leadership. In fact, the legislation required that a minimum 51% of their time be focused on instructional leadership. This provided for more teachers focusing on being conductors of student learning activities that promote higher-order thinking, problem-solving, creativity and analysis. While the 1985 process vaulted many programs and services forward, it is clear that continued growth in Illinois education policy was necessary.

As legislative leaders of late have done, the 84th general assembly did not leave school districts wondering how to pay for the comprehensive reform plan. An unprecedented increase in school funding was granted, which amounted to approximately 17.3% over the previous year's funding. Each of the 169 items that required funds was given an allocation and authority for the Illinois State Board of Education to fund in the future. This is where our legislature of late has begun to tear apart the system. Many of the programs and services now required are specifically mandated, but without funding. Program funds have been drastically diminished or completely cut.

The Standards Movement

By the mid-1980s, the National Governor's Association had become a centerpiece for action in American education policy. A summit was held in 1989 where all of the governors were invited by the President for a summit specifically designated for education policy. Shortly after the summit, a task force was created to begin drafting goals. Six completely unattainable goals were released by President Bush in his 1990 State of the Union address:

- By the year 2000, every child must start school ready to learn.
- The United States must increase the high school graduation rate to no less than 90 percent.
- In critical subjects—at the 4th, 8th, and 12th grades—we must assess our students' performance.
- By the year 2000, U.S. students must be first in the world in math and science achievement.
- Every American adult must be a skilled, literate worker and citizen.
- Every school must offer the kind of disciplined environment that makes it possible for our kids to learn. And every school in America must be drug-free. (Bush, 1990)

While many Americans appreciated the President's commitment to education and his national goals, many also saw this as another fruitless endeavor, wrought with empty

promises. Keith Geiger, the President of the National Education Association stated, "The President has provided us with a hearty menu, but has left the cupboard virtually bare (Vinovskis, 2009, p. 28).

The Governor's association accepted the goals offered by the President, although they saw them as a template, rather than a prescription. They added 21 objectives, dispersed throughout the six goals, "focused on enabling or opportunity to learn indicators rather than referred just to student and adult outcomes" (p. 28). However, this rhetoric was not able to make it through congress as a school reform bill until 1994, now under a new President, Bill Clinton. Congress also added two more goals, one for "professional development of teachers" and one for "parental participation" (p. 74). None of these actions dramatically affected teacher evaluation in Illinois or across the country.

The Illinois legislature enacted a minor language change in 1997 that would have a drastic effect on teachers in Illinois. According to the Illinois School Code (2012, Article 24-11), a teacher could gain tenure with a school district after just 2 consecutive years of service. However, effective January 1, 1998, teachers had to have 4 consecutive years of service to become tenured. While this change resulted in a tremendous accountability increase for teachers in Illinois, it did not come without a cost. In order to get this legislation passed with the two giants in lobbying and political contributions, the Illinois Education Association (IEA) and the Illinois Federation of Teachers (IFT), legislators bargained an increase in the retirement calculation to 2.2%. With this sweetener in the pot, school districts could now evaluate the performance of a teacher, and, for up to 3 years, they could dismiss that teacher, without reason, and without fear of retaliation. The Illinois evaluation laws enacted in the 1985 reform package required that

all teachers be evaluated at least every other year; however, most school districts' collective bargaining agreements required non-tenured teachers to be evaluated twice annually. By increasing the non-tenured service period from 2 years to 4 years, the legislature effectively caused a tremendous increase in the substance of the evaluation process for non-tenured teachers.

No Child Left Behind

It was not until the passage of the *No Child Left Behind Act* on January 8, 2002, that the national spotlight would, once again, shine on teacher performance and evaluation. *No Child Left Behind* was the reauthorization of the *Elementary and Secondary Education Act of 1965*. Its passage ushered in another movement of accountability. This time, accountability was extremely prescriptive for school districts, buildings, classrooms, teachers and special student populations. Students in grades 3 through 8, and 11 would be tested annually to determine if they were making Adequate Yearly Progress or AYP (No Child Left Behind, Part A, Subpart I, Sec. 1111, Accountability). Not only were individual student results examined by teachers, administrators and parents, but students from special populations were clustered to identify achievement gaps. For decades, affluent school districts ignored poor and minority students' test scores and other measures of achievement. They paid no mind to the fact that a portion of their student population was failing to meet standards, because the majority, affluent, white students scored well.

According to ISBE's interpretation of the requirements for statistical significance, when at least 40 students from any one of eight sub-groups were tested, their scores were disaggregated to determine if the sub-group met AYP on its own. The subgroups were

identified as: Asian & Pacific Islander, Black, Hispanic, American Indian, white, free and reduced price lunch students (low income), special education students, and Limited-English proficient students. The goal of the sub-group analysis was to increase focus on, and end the growth of, the ever present achievement gap between white and minority students that was brought forth so vigorously in *A Nation at Risk*. This requirement would force affluent schools to educate *all* of the students under their roof, not just the capable ones.

While much attention has been directed toward the sub-groups and their achievement, critics argue that the manner in which *No Child Left Behind* measures achievement, and the sanctions it places on districts who cannot meet its impossible standards, has not helped close the gap. Hoerandner and Lemke (2006) stated, "The punishments for not 'progressing' are quick to be threatened and are overly severe. Standardized tests fail to accurately measure skill levels. It encourages teaching to the test in place of a well-rounded curriculum" (p.1).

Best practice suggests that all students receive a rich curriculum designed to foster higher-level thinking (Sappington, 2008). This type of instruction will lead to positive test results. However, overemphasis on subgroup analysis could undermine these efforts. Nichols and Berliner (2007) illustrate how the significant social consequences associated with a quantitative indicator, such as test scores, can result in corrupting the social processes they were designed to monitor.

No Child Left Behind did not just demand that schools close the achievement gap by requiring testing and breaking apart results in unimaginable ways. It also sought to improve the quality of teachers by legislating new standards for certification,

performance, and evaluation. Smith and Gorand (2007) noted, "For the nation's public school teachers, *No Child Left Behind* means complex systems of performance and accountability measures aimed at addressing concerns over teacher quality and increasing the number of highly qualified teachers in America's schools" (p. 192).

The term "highly qualified" gained a whole new meaning after the passage of *No Child Left Behind*. The act legislated that by 2006, all teachers in the U.S. would meet a strict set of requirements for the subject-area that they were teaching. *No Child Left Behind* required all teachers to have a full certificate, bachelor's degree and to demonstrate competence in every core subject they teach. The core subjects were defined as: English, reading or language arts, mathematics, science, foreign languages, civics and government, economics, arts (music and art), history, and geography (*No Child Left Behind*, 2002). States were left to define exactly how they would assess whether their teachers were meeting the new requirements. They would be forced to change their certification systems to address the new standards for incoming teachers. However, millions of teachers across the country were currently certified and teaching and would now have to become highly qualified in just 4 years.

Teachers were not the only group that was required to become highly qualified. Instructional aides or teaching assistants were also required to become highly qualified by passing 30 hours in education courses from a regionally accredited college or university, passing the ACT WorkKeys, or passing the ETS Parapro assessment (ISBE, 2012).

According to the Illinois State Board of Education website (2012), a complex system of points for activities was developed called HOUSSE, which stood for High Objective Uniform State Standard of Evaluation. This was the format for determining if

veteran teachers, those with one or more years of experience, were now highly qualified. Students in classrooms that were not staffed with highly qualified teachers were required to have a note sent home to their parents explaining what qualification(s) the teacher lacked.

The HOUSSE system was analyzed through worksheets developed by the Illinois Education Association (IEA) and Illinois Federation of Teachers (IFT). These two labor leaders created tools for educators to score themselves to determine if they were meeting the new federal guidelines and, in turn, provide that evidence to their district. Smith and Gorand also noted the goals evident in certification portions of *No Child Left Behind;* "It is quite clear that the emphasis throughout HOUSSE, and indeed all aspects of the highly qualified teacher requirements, is that teachers must prove that they have sufficient content knowledge for the subjects that they teach" (p. 193).

No Child Left Behind did not, in fact, focus on evaluation of instructional staff, but rather on the preparation of teachers. This may have been its biggest flaw. The act did not seek to improve instruction but to improve outcomes which it was not measuring. It effectively caused the complete overhaul of the preparation system for teacher candidates. According to some critics, it weakened the preparation by taking the focus away from instructional skills, and moving it toward content knowledge. As one author suggested, "We should not confuse a highly qualified taker of tests about teaching with a highly qualified classroom teacher" (Berliner, 2005, p. 208).

While much of the law is dedicated to the testing of students and the ramifications for not making AYP, it is hard to argue that the major focus of *No Child Left Behind* was to improve the quality of teachers. The act required a report from all 50 states in 2003,

which was to be a progress report on the process toward full attainment of all requirements for teacher certification by 2006. The report required that states report on:

- state certification and licensing requirements for new teachers, including those that enter through traditional and alternative routes;
- statewide pass rates on recent state assessments of graduates from teacher preparation programs;
- numbers of teachers on waivers or emergency and temporary permits;
- information on teacher standards and their alignment with student standards;
- criteria for identifying low performing schools of education. (*No Child Left Behind*, 2002)

Much of the specificity on how to implement *No Child Left Behind* was left out of the law and given to states. It would stand to reason that states would then implement many laws and policies regarding every facet of teacher performance, including the evaluation of such. According to Hazi and Rucinski (2007), 42 states passed laws affecting how teachers were evaluated (p. 11). Illinois, however, was not one of them; not for several more years.

Race to the Top

On November 4, 2009, at an education rally in Wisconsin, President Barak Obama explained to on-lookers how Washington had, once again, attempted to legislate education reform. "It's time to stop just talking about education reform and start actually doing it. It's time to make education America's national mission" (USDE, 2009, p. 2). He was speaking about the passage of the *American Recovery and Reinvestment Act*, which included $4.35 billion for a program called Race to the Top.

Race to the Top was a competitive grant program; a new style of education reform. Rather than casting reforms down on the states, or even directly to local school districts, this time Congress dangled the largest education carrot they had ever created in

front of states and school districts. The goal of Race to the Top was to allow states that had already developed successful reform strategies, but could not afford to implement them on a large scale, to apply for a grant that would fund the roll-out of the initiative. Not only would this broaden school reform efforts to state-wide scale, but it would drag them across state lines to initiate wide-spread national reform.

According to the *United States Department of Education Race to the Top Executive Summary* (2009), competitive grant monies would be allocated based on a scoring rubric that centered state applications on six key areas: state success factors (125 points), standards and assessments (70 points), data systems to support instruction (47 points), great teachers and leaders (138 points), turning around the lowest-achieving schools (50 points), and a general area (55 points). There is no coincidence that the "great teachers and leaders" section was weighted the highest. As previous reform efforts and research have shown, there is a strong correlation between student achievement and teacher effectiveness. If you want to increase the former, you must increase the latter.

Section D of the rubric is the crux of the reform package focusing on teacher performance (see Exhibit 1).

Exhibit 1—Section D of the USDE Race to the Top Executive Summary

D. Great Teachers and Leaders *(138 points)*
State Reform Conditions Criteria

(D)(1) Providing high-quality pathways for aspiring teachers and principals *(21 points)*
The extent to which the State has—
(i) Legal, statutory, or regulatory provisions that allow alternative routes to certification (as defined in this notice) for teachers and principals, particularly routes that allow for providers in addition to institutions of higher education;
(ii) Alternative routes to certification (as defined in this notice) that are in use; and

(iii) A process for monitoring, evaluating, and identifying areas of teacher and principal shortage and for preparing teachers and principals to fill these areas of shortage.

Reform Plan Criteria

(D)(2) Improving teacher and principal effectiveness based on performance *(58 points)*

The extent to which the State, in collaboration with its participating LEAs (as defined in this notice), has a high-quality plan and ambitious yet achievable annual targets to ensure that participating LEAs (as defined in this notice)—

(i) Establish clear approaches to measuring student growth (as defined in this notice) and measure it for each individual student; *(5 points)*

(ii) Design and implement rigorous, transparent, and fair evaluation systems for teachers and principals that (a) differentiate effectiveness using multiple rating categories that take into account data on student growth (as defined in this notice) as a significant factor, and (b) are designed and developed with teacher and principal involvement; *(15 points)*

(iii) Conduct annual evaluations of teachers and principals that include timely and constructive feedback; as part of such evaluations, provide teachers and principals with data on student growth for their students, classes, and schools; and *(10 points)*

(iv) Use these evaluations, at a minimum, to inform decisions regarding— *(28 points)*

(a) Developing teachers and principals, including by providing relevant coaching, induction support, and/or professional development;

(b) Compensating, promoting, and retaining teachers and principals, including by providing opportunities for highly effective teachers and principals (both as defined in this notice) to obtain additional compensation and be given additional responsibilities;

(c) Whether to grant tenure and/or full certification (where applicable) to teachers and principals using rigorous standards and streamlined, transparent, and fair procedures; and

(d) Removing ineffective tenured and untenured teachers and principals after they have had ample opportunities to improve, and ensuring that such decisions are made using rigorous standards and streamlined, transparent, and fair procedures.

(D)(3) Ensuring equitable distribution of effective teachers and principals *(25 points)*

The extent to which the State, in collaboration with its participating LEAs (as defined in this notice), has a high-quality plan and ambitious yet achievable annual targets to—

(i) Ensure the equitable distribution of teachers and principals by developing a plan, informed by reviews of prior actions and data, to ensure that students in high-poverty and/or high-minority schools (both as defined in this notice) have equitable access to highly effective teachers and principals (both as defined in

Race to the this notice) and are not served by ineffective teachers and principals at higher rates than other students; and *(15 points)*
(ii) Increase the number and percentage of effective teachers (as defined in this notice) teaching hard-to-staff subjects and specialty areas including mathematics, science, and special education; teaching in language instruction educational programs (as defined under Title III of the ESEA); and teaching in other areas as identified by the State or LEA. *(10 points)*
Plans for (i) and (ii) may include, but are not limited to, the implementation of incentives and strategies in such areas as recruitment, compensation, teaching and learning environments, professional development, and human resources practices and processes.

(D)(4) Improving the effectiveness of teacher and principal preparation programs *(14 points)*
The extent to which the State has a high-quality plan and ambitious yet achievable annual targets to—
(i) Link student achievement and student growth (both as defined in this notice) data to the students' teachers and principals, to link this information to the in-State programs where those teachers and principals were prepared for credentialing, and to publicly report the data for each credentialing program in the State; and
(ii) Expand preparation and credentialing options and programs that are successful at producing effective teachers and principals (both as defined in this notice).

(D)(5) Providing effective support to teachers and principals *(20 points)*
The extent to which the State, in collaboration with its participating LEAs (as defined in this notice), has a high-quality plan for its participating LEAs (as defined in this notice) to—
(i) Provide effective, data-informed professional development, coaching, induction, and common planning and collaboration time to teachers and principals that are, where appropriate, ongoing and job-embedded. Such support might focus on, for example, gathering, analyzing, and using data; designing instructional strategies for improvement; differentiating instruction; creating school environments supportive of data-informed decisions; designing instruction to meet the specific needs of high-need students (as defined in this notice); and aligning systems and removing barriers to effective implementation of practices designed to improve student learning outcomes; and (ii) Measure, evaluate, and continuously improve the effectiveness of those supports in order to improve student achievement (as defined in this notice).

When examining Section D, the greatest value is given to part 2, "Improving teacher and principal effectiveness based on performance." This section calls for the applicants to demonstrate they have measures for student growth for each teacher and a

"rigorous, transparent, and fair evaluation system for teachers and principals." Furthermore, the document calls for these evaluations to be the centerpiece of human resource management in school districts. It is to be the driving force behind professional development, salary, promotion, retention, and dismissal. This area is of such importance in the legislation that only one other sub-point has a higher point value in the entire rubric, behind only "support for LEA's in using one of four selected turnaround models for underperforming schools" (section E).

Performance Evaluation Reform Act

States immediately began clamoring to get their piece of the $4 billion pie. With Illinois being the home state of President Obama and Secretary of Education Arne Duncan (former CEO of Chicago Public Schools), it was essential that Illinois prepare an application that would allow it to access the funds. Governor Pat Quinn, State Superintendent Chris Koch, members of the Illinois Statewide School Management Alliance, which included the Illinois Association of School Administrators (IASA), Illinois Association of School Business Officials (IASBO), Illinois Principals Association (IPA), Illinois Association of School Boards (IASB), leadership of the IEA and IFT and members of several reform organizations all began working together to write Illinois' Race to the Top application.

Unfortunately, anyone hoping that Illinois would get a share of the Race to the Top money knew that it had no chance without significant changes to the current system. Illinois is a stronghold among educational unions, with a high percentage of educators being members of one of the three main associations: Illinois Federation of Teachers (IFT), Illinois Education Association (IEA) or the Chicago Federation of Teachers (CFT).

Illinois simply could not document success in many of the key areas outlined in the rubric, most notably in section D. The focus of the working group was on drafting legislation that would strengthen Illinois' bid for Race to the Top money. On January 15, 2010, Governor Quinn signed into legislation the *Performance Evaluation Reform Act of 2010*, which was the stakeholder group's effort to bolster Illinois' standing in the Race to the Top process.

While the *Performance Evaluation Reform Act of 2010* may not have had the breadth that the 169 point reform package of 1985 had, it certainly had all of the depth. The act began with a call to action stating:

1. Effective teachers and school leaders are a critical factor contributing to student achievement.
2. Many existing district performance evaluation systems fail to adequately distinguish between effective and ineffective teachers and principals. A recent study of evaluation systems in 3 of the largest Illinois districts found that out of 41,174 teacher evaluations performed over a 5-year period, 92.6% of teachers were rated "superior" or "excellent", 7% were rated "satisfactory", and only 0.4% were rated "unsatisfactory".
3. Performance evaluation systems must assess professional competencies as well as student growth.
4. School districts and the State must ensure that performance evaluation systems are valid and reliable and contribute to the development of staff and improved student achievement outcomes. (Illinois General Assembly, 2010)

Clearly, the impetus of the act was to reform education, and this time it would be done by changing the way teachers were evaluated. Past reform efforts, both in Illinois and across the country, had touched on evaluation as a component, but never before had a reform package been dedicated solely to fundamentally changing the way educators were evaluated.

To address the issue cited in (2) above, the legislature changed to four distinct categories to allow for greater differentiation of teacher effectiveness. The *Performance*

Evaluation Reform Act of 2010 (Illinois General Assembly, 2010) called for both principals and teachers to be evaluated using the categories of "excellent, proficient, needs improvement or unsatisfactory" (p. 2). To satisfy their requirement in (3) above, the act required that all school districts implement an evaluation system for their certificated employees that, "incorporates the use of data and indicators of student growth as a significant factor in rating teacher performance" (p. 6). This section of the law began to demonstrate a completely renewed direction for education reform. It bears reminding at this point that the teachers' unions and management alliance supported legislators in the endeavor to create this language. These organizations have active memberships, not only contributing millions of dollars into campaign funds for Illinois legislators each year, but also taking to the street to make sure legislators are well-informed of their position on issues.

The *Performance Evaluation Reform Act of 2010* was very specific, not only in the content and nature of the evaluation systems but also in the timeliness of evaluations. School districts no longer had the local control to decide how often different classes of teachers would be evaluated. The *Performance Evaluation Reform Act of 2010* states:

- Each teacher not in contractual continued service [tenure] is evaluated at least once every school year; and
- Each teacher in contractual continued service [tenure] is evaluated at least once in the course of every 2 school years. However, any teacher in contractual continued service whose performance is rated as either "needs improvement" or "unsatisfactory" must be evaluated at least once in the school year following the receipt of such rating. Notwithstanding anything to the contrary in this Section or any other Section of the School Code, a principal shall not be prohibited from evaluating any teachers within a school during his or her first year as principal of such school. (p. 7)

Not only are these requirements a new undertaking for Illinois legislators, they are much stricter than those required by the 1985 legislative package, and more than most Illinois school districts have previously negotiated with their local education associations. This has a left a litany of negotiated agreements between local school boards and local labor unions in direct violation of school code. In most cases, this circumstance nullifies the affected section of the agreement. School districts with current collective bargaining agreements that have conflicting sections with the new law now have no language in effect for the conflicting sections. School districts need only to follow the new school code on evaluation.

The *Performance Evaluation Reform Act of 2010* remained very similar to previous legislation for teachers who received a cumulative rating of "unsatisfactory." Those teachers were now required to participate in a remediation plan and be re-evaluated. Previously, if the succeeding evaluation was "satisfactory", the teacher was reinstated to the regular evaluation plan. The *Performance Evaluation Reform Act of 2010* required that the teacher achieve a "proficient" rating. If "unsatisfactory" or even "needs improvement" is attained on the succeeding evaluation, the teacher must be dismissed (p. 9).

Administrators were not left out of the fray of the *Performance Evaluation Reform Act of 2010*. Principals, and any other certificated employees who evaluate teachers, were also required to have student growth as a significant factor in their evaluation, and they must be evaluated annually. They must also receive one of the four previously named categories as their summative rating (p. 12). The main difference is that, for most school districts in Illinois, outside of Chicago Public Schools and not in the

bottom 20% of schools in the state as measured by standardized tests of achievement, the student growth portion of the evaluation does not have to be in place until the 2016-17 school year. For principals, it must be in place by the 2012-13 school year.

The final cog in the *Performance Evaluation Reform Act of 2010* wheel of evaluation is that evaluators must be trained by a formal process, to be developed by the Illinois State Board of Education, before they can begin evaluating in the 2012-13 school year. The process will be on-line and require approximately 20 hours of training for those who will be evaluating teachers. An additional 20 hours will be required of those who will be evaluating principals.

While the obvious goal of the *Performance Evaluation Reform Act of 2010* was to improve the quality of teaching in Illinois by increasing the rigor of the evaluation process, the underlying goal all along was for Illinois to be able to access its share of the $4 billion is federal money available through Race to the Top. The first round of recipients was announced in August of 2010, and included 10 states. Illinois, however, was left out. The federal reviewers saw that Illinois had implemented changes but not enough to score in the top 10. Illinois would need to do further work to strengthen its position if it were to access funds in later rounds. As Illinois commonly does, it went back to the legislators.

Senate Bill 7

The stakeholder group went back to the table, once again, to improve upon the already gigantic undertaking that resulted in the *Performance Evaluation Reform Act of 2010*. Although none of the previous legislation had taken effect, the group had hoped that providing for it in legislation would be enough to gather the necessary points on the

scoring rubric to be funded. Since this was not the case, the group drafted new legislation to assure that the effects of the *Performance Evaluation Reform Act of 2010* would be applied to the greatest extent possible.

In June of 2011, Governor Quinn signed the next education reform legislation in Illinois. This time it was the *Illinois General Assembly Public Act 097-0008* (2011), which is more commonly referred to as *Senate Bill 7*. This legislation required that evaluations be used in determining who gets released when a district is forced to cut personnel under reduction-in-force (RIF) procedures (pp. 29-38). Previous methods called for the staff members with the least seniority to be released. Now that evaluations were to be more effective, it stood to reason that they should be the decision-making tool in personnel decisions. *Senate Bill 7* also made qualifications and evaluations the deciding factor in filling new positions (p. 13).

Tenure was again changed in *Senate Bill 7*, basing it on performance evaluation outcomes over time, rather than only consecutive years of service. Teachers who achieve all "excellent" evaluations would now be tenured after just 3 years. Teachers who receive "proficient" will have to wait until 4 years have passed to gain tenure, but will not be granted tenure if they receive "needs improvement" or "unsatisfactory" (p. 16).

A new fate was added for teachers who receive multiple "unsatisfactory" ratings in a 7-year period. Those teachers would be eligible for certificate revocation by the State Certification Board (pp. 7-8). This put a great deal of meaning into the evaluation, which, in and of itself, is not grievable in most school districts. If a certificate is revoked because of "unsatisfactory" ratings, the district would not be able to employ the individual and would not have any liability challenges for terminating them.

Once again, with full support of the education stakeholders in Illinois, the legislation reduced the power of unions in the area of strikes. In any district except the City of Chicago, if during the collective bargaining process an impasse is declared, the last, best offer and the costs associated with each offer will be made available to the public by the mediator. In Chicago, impasse would result in 90 days of fact-finding, followed by 30 days of public viewing of the last best offers. If the Chicago Teachers Association (CTA) wishes to strike after this time period, it would require a 75% passage rate in a union vote. While this language does not prevent a strike from happening, it severely damages the unions' ability to use this tactic in the bargaining process.

Other changes to the school code implemented with the passage of *Senate Bill 7* include a streamlined process for the dismissal of tenured teachers (pp. 38-94) and mandatory training for members of local school boards. The dust had yet to settle on the *Performance Evaluation Reform Act of 2010* when *Senate Bill 7* was passed, increasing the rigor and relevance of its mandates.

A second round of Race to the Top fund availability was announced in 2011 and again, Illinois' name did not appear on the list of states to receive funds, even after adding *Senate Bill 7* requirements to its application. The good news for Illinois is that the state was named as one of the top seven applications in 2011 who did not get funded. In April of 2012, those seven states were invited to revamp their application and try, once again. The prize is dwindling, however, with only 50% of the 2011 award amount up for grabs in 2012 at $133 million.

Summary

Dating back to *A Nation at Risk*, there have been numerous attempts by the federal government, as well as all 50 states, including Illinois, to improve the quality of education. Many of these reform efforts have been predicated upon unreliable, incomplete, or even inaccurate data to support their claims that U.S. students are not performing as well as their peers across the globe.

Reformists have attempted to improve education through a variety of means, most notably, a focus on accountability. This movement evolved after *A Nation at Risk.* As noted by Darling-Hammond, Wise, and Pease (1983): "The demand for accountability in education has shifted from broad issues of finance and program management to specific concerns about the quality of classroom teaching and teachers" (p. 285).

Since 1983, efforts have grown rapidly, and are well-documented in Illinois, to increase teacher accountability for their students' performance. The basis for many of these efforts is to legislate teacher and principal evaluation packages that include student growth and performance as a part of the system. Hargreaves and Fullan (2012) argue for a very different approach. Schools are not businesses, although many business leaders will argue otherwise. They are not stocked with the highest performing parts, to be carefully assembled by skilled craftsmen, or even robots, with the ultimate goal of developing a perfect product (or service) that many will buy for a significantly greater price than it cost to produce. We must seek efforts to break down barriers and foster collaboration, not competition across classrooms to truly improve our students' achievement. Illinois legislators chose a different formula.

The *Performance Evaluation Reform Act of 2010* and *Senate Bill 7* are the two most recent examples of Illinois legislation that focuses on teacher evaluation. It is likely that other states will follow the path that Illinois has set in this endeavor, and pass similar legislative packages. These policies not only require examination of student achievement data to support focus on continued growth, they actually legislate ways that teachers and administrators will lose their jobs if effective growth is not achieved. Indeed, this reform movement can be characterized as land-mark. The question that remains is whether or not it will be the final chapter in the school reform movement.

CHAPTER IV

THE LITERATURE ON TEACHER EVALUATION

Darling-Hammond, Wise, and Pease (1983) described four main purposes for evaluation: "individual staff development, individual personnel (job status) decisions, school improvement and school status (e.g., certification) decisions" (p. 302). These purposes are very distinctive and are difficult to measure with any one tool for evaluation. Instead, teacher appraisal practices must be wide-reaching and all-encompassing to be able to achieve both types of objectives, individual and institutional. The authors went on to detail the requirements for an effective evaluation plan:

> In general, teacher evaluation processes most suited to accountability purposes must be capable of yielding fairly objective, standardized, and externally defensible information about teacher performance. Evaluation processes useful for improvement objectives must yield rich, descriptive information that illuminates sources of difficulty as well as viable courses for change. Teacher evaluation methods designed to inform organizational decisions must be hierarchically administered and controlled to ensure credibility and uniformity. Evaluation methods designed to assist decision-making about individuals must consider the context in which individual performance occurs to ensure appropriateness and sufficiency of data. (p. 303)

Medley (1982) discussed how many different terms are used synonymously, when discussing teacher performance, yet they may have very different meanings. He defined the following terms:

- *Teacher competency* refers to any single knowledge, skill, or professional value position, the possession of which is believed to be relevant to the successful practice of teaching. Competencies refer to specific things that teachers know, do, or believe but not to the effects of these attributes on others.
- *Teacher competence* refers to the repertoire of competencies a teacher possesses. Overall competence is a matter of the degree to which a teacher has

mastered a set of individual competencies, some of which are more critical to a judgment of overall competence than others.

- *Teacher performance* refers to what the teacher does on the job rather than to what she or he can do (that is, how competent she or he is). Teacher performance is specific to the job situation; it depends on the competence of the teacher, the context in which the teacher works, and the teacher's ability to apply his or her competencies at any given point in time.
- *Teacher effectiveness* refers to the effect that the teacher's performance has on pupils. Teacher effectiveness depends not only on competence and performance, but also on the responses pupils make. Just as competence cannot predict performance under different situations, teacher performance cannot predict outcomes under different situations. (p. 41)

While many authors, educators, policy-makers and parents use these four terms interchangeably, much of the time the intended definition is that of "teacher effectiveness," or the effect that teachers have on students.

McGreal (1982) identified nine characteristics that were common among the most effective evaluation systems. All of the items were related to the relationship that took place between the observing administrator and the teacher. They included the following:

1. A shared attitude on the purpose of evaluation being improvement of instruction.
2. The requirements of the system reflect the purpose.
3. *Teacher* evaluation is separate from *teaching* evaluation.
4. Goal-setting takes place.
5. There is a narrow focus on the act of teaching
6. Pre-conferences are held prior to observation.
7. Student evaluation artifacts are examined.
8. There are different requirements for tenured and non-tenured staff members.
9. A complete training program is in place for evaluators and teachers.

Relationship Between Evaluation Score and Student Achievement

A growing body of research supports the notion that teachers are the primary factor affecting student achievement. Hanushek (1992) asserted that good and bad teaching can account for as much as one full grade level difference in just one school year (p. 113). Horn and Sanders (1997) found, "The most important factor affecting student

learning is the teacher" (p. 63). Their examination of standardized test scores of third, fourth and fifth grade students in Tennessee sought to control for classroom contextual variables, such as class size and heterogeneity, to identify a causal relationship between teacher behavior and student learning. Results further indicated that "Effective teachers appear to be effective with students of all achievement levels," and are "additive and cumulative over grade levels with little evidence of compensatory effects" (p. 63). Odden, Borman, and Fermanich (2004) had similar findings that demonstrated the significant impact teachers have on student achievement, and that classroom activities are likely the greatest factor. Their review of research indicated, however, that the current body of research is limited in scope. They pointed out that researchers have focused on controlling for many of the variables associated with the classroom and teacher, but have been unable to account for several.

Borman and Kimball (2005) used an analysis of variance (ANOVA) to determine if teachers with higher standards-based evaluation scores increased student achievement more than teachers with lower standards-based evaluation scores. They analyzed the evaluation scores of 500 teachers and the standardized test scores of 7,500 fifth and sixth grade students. When defining a good teacher as one above the 84th percentile and a bad teacher as one below the 16th percentile, they found differences between 1/10th and 1/5th of a standard deviation in the achievement of the students (p. 16).

Milanowski (2004) analyzed the test scores of 48,000 students and 3,000 teachers over a 2-year period to determine if there was a statistical correlation between teacher evaluation and student achievement. He found that while the correlation was relatively small (0.3 – 0.4), this was generally due to errors in measuring both the teachers' and the

students' performance, lack of alignment between the curriculum and assessment, and a few general student characteristics (p. 50).

While Darling-Hammond disqualified many teacher traits from having an impact on student achievement, she went on to examine the statewide results from the National Assessment of Educational Progress (NAEP) to determine if there was a correlation between teacher quality and student achievement. Her findings indicated:

> Partial correlations confirm a strong, significant relationship of teacher quality variable to student achievement even after controlling for student poverty and for student language background (LEP status). The most consistent highly significant predictor of student achievement in reading and mathematics in each year tested is the proportion of well-qualified teachers in a state: those with full certification and a major in the field they teach. The strongest, consistently negative predictors of student achievement, also significant in almost all cases, are the proportions of new teachers who are uncertified and the proportions of teachers who hold less than a minor in the field they teach. (p. 23)

This study encompassed a tremendous amount of data on students across the country. General results that demonstrate such a strong correlation are clear evidence for the need for effective teacher evaluation.

A study that quantified the statistical effect teachers had on their students was the 2004 work done by Barbara Nye and Spyros Konstantopoulos. Their study examined longitudinal data from the Stanford Achievement Test (SAT) scores of kindergarten through third grade students in 79 elementary schools. The study included a diverse population of students from both rural and large urban districts, as well as high poverty and low poverty schools. They assert:

> The results of this study support the idea there are substantial differences among teachers in the ability to produce achievement gains in their students.
>
> If teacher effects are normally distributed, these findings would suggest that the difference in achievement gains between having a 25th percentile teacher (a not so effective teacher) and a 75th percentile teacher (an effective teacher) is

> over one third of a standard deviation (0.33) in reading and almost half a standard deviation (0.48) in mathematics. Similarly, the difference in achievement gains between having a 50th percentile teacher (an average teacher) and a 90th percentile teacher (a very effective teacher) is about one third of a standard deviation (0.33) in reading and somewhat small than half a standard deviation (0.46) in mathematics. (p. 253)

These findings were representative of a large sample of students across age and ability levels. Students who are struggling will have their delays compounded by the teacher ineffectiveness, leading to the need for years of intervention through Response to Intervention (RtI) or special education services.

Another quantitative study on the correlation between teacher evaluation and student achievement was published by Gallagher (2004). His study was conducted on data provided by Vaughn Elementary Charter School in the Los Angeles Unified School District. The results support the body of research presented thus far that indicates there is a correlation between the two variables, teacher quality and student achievement. He noted the following in summary of his statistical analysis:

- Overall, the Vaughn teacher evaluation system had a statistically significant relationship to classroom effects, that is, value-added learning growth. The strength of the relationship in literacy was much stronger than would have been anticipated from previous research. [The school system went through an intensive professional development process focused on literacy in the years prior to the study]
- Traditional teacher quality variables (e.g., licensure, experience) appeared to be insignificant predictors of variation in student achievement, especially when compared to some more proximal indicators of instruction such as the literacy evaluation score. (p. 100).

By identifying that, in fact, teachers are the factor that has the greatest impact on student achievement, one could assert that through observation of good teachers, a checklist of knowledge, skills, traits, and behaviors could be identified and used as a primer for training and evaluation models. Rockoff (2003) suggested, however, that we

may not be able to specifically identify or observe the teacher characteristics responsible for producing good teaching. He suggested that we focus on policies that reward performance, rather than attempting to legislate a set of credentials that may or may not produce the desired outcome of good instructional practices.

The basic understanding that teachers' skills and behaviors have an effect on the achievement of their students leads researchers to believe the notion that the field should identify which knowledge, behaviors, and skills are responsible for student growth. In a study that examined several of these key factors, Darling-Hammond (2000) generalized that neither the general academic ability and intelligence, subject matter knowledge, experience, nor certification status of a teacher had a significant impact on the teacher's ability. Wayne and Youngs (2003) found the same to be true when looking at identical characteristics. Darling-Hammond's same study found that instructional knowledge had a "somewhat stronger" correlation (p. 5). There was variance in the results of the study, which indicated that math teachers performed better when they had full certification and a major in math (p. 4). They determined that, overall, teacher evaluation practices are superficial and lack any real depth of understanding that would promote the future growth of the teacher.

Another important researcher in the field of education in general, but specifically related to teacher effectiveness, is Robert Marzano. Dr. Marzano's, *What Works In Schools* (2003) used meta-analysis of several studies to examine the impact of the teacher on student achievement. The results indicated that the difference in student achievement, on a standardized test, between students who had the "most effective" teacher versus those students who had the "least effective" teacher, is as much as 39 percentile in just

one school year (p. 72). Even more troubling was the fact that the gap grew to a difference of 54 percentile with the same pairings over 3 years.

The research on this topic is clear; the teacher has a significant impact on student achievement. While non-school factors are frequently attributed to low achievement scores, the quality of instruction is clearly the most important school-level factor. In order for schools and districts to ensure that they are putting students in classes with teachers who use appropriate methods, the teacher evaluation system must be effective.

Evaluation Methods and Models

Principal Observation

Teacher evaluation systems, or performance appraisals, as they are sometimes referred to, have historically included a variety of styles and options. Darling-Hammond et al. (1983) found that principal observation was one of the most widely used techniques in teacher evaluation (p. 306). McGreal, Broderick, and Jones (1984) stated even more plainly that "Contemporary supervisory practice relies almost exclusively on classroom observation for collecting data about teaching" (p. 20). A decade later, Peterson (2004) stated that "the most common method of teacher evaluation in current practice is to use administrator reports based on one or two classroom visits" (p. 60). Most school districts still employ principal observation as their primary source for data in teacher evaluation, even though research indicates that it is wrought with inconsistent findings. Several studies indicate that observer bias, insufficient sampling of performance, and poor measurement instruments threaten the reliability and validity of the teacher appraisal process (Haefle, 1980; Lewis, 1982; Peterson & Kauchak, 1982). Scriven (1981) cited further problems such as "change in the usual teaching practice caused by the visit itself,

unreliable number of samples, personal biases of the evaluator, adult raters who do not think like students, style preferences of the evaluator, and costs in time of lengthy classroom visits" (p. 271).

Downey, Steffy, English, Frase, and Poston (2004) discussed an observation style that accounts for many of the frequent causes for the lack of validity and reliability that have been discussed. The authors have developed a classroom walk-through tool that has the potential for inclusion in good teacher evaluation practices. This measurement utilizes a number of evaluators who are all trained to identify specific types of instruction. On random days throughout the school year, the evaluators spend the day doing 3-minute drop-ins to determine the type of instruction and engagement that are occurring in the classroom. The observations are purposely scheduled throughout the building to ensure that classes are not observed during transitional times, and that all classes and teachers have a representative number of data points throughout the year. If an observer enters a class during a break or change instructional techniques, the observation formally begins when the next activity begins.

This strategy can be much more effective in the school improvement goal of evaluation, as building teams can set goals for the percent of time spent on different types of instruction (e.g., lecture, small group, independent work, etc.), and then the observational data can be compared to determine progress toward the stated goals. Frequently, these observations are not done by the same person (principal) who completes the individual performance appraisal of staff members for personnel decisions, as authentic data should be gathered with no fear of judgment. The evaluators code the instruction type to generic classroom descriptors so that the data are not linked to

individual teachers. This type of assessment system can yield thousands of observations in a school year to truly define the type of instruction students are receiving. While this is extremely useful for school improvement, it is useless for individual assessment.

Student Surveys

Xu and Sinclair (2002) extrapolated that multiple data sources were required to get an accurate picture of teacher performance. Another possible source of data, according to both Darling-Hammond et al. (1983) and Peterson (2004) is student surveys. Students have a somewhat surprisingly profound knowledge about their teacher's effectiveness. However, the validity and reliability of this technique relies on the students understanding when they learn, students needing to learn, and student feedback being directed to the teacher (Darling-Hammond et al.; Peterson).

Ensuring that these criteria are accounted for, along with outliers in student reporting, makes it extremely difficult to have a suitable confidence level with this type of evaluation. It is another source of data that has value, especially in identifying potential areas for further investigation, and could be an effective component of a comprehensive evaluation program.

John Hattie (2012) also noted that student surveys had a very high correlation with effective teaching. Hattie compiled a meta-analysis of over 500,000 studies to try to isolate those teaching strategies that are most likely to produce results. In his analysis, Hattie notes that almost everything done in a school has a positive effect on student achievement. For that reason, anything less than a 0.4 effect size is considered to be usual and customary, and not having a profound effect on student learning. However, the strategies that have greater than 0.4 are considered to be the strategies that teachers

should seek to employ regularly.

Student feedback was ranked as a 0.6. Hattie explains that kids generally have a very good sense about their teacher's level of effectiveness because the key ingredient is a connection with students. If the teacher has it, the kids know it and score the teacher high. If they do not have a strong bond with their students, the kids know it and rate them low. This leads to the 0.6 correlation between student perception of teacher effectiveness and student achievement.

Peer Feedback

Another source of information that both Darling-Hammond et al. and Peterson cite as a possible source of valuable data is peer feedback. Even at the primary and secondary levels, peer review can provide relevant data from respected colleagues who are knowledgeable about the expectations, demands, and situational data that affects classroom teaching.

Marzano (2003) also found a correlation between peer feedback and student achievement. Authentic, professional discussion about instruction had a correlation of 0.326, suggesting a strong relationship with student achievement. Interestingly, teacher friendship had a *negative* correlation (-0.252) with student achievement. This is not to say that teachers cannot be friends and still promote student achievement. However, it does suggest that friendship may limit professional dialogue, thereby limiting the growth that teachers experience by engaging in meaningful discussion with peers.

The term "relational trust" was first used by Bryk and Schneider (2002, p. 12). They defined relational trust as "the distinctive qualities of interpersonal social exchanges in school communities, and how these cumulate in an organizational property."

According to Blankstein, relational trust among staff members is one of the key elements of a successful professional learning community (2004). In his meta-analysis, Blankstein points out the relative ease of implementing systems to promote a professional learning community. However, developing relational trust amongst staff members is one of the more difficult, yet critical tasks associated with school improvement. Without people, leaders are not leading and have no place to lead to.

Careful consideration must be given in the assignment of peer review teams to be certain that the benefits are fully recognized. Utilizing teachers whose performance is sub-standard, or who lack the ability to protect confidentiality, poses a threat not only to the integrity of the peer review but also to the meaning that is gained by the teacher being reviewed.

Standards-based Evaluation

In 1996, Charlotte Danielson introduced her framework for teaching to educators. She stated, "A framework of professional practice for teaching is useful not only to practicing educators but also to the larger community, because it conveys that educators, like other professionals, hold themselves to the highest standards" (p. 2) This quotation represents a thesis statement for the evaluative theory that many authors refer to as "standards-based evaluation."

Milanowski, Kimball, and White (2004) defined standards-based evaluation as a "comprehensive model or description of what teachers should know and be able to do, represented by explicit standards covering multiple domains and including multiple levels of performance defined by detailed behavior rating scales" (p. 2) Danielson's model is still widely regarded today as the definitive standards-based model as its four domains

break apart the art of teaching into twenty-two components. A rubric differentiates levels of performance within each domain as "unsatisfactory, basic, proficient and distinguished." Exhibit 2 is an example from Danielson's rubric.

Exhibit 2—Portion of Domain 1 of the Danielson Rubric for Teacher Evaluation

DOMAIN 1: PLANNING AND PREPARATION
Component 1D: Demonstrating Knowledge of Resources
Elements: Resources for teaching • Resources for students

	LEVEL OF PERFORMANCE			
Element	**Unsatisfactory**	**Basic**	**Proficient**	**Distinguished**
Resources for Teaching	Teacher is unaware of resources available through the school or district.	Teacher displays limited awareness of resources available through the school or district.	Teacher is fully aware of all resources available through the school or district.	In addition to being aware of school and district resources, teacher actively seeks other materials to enhance instruction, for example, from professional organization or through the community.
Resources for Students	Teacher is unaware of resources available to assist the students who need them.	Teacher displays limited awareness of resources available through the school or district	Teacher is fully aware of all resources available through the school or district and knows how to gain access for students.	In addition to being aware of school and district resources, teacher is aware of additional resources available through the community.

This model is so widely accepted as the preeminent standards-based evaluation model that the Illinois General Assembly specifically cited this model as the default model for districts to use when implementing the requirements of the *Performance Evaluation Reform Act*.

The National Board for Professional Teaching Standards is an organization that was created based upon the recommendations of the *Carnegie Forum on Education and the Economy*'s 1986 research, a result of *A Nation at Risk*. The impetus for the organization is, "developing the policies on which to base its assessment system, and developing prototype performance assessments" (Baratz-Snowden, p. 82). The

organization has developed a standards-based appraisal system that teachers must complete in order to become National Board certified. The tests consist of interviews, simulations, and pencil-paper exams used to determine if the candidate has the knowledge, skills, and traits identified as critical to the art of teaching.

Borman and Kimball (2005) noted in their assessment of standards-based evaluation that its inherent goal was not only to improve instructional practice, but also to "strengthen educational accountability" (p. 5). Danielson and the other advocates of standards-based evaluation have clearly identified with the pulse of the reform movement and attempted to provide a research base for good teachers to use with reformists. The findings of Milanowski et al. (2004) support this objective by suggesting that "evaluation scores from well-designed and implemented standards-based teacher evaluation systems can be used for decisions about teachers, and have potential for use in studies of teacher effects on student achievement" (p. 19).

Self-Assessment

Self-assessment is a strategy employed by some school districts as a beginning point for the evaluation process (Darling-Hammond et al., 1983; Peterson, 2004). This process is especially effective in the practices of goal-setting and reflection. Both of these methods are successful in motivating staff members to grow and improve their craft as professionals. While it may be a useful component in the teacher appraisal process, clearly self-assessment cannot be the foundation for personnel decisions or even system analysis. The validity and reliability of self-reporting is critically low (Darling-Hammond et al., 1983).

Value-Added Models of Evaluation

Several articles have begun to focus on the latest trend in teacher performance appraisal. Student achievement, pupil achievement, or value-added models are all terms used in research to discuss the application of student assessment data to the teacher evaluation process. This topic has reached a pinnacle of significance in the past year, especially in Illinois, with the passage of the *Performance Evaluation Reform Act* and *Senate Bill 7*. While these pieces of legislation have gained a great deal of notoriety of late, the idea of using student achievement data in teacher performance appraisal is not a new idea. McCaffrey, Lockwood, Koretz, and Hamilton (2003) provided a definition for the goal of value-added modeling: "to estimate the effect of educational inputs on student outcomes, in particular student achievement as measured by standardized tests" (p. 7).

Darling-Hammond et al. discussed in 1983 that student achievement may be the only factor important to some educational audiences (p. 307). They argued, however, that "numerous assumptions [have to] be made to link them to teacher competence or even teacher performance" (p. 307). Peterson (2004) also issued words of caution regarding the use of student achievement data for teacher evaluation purposes. He stated that it "should only be used where reliable and valid data are available and where teachers individually select and control the information—including the ability to withhold their own results" (p. 64). The new Illinois legislative packages do contain provisions for administrators and educators to work together to choose the measures and benchmarks for student growth and proficiency that will be part of the individual teacher evaluations. However, the ability for teachers to "withhold their own results" is not part of the package.

Another concern some researchers have pointed toward in the discussion on use of value-added measures is the lack of availability of data for all teachers, especially special-area teachers (e.g., art, music, and physical education). Researchers also point to the notion that is important not to assume a causal effect in all situations. McCaffrey et al. (2003) noted:

> The causal effect of a teacher will depend on the measure of achievement. Effects on one measure of achievement will not necessarily equal effects on other measures. Users should choose a measure of achievement that suits the desired inferences. The most commonly used measures of student achievement are scale scores from standardized tests. However, these are not the only available measures. For example, criterion referenced test scores that are not on a single developmental scale might also be used. (pp. 15-16)

Missing data is another concern with standardized tests. Frequently, the student population in any group will have some variance from year to year. These cumulative holes can have a significant impact on the analysis of test scores when attempting to determine teacher effectiveness in a cohort over time. Braun (2005) reported similar concerns in his primer on the use of value-added models. He noted the following:

> These models require data that track individual students' academic growth over several years and different subjects in order to estimate the contributions that teachers make to that growth…Indeed, the implementation of such models and the proposed uses of the results raise a host of practical, technical, and even philosophical issues. (p. 3)

Additional concerns about making a "causal attribution" (p. 7) to student achievement data are found in the research. Many factors have an influence on the overall learning process, and many of them can be controlled for. Nonetheless, "No statistical model, however complex, and no method of analysis, however sophisticated, can fully compensate for the lack of randomization" (p. 8). In discussing randomization, Braun is referring to the lack of randomness about how students are placed in classes. While the

process is not necessarily complex, students are traditionally placed in classes with specific teachers to match learning styles, abilities, behaviors, skill levels, and many other factors. This matching process degrades the statistical analysis that would need to be completed in order to accurately and precisely determine the true effect a teacher has on a group of students, based on their level of skill.

One example Braun shares that illuminates this deficit in the process deals with teacher seniority. In many cases, senior teachers have first choice of assignments and are able to choose classes and buildings that have more resources and students with higher levels of preparation, readiness, and intrinsic motivation. A statistical analysis of student achievement data would very likely reveal a higher level of growth and performance than that of students in a building with fewer resources, low levels of preparation, and a lack of intrinsic motivation (p. 9).

This observation leads to a significant concern about the overall state of data on student achievement across the country. The achievement gap is frequently pointed to when debates about educational effectiveness begin. Data frequently shows students from poverty, limited-English proficiency, or minorities achieving at lower levels than their non-disadvantaged peers. Many government officials would like to make a causal attribution to the teachers' skill. However, Braun's (2005) research indicates that exchanging teachers from high performing schools with those from low performing schools may not have a drastic effect, as the resources, level of preparation, and motivational factors may also have a significant impact on the achievement of the students and would not change with a new face.

Darling-Hammond, Amrein-Beardsley, Haertel, and Rothstein (2012) engaged in further research on the topic of value-added modeling. They found that this technique attempts to use statistical analysis methods that control for all of the following:

- School factors such as class sizes, curriculum materials, instructional time, availability of specialists and tutors, and resources for learning (books, computers, science labs, and more);
- Home and community supports or challenges;
- Individual student needs and abilities, health, and attendance;
- Peer culture and achievement;
- Prior teachers and schooling, as well as other current teachers;
- Differential summer learning loss, which especially affects low-income children; and
- The specific tests used, which emphasize some kinds of learning and not others and which rarely measure achievement that is well above or below grade level. (p. 8)

While they reported that the variables were, indeed, well controlled, the researchers also found a great deal of inconsistency in value-added models. For example, individual teacher effectiveness scores vary greatly from class to class and year to year. It would be reasonable to expect that a good teacher would score well across classes and years. Obviously, there can be changes in teachers' performance over time, but there should be some consistency in their scoring across classes.

Darling-Hammond et al. also found significant differences in teacher effectiveness when different measures were utilized. If the same teacher and pupils are being measured at that same period of time using two distinct measures, one should expect similar results. Without a high level of reliability, value added measures cannot possibly be successfully instituted as the definitive performance appraisal strategy.

While Braun's assertions about the lack of ability to effectively control all variables are warranted, the research that has been presented previously, documenting the

overall effect of good teaching, is profound as well. Both schools of thought are accurate; good teaching is the most important factor in the classroom, and external factors (resources, preparedness, motivation) also combine to provide a very strong influence. To address the achievement gap and reach students who need the most support through programs like Response to Intervention (RtI), educational institutions need to provide a system that attracts good teachers to schools, classes, and buildings with the greatest need. In order to accomplish this, there will need to be a compensation program that adjusts for the difficulty of assignment. Further, society will need to recognize that even the greatest teacher is not going to have a significant effect on all students in all situations. Consideration for policies and programs that focus on other factors as the cause for the lack of educational attainment of these students will need to be considered.

A final concern about value-added models that Braun discussed was that, in some situations, students' test scores from a variety of subjects over as many as 5 years are analyzed in databases. A primary example would be to use *Illinois Standards Achievement Test (ISAT)* scores in a building which houses third through eighth grade students. While the name and type of assessment are the same, the measure is different for each grade level. An assumption is made that the test is grade-appropriate at each level. Braun stated, "Consequently, as we move to higher grades, the detailed specifications that govern the construction of each test will reflect the greater dimensionality and expanded knowledge base of the subject" (p. 13). In other words, a great deal of test analysis would need to be completed in order to validate the notion that test scores over a period of years, from different grade levels, are an equivalent measure of student progress. Without this analysis, statistical comparison is invalid.

Portfolios

Another aspect of the National Board for Professional Teaching Standards' assessment process is the creation of a portfolio to document the candidates' application of National Board Standards into their classroom planning, organization, and execution (Baratz-Snowden, 1993, p. 84). This assessment method engages the candidate in a reflection process, which first provides them an opportunity to determine whether National Board certification is a legitimate, attainable goal, or whether further remediation and study will be required.

Blake, Bachman, Frys, Holbert, Ivan, and Sellitto (1995) stated, "The inclusion of reflective practices must become part of this new method if professional growth is to occur. One way schools can promote reflection and self-assessment is to encourage the use of teacher portfolios" (p. 38). The gathering of artifacts over time not only encourages reflection, but leads to continuous improvement. This is a foundation of any effective school improvement planning process.

Blankstein (2004) wrote about six key principles that were present in high-performing schools. Principle number four was "Using Data to Guide Decision-Making and Continuous Improvement" (pp. 141-165). This section of Blankstein's book discussed the critical role that gathering data sources and reflecting on its relevance plays in continuously improving results. Regardless of the success that the school or district has achieved at any point in time, current accountability measures, such as those cited in the previous sections of this paper, require that schools continue to improve.

Portfolio assessment practices for teachers address an aspect that is rarely thought of when considering school improvement and reform: adult learning styles (Blake et al.,

1995). Adults require more intrinsic motivation, unlike children and adolescents who require more extrinsic motivation. Portfolios allow adults to display their accomplishments and highlight their strengths, which provide further job satisfaction (intrinsic motivation). They allow teachers to build upon their assets, yet areas of weakness that are underrepresented in the portfolio also draw attention. However, it is focused in a manner that motivates the individual to achieve greater proficiency in the area, rather than punishing them for a deficiency.

Blankstein discussed the usefulness of portfolios in the continuous improvement process:

> Teams will invariably look at data to assess how they are doing relative to SMART goals. Members collectively brainstorm ways to improve and celebrate successes. Being committed to constant improvement, these teams will always find ways to 'raise the bar' once their current goals are accomplished. (p. 130)

His statement demonstrates that when teachers are organized in groups, with a common SMART goal, specific, measurable, achievable, realistic, time-related (Doran, 1981), they will need specific pieces of data to help them understand the progress they are making. Portfolios are one method for tracking instructional practices, and changes in them, to determine overall effectiveness.

Marzano also presents an argument for teacher portfolios. He stated, "…collegiality is characterized by authentic interactions that are professional in nature" (p. 61). This is not achieved simply by causing teachers to work together or engage in co-planning processes. Having a defined practice of documenting artifacts and engaging in professional dialogue and critique helps to form the level of collegiality that Marzano suggested we strive toward.

The constructivist theory suggests that the learner makes meaning from real-world situations and applications of knowledge (Attinello, Lare, & Waters, 2006, p. 135). This theory is one that is being heavily supported in the advancement of student learning. Some researchers are now suggesting that it can also be a powerful tool for teachers to increase their knowledge and skill level. Early evaluation methods focused on summative ratings that provided little opportunity to make meaning and grow. The constructivist paradigm, driving for authentic assessment, would direct school leaders toward practices, such as portfolios, that are formative and allow the teacher to become the learner (Attinello et al.).

Performance evaluations rarely contain portfolios in teacher appraisals today. While there is a great deal of benefit from the process, the reliability of portfolios is very low, according to the empirical research completed by Attinello et al. Many teachers felt that their portfolios were accurate reflections of their skills and abilities, but also recognized that it would be very easy for unskilled educators to put together an excellent portfolio. Likewise, a highly effective teacher could have very poor documentation of their effect through portfolio assessment. For these reasons, it is difficult to use portfolios as a measure for personnel decisions. The need for other evaluation tools and models makes the addition of portfolios cumbersome.

Summary

There is a growing body of well-documented research in this section that discusses how influential the teacher is in affecting student achievement. Much of the research has shown teacher effects being constant across grade and student performance levels, when other variables are controlled for. Such measures show effects as high as one

full grade level difference in achievement for a given year.

To design and implement an effective teacher evaluation system, it essential the system be valid and reliable. These two elements are key to any evaluation system, and are critically important when assessing whether or not a staff member should be re-employed. Elements such as self-assessment, principal observation, peer observation, student growth data, and evidence portfolios can all be valuable components to a comprehensive evaluation plan.

CHAPTER V

FINDINGS

The first and most important goal of any school district is to maximize student achievement, or in other words, to give students the greatest opportunity for success. As this study has cited, the teacher is the primary factor that influences students' success once they reach the school building. With this understanding, it is imperative that an effective evaluation system is in place to monitor teacher performance, as well as provide feedback to promote continuous growth and development as an educator.

With the adoption of PERA and Senate Bill 7, all Illinois public school districts are required to make changes to their evaluation plans. The new plans will be required to contain student growth and performance data as a significant factor in the teachers' overall summative evaluation rating. Simply modifying the current tool to "add in" student growth misses an opportunity to redesign the evaluation plan into a process that recognizes excellent practices, acknowledges needed areas of growth and provides guidance toward increased teacher performance and ultimately student achievement. This chapter outlines a process for developing a tool that can successfully meet these goals.

Assemble the Stakeholders

The starting point for any instructional redesign process should be to identify the stakeholders. In the teacher evaluation plan, there are a number of stakeholders, depending on the size of the district, including teachers and administrators. While

students, parents, school board members, and taxpayers are all stakeholders in the school district, they do not have a vested interest in the *process* of teacher evaluation itself. They are quite concerned with the *outcomes* of the system, but have neither the knowledge nor experience to build or evaluate the plan.

A working stakeholder group needs to be identified that can put forth a significant amount of time and energy, specifically devoted to designing the evaluation plan for the district. As Heifetz (2004) discussed, we must mobilize those affected to deal with the problem. While implementing a teacher evaluation incorporating student growth may not be a problem, it certainly meets Heifetz' definition of an adaptive challenge and will require growth by the organization.

Under PERA, the evaluation plan is a mandatory topic of bargaining with the joint committee of the local education association and must be completed within 180 days from the start of the negotiations. This is not the stakeholder work that is being suggested at this point. In fact, it is important that once this stakeholder group is identified and formed, they should explicitly discuss the fact that they are not to be considered the joint committee as defined by PERA. To assist with this clarity, giving this committee the title of "Appraisal Committee" will avoid legal bargaining implications.

In identifying the Appraisal Committee, the district should endeavor to reach the key players: teachers who carry a high level of respect among their colleagues, have a high level of effectiveness in the classroom, and will be able to communicate well with both administrators and their colleagues. While this first step may seem trivial, it is vital to the outcome of the plan. The Appraisal Committee must not only create an effective

evaluation plan, but must ensure that the plan is ultimately approved in all of the necessary channels following their work and implemented with integrity to their original design.

A significant amount of work is to be undertaken by the committee, so having enough people to be representative of the district, as well as to break up the work, is essential. However, having too many people will lead to excessive debate and will deter the process. An Appraisal Committee of approximately 20 to 25 people composes a manageable group that can accomplish a great deal.

The committee should begin its work by discussing why they are gathered and outlining the tasks at hand. They should begin with a foundation for their work that includes what is important to them in the evaluation process. At this point, the committee should be working from a "blank slate" mentality. Forget about the laws and regulations and focus on what is good and helpful in an evaluation process and commit to those right away. Some of the items may include trust, valid, fair, focused on growth, collaborative, meaningful, etc. When the committee gets into debate about content or process, revisit these foundational principles that the Appraisal Committee deemed to be critical and put the item in question to the test.

Identify the Evaluation Components

The Appraisal Committee is formed and understands their charge and vision, and now it is important that everyone understands the best practices in teacher evaluation. Chapter III provided detailed reviews of many different techniques for evaluating teachers. The Appraisal Committee should review each of the techniques presented to get a complete understanding of its value, as well as its flaws. There is not one perfect

tool, nor is there one perfect method for evaluating teachers. There is much about teaching that is a craft, or art form. To assess an art form, one must obtain a 360-degree view of the entire picture to get a true understanding of it (Xu & Sinclair, 2002). While choosing which components are included in the tool should be a valuable discussion undertaken by the Appraisal Committee, several strategies or techniques are suggested below, including the reasoning for their suggested inclusion. Some of the strategies are required by law, while others are suggested to meet the desired outcomes of the evaluation plan.

Professional Goal-Setting

An effective evaluation system should begin with goal setting. As Blankstein (2004) discussed, staff members will strive for greater, continuous improvement when they are aware of their professional performance and plan for specific improvement.

Student growth goals are required under PERA. Additionally, it is important the teachers reflect on their previous performance and evaluations to set professional improvement goals for themselves. The principal plays an important role in this process by reviewing the goals with the teachers to ensure their targets are S.M.A.R.T., and that the teachers have correctly identified the areas most in need of improvement.

The teachers' professional improvement goals are not required under PERA or SB 7 and do not need to have a direct impact on their summative rating. However, as previously discussed in this study, setting S.M.A.R.T. goals is necessary if improved performance is desired. These goals will assist the teachers in remaining focused on the aspects of their performance that are targeted in the goals. This goal setting process also provides a framework for the teachers' individual professional development program

during the evaluation cycle. This could include professional reading, as well as attending workshops or conferences on specifically identified topics in the goals.

Student Growth Goal-Setting

Goals for each teacher, based upon the student growth and performance data of their students, is required beginning in the 2016-17 school year. Teachers in the lowest-performing 20% of districts as measured by standardized test performance are required to begin implementation in 2015-16. At least 25% of the summative evaluation rating must be comprised by student growth data. In subsequent years, the summative evaluation rating must be comprised of at least 30% student growth data.

While the incorporation of student growth is required in law by Senate Bill 7, most likely to satisfy the Race to the Top requirements, there is also considerable research that supports it as a measure of teacher effectiveness. Hanushek (1992), Marzano (2003), and Hattie (2012) all documented high correlation between teacher effectiveness and student growth and performance on standardized tests. The difficulty in implementing the requirements of Senate Bill 7 is developing a system that will accurately account for inconsistencies.

Student growth data comes from three distinct types of assessments as described in PERA:

> **Type I** assessment means a reliable assessment that measures a certain group or subset of students in the same manner with the same potential assessment items, is scored by a non-district entity, and is administered with either statewide or beyond Illinois. Examples include assessments available from the NWEA, Scantron Performance Series, Star Reading Enterprise, SAT, AP, or ACT EPAS.
>
> **Type II** assessment means any assessment developed or adopted and approved for use by the school district and used on a district-wide basis by all teachers in a given grade or subject area. Examples include collaboratively developed

> common assessments, curriculum tests, and assessments designed by textbook publishers.
>
> **Type III** assessment means any assessment that is rigorous and is aligned to the course's curriculum, and that the qualified evaluator and teacher determine measures student learning in that course. Examples include teacher-created assessments, assessments designed by textbook publishers, student work samples or portfolios, assessments of student performance, and assessments designed by staff who are subject or grade-level experts, that are administered commonly across a given grade or subject.

Whenever possible, Type I assessments must be used. These standardized assessments are the most reliable assessments because they are rigorously built and tested to ensure reliability. However, two very significant problems result from their use. First, there is no certainty of validity. The simple fact that students in fifth grade are given an assessment designed to measure fifth grade reading ability does not necessary mean that the teacher had anything to do with the assessment results. The test may not be aligned to the local curriculum, if one exists at all. Additionally, the student could have had a number of issues that resulted in a poor test score. To the greatest extent possible, the evaluator and teacher must work to ensure that the tests used are aligned to standards and that the population tested, for purposes of the teacher evaluation, have an opportunity for success on the assessment. They should not have been excessively absent, nor subject to any instructional or testing conditions that could negatively affect their score.

The second significant problem with Type I assessments is their availability. While there are a number of Type I assessments available for regular classroom teachers of math and language arts across all grade levels, there are far fewer available for science and social studies. There are few, if any, Type I assessments available for most special area classes, such as art, physical education, music, band, family and consumer sciences,

industrial technology, and many more. Teachers of these classes, as well as special education teachers, are forced to use Type II assessments, which are nearly as scarce in these areas as Type I assessments.

All teachers must use a minimum of two assessments, one of which must be a Type III assessment. These local assessments offer the highest level of validity because they are a direct match to the curriculum being taught. However, because they are created by local educators for a specific unit or course of study, their reliability is highly questionable. To ensure the highest level of reliability possible in this circumstance, the teacher is required to use the Specific Learning Objective (SLO) process for all Type III assessments.

The SLO process is a detailed analysis that the teacher and principal go through to determine the objectives of instruction, appropriate student population, expected learning and growth targets, and measurement tools and models that will be implemented to understand the students' growth and performance. This information is used along with the Type I or II assessment data, if available; otherwise, the other Type III data is used to calculate an overall student growth score that is applied to the summative rating. The state model SLO template (Appendix) is attached to provide an overview of what is required in the process. However, it is important to note that local districts are allowed to change the form to fit their distinct needs, as long as all legal aspects of the process remain in place.

It is important that the teacher and administrator begin the goal setting process early in the year to properly identify baseline data and then navigate through the determination of assessments and SLO process outlined. Frequent checks of data are

encouraged to drive the instructional path taken by the teacher. PERA requires that student achievement data is available to the teacher at least once during the mid-point of the evaluation cycle.

Self-Assessment

A teacher self-assessment based on a specific set of clearly defined standards is required under PERA and is supported by the best practice research as presented in Chapter IV. There is no requirement for the self-assessment to have any weight or impact on the final summative rating. This assessment does, however, provide an opportunity for teachers to review the performance standards identified in the rubric and reflect on where their performance is relative to the standards.

Completing the self-assessment at the beginning of the year will allow help facilitate that professional goal-setting process for the teacher. The tool should assist in identification of which standards and objectives are more difficult for the teacher to achieve top ratings. However, PERA does not require the self-assessment to be submitted to the evaluator until February 1st of the summative evaluation year.

Standards-based Rubric

A rubric containing a detailed description of levels of accomplishment of teacher activities is essential in measuring teacher performance. Many traditional evaluation systems employ a checklist-style document that calls for the evaluator to make judgments about the level of proficiency on a set, or sets, of skills. This system leaves a great deal of objectivity and bias in the hands of the administrator by not identifying the specific, describable activities and behaviors the teacher is expected to demonstrate.

Charlotte Danielson (1996) authored one of the most popular standards-based evaluation rubrics. The Illinois General Assembly made this the default rubric for Illinois' model evaluation plan. It is not required that local districts use the Danielson model, but rather that they see a clear example of what a framework should look like. (See Exhibit 2, Chapter IV, p. 50.)

There are other standards-based rubrics that are valuable instructional guides, such as Marzano (2003). It is important that the Appraisal Committee review the rubrics and choose the one that most closely aligns to their beliefs about instruction. The rubric should be specifically adjusted in any area not completely clear and aligned to the local expectations. It is also suggested that the Appraisal Committee weigh the scoring on the rubric by domain based on the level of necessity for student achievement.

Principal Observation

Principal observation is the most commonly utilized type of teacher performance assessment. It provides a wealth of data to the evaluator, in a relatively short period of time. However, this technique can be extremely unreliable, with a number of factors and biases influencing the validity and reliability of data. Reliability can be significantly improved when observations are given a purpose and paired with a performance assessment rubric.

Two types of observations are commonly used in teacher evaluation. Formal observation includes pre-scheduled times when the principal spends most of or all of one full class period seeing a lesson from start to finish. This observation is arranged through a pre-conference, where the teacher and administrator discuss the lesson that will be observed, the learning objectives or expected outcomes, assessment techniques, and

instructional tools that will be utilized. The observation is held shortly after the pre-conference; a post-conference is held shortly after the observation to review the findings.

The formal observation provides significant data to the evaluator. With a performance rubric, the principal is able to identify which behaviors the teacher is engaging in and at what level. This leads to a prediction about their effectiveness as it relates to student achievement. A minimum of one formal observation is required under PERA, but more observations are suggested to increase reliability of the observational data.

Informal observation can be an even more powerful method of teacher evaluation, when used appropriately. This type of observation occurs when an administrator observes a teacher without pre-scheduling the observation. This "drop in" or "unannounced visit" can include regular classroom observation or other areas under the teacher's responsibility, including hallway, recess, parent meetings, etc. Under PERA, administrators are allowed to use data gathered during an informal observation, provided the data is reduced to writing and the teacher is provided with an opportunity to meet and discuss the data.

Informal observation increases the validity of the data over formal observation from the standpoint that the teacher is engaged in their typical, regular behavior, in most cases. During a formal observation, the teacher is likely to "show off" or give their very best performance. While it is valuable to see what levels a teacher is capable of reaching, the point of the observation is to see what the teacher is doing most of the time. This allows not only for an accurate rating of the teacher's ability, but also for appropriate feedback to influence the teacher's professional growth.

The possible flaw in informal observational data is that it, too, can be outside the norm. Not everyone is performing at their peak of regular performance at all times.

Difficult situations and circumstances can cause teachers to compromise their standard of performance in order to resolve a conflict. These situations are noticeable and can easily become informal observations that skew the perception of the evaluator. However, this is the reason for the safeguard requirement of putting the data in written form and offering a discussion. This assists the evaluator in determining to what level the observed data is regular behavior.

It should be the goal of the evaluator to identify as much "normal" teaching behavior as possible throughout the evaluation cycle; the greater the amount of "typical" data on a given teacher, the greater the validity of the evaluation. As has been previously mentioned, it is critical that the observations be aligned to a performance-based rubric.

Portfolios

Portfolios are another type of teacher appraisal tool that researchers in the field have identified as providing a great deal of promise. Blake and his colleagues (1995) went so far as to state that reflective portfolios must be included if professional growth is expected to occur. This technique, in isolation, does not provide a high level of validity nor reliability. Rubrics provide intrinsic motivation for educators and promote continuous improvement of the staff member and, in turn, the building as a whole.

Teachers should be required to assemble a portfolio of artifacts throughout the school year that documents their growth toward, and mastery of, the standards presented in the rubric, as well as formative assessment data on their students' progress toward the specified learning objectives. In the section on Principal Observation, it was suggested that the evaluator seek as much evidence as possible that documents the teacher's "normal" teaching behaviors. Portfolios can assist with this collection of evidence by

providing a much more thorough view of the teacher's class. Examples of evidence could include lesson plans, assessments, technology uses, differentiated instruction practices, grading practices, student surveys, peer observational feedback, photos/videos of student growth, learning, and achievement.

Portfolios can provide the teacher with the freedom to present in a manner that is most comfortable to them. The contents can show clearly through print, or other media, what happens inside the walls of the classroom. This evidence can be very powerful, as classroom observation, even informal observation, can still only account for a fraction of the time a teacher spends with their students throughout the year. A portfolio, however, can show the types of activities and behaviors that a teacher engages in daily.

An important role for the evaluator when reviewing the portfolio is to engage in a critical analysis of the contents. The evaluator must review the artifacts along with the observational data gathered to determine if the artifacts are representative of what normally happens in the classroom, or what the teacher wants to appear as what normally happens in the classroom.

Student Surveys

An evaluation strategy with a 0.6 effect size related to student achievement should certainly be included in an effective evaluation program (Hattie, 2012). However, due to the inconsistency with different age and ability level of students, it is difficult to make student surveys a required, weighted component of an effective evaluation system. That does not mean, however, that there is not an appropriate place for this technique in the teacher evaluation plan.

Teachers should be encouraged to utilize effective student surveys of teacher performance. There are no special requirements for the contents of the survey. Their objective is to gain an understanding, from the students' perspective, on whether or not the teacher is effectively teaching them. Teachers should strive to make meaningful connections with their students, which will be evidenced in student surveys.

Not only can the surveys provide data through an annual administration, but they can also provide data on the teacher's growth in this area by administering the survey with the same population at two or three points throughout the year. The raw and disaggregated data, along with planned and achieved improvements, are excellent examples of evidence sources for a portfolio.

Peer Observation

With one of the primary goals of teacher evaluation being professional growth, peer observation is a useful component of the plan. This feedback is another component that is well suited for a portfolio. As was described with student surveys, peer observation also has no special requirements to have an impact on teacher growth.

The teacher should seek a trusted colleague from a pre-approved list of mentor teachers from whom they can request feedback. The teacher is required to work out arrangements to provide the mentor teacher with necessary materials (lesson plans, assessment practices, differentiation strategies, etc.) and set up a time for observation. Once complete, the mentor provides the teacher with feedback on how well they have accomplished the observed standards from the rubric. Not only could the rubric feedback itself be included in an evidence portfolio, but also the teacher's reactions and summary/evidence of implemented changes, based upon the feedback.

Marzano (2003) found a strong correlation between professional dialogue and increased student achievement. Fostering professional learning communities that regularly engage in honest, open and constructive feedback leads to gains by the individuals, and is critical to achieve gains by the school as a whole.

Outline the Process

After the Appraisal Committee has reviewed all of the possible components for the evaluation system and decided on which are appropriate for their local evaluation plan, they must outline the process that will be used to complete the teacher evaluation. Figure 1 illustrates a vision for the teacher evaluation cycle under the PERA/Senate Bill 7 era. This figure was created to provide a visual representation of the interaction and components of an effective teacher evaluation program.

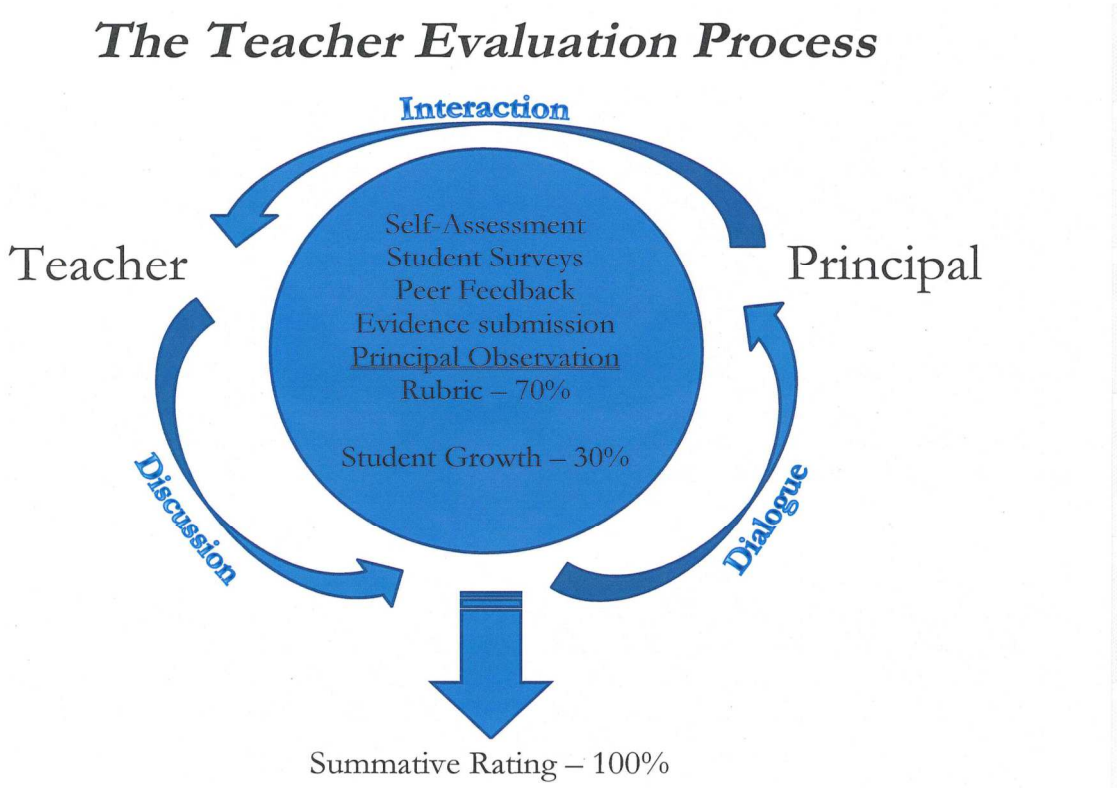

Figure 1. The Teacher Evaluation Process

To be successful, the evaluation cannot be a week-long process where the teacher meets with the principal for a pre-conference, has an observation, meets for a post-conference, and then receives their summative evaluation. Not only will this antiquated process not meet the demands of the PERA/Senate Bill 7 legislation, it certainly will not be sufficient to promote teacher improvement and student achievement. To be highly effective, there needs to be on-going interaction between the teacher and principal throughout the school year, including continuous dialogue and discussion about the teacher's performance, students' achievement, and the learning process taking place by all.

The required components of the evaluation plan for both tenured teachers and non-tenured teachers are included in this sample timeline.

Exhibit 3 – Sample Teacher Evaluation Timeline

Evaluation Key Dates - Tenured

DATE DUE	PURPOSE	OUTCOME
10/31	Set Student Growth Goals (2) and Professional Goal (1)	Goals set
1/15	Mid-point data check	Review student data, determine necessary changes
2/1	Pre-Conference/Observation/Post-Conference (The Formative Process)	Some of eval doc completed NO SUMMATIVE SCORE AT THIS TIME
2/15	Student Growth/Teacher Evidence/ Self-Reflection (Meeting)	Final growth data/evidence submitted to principal
3/1	Summative Meeting - Final	Summative document signed and delivered

****Additional check-in meetings are optional.*

Evaluation Key Dates – Non Tenured

DATE	PURPOSE	OUTCOME
10/31	Set Student Growth Goal (2) and Professional Goal (1)	Goals set
12/1	Pre-Conference/Observation/Post-Conference (The Formative Process)	Some of eval doc completed NO SUMMATIVE SCORE AT THIS TIME
1/15	Mid-point data check	Review student data, determine necessary changes
2/1	Pre-Conference/Observation/Post-Conference (The Formative Process)	Some of eval doc completed NO SUMMATIVE SCORE AT THIS TIME
2/15	Student Growth/Teacher Evidence/ Self-Reflection (Meeting)	Final growth data/evidence submitted to principal
3/1	Summative Meeting - Final	Summative document signed and delivered

****Additional check-in meetings are optional.*

As the timeline suggests, there will be a variety of opportunities for the teacher and principal to meet throughout the year to discuss both teacher and student progress. The critical element to the success of the evaluation program is a continuous dialogue focused on teacher improvement and student achievement. The teachers must feel empowered to share their successes and failures, while being supported with resources to aide in growth. While there are due dates in the timeline, these due dates are latest possible dates for the outcomes to occur. Portions of the exercises will be started months before their final due date.

Ideally, this system would culminate near the end of the year with a discussion of progress between the administrator and teacher. However, due to the Senate Bill 7 requirement to develop the Sequence of Honorable Dismissal no less than 75 days prior

to the end of the school year, the evaluation cycle must be complete by March 1 annually. The teacher's progress toward the stated goals would be examined by using the evidence portfolio and the principal's observational data. Students' attainment of achievement goals would be reviewed and discussed to determine if there were unexpected circumstances that resulted in unreliable or incomplete data, and then ultimately whether the students remaining in the pool for evaluation met their targets or not. As described in the state model SLO, 75% of students must achieve their target to receive an "Excellent" rating on the student growth portion. To receive a "Proficient," 50% must achieve their target. Having 25% to 50% of students reach their target will result in a "Needs Improvement." Finally, with less than 25% of students reaching their target, the teacher receives a rating of "Unsatisfactory" for the student growth portion of the evaluation. For year one of implementation, this must be at least 25% of the total summative rating, and 30% in subsequent evaluation years.

The principal would use all of the data available from the observations and evidence submission to determine the final scoring of the teacher's performance on the standards-based rubric. A summative evaluation score would be given based upon the teacher's progress on all aspects of this evaluation process. This summative rating, Excellent, Proficient, Needs Improvement or Unsatisfactory, as outlined in the *Performance Evaluation Reform Act of 2010* and *Senate Bill 7*, would drive retention and sequence of honorable dismissal list group placements.

The Sequence of Honorable Dismissal list replaces the reduction in force (RIF) lists of past years. Now teachers are placed into one of four groups based upon the outcome of their most recent evaluations. Teachers who achieve Excellent on two of

their last three evaluations with no less than a Proficient on the third are placed into Group IV. Teachers who receive at least Proficient on their two most recent evaluations are placed into Group III. Teachers who receive one Needs Improvement or Unsatisfactory on either of their last two evaluations are placed into Group II. Group I is reserved for only those teachers who have yet to receive a summative rating by the deadline for completion of the Sequence of Honorable Dismissal list.

When reductions in staff occur, teachers are dismissed by position from Group I first, then Group II, Group III, and Group IV last. Within each group, teachers are ranked from top to bottom by seniority. Teachers with less seniority are dismissed first, when they are in an equal or lesser group. This is a tremendous change from the days prior to the adoption of Senate Bill 7, when teachers were strictly released or retained in order of seniority, regardless of performance in the classroom. To even further reduce the rights of local bargaining units, teachers in Group I or II who are released do not have recall rights if their position opens later. They *may* be recalled, but are not entitled to positions. Again, this is a swift blow to the educational associations of the State of Illinois.

Summary

Teacher evaluation, dismissal, and recall are all vastly changed by the enacting of PERA and Senate Bill 7. These bills significantly reduce the rights of local bargaining units (unions). However, the intention is not to get rid of effective teachers. The intention is to foster improvement of instruction and increased student achievement. The laws seek to hold teachers accountable to these standards, and demand performance that advances both goals.

What should not be missed in the implementation of these laws in Illinois School Districts is the opportunity to create an evaluation system that supports the teachers. Their rights needed to be diminished, as too many inept teachers were doing far too little for their students but remaining employed because of strict regulations protecting their rights to employment. With these rights diminished, administrators across Illinois must be resources for teachers who want to serve their students and help them achieve to their greatest potential.

The techniques addressed in this section can be assembled to provide a valid, reliable appraisal program. However, as the research has indicated, it takes a variety of the techniques coupled together, which will result in great deal of time spent by both the teacher and the administrator in this process. This is a reasonable feature of the program, as student achievement is the ultimate goal of education and the evaluation process.

CHAPTER VI

IMPLICATIONS

The goal of the evaluation system is, ultimately, to increase student achievement. In seeking to do so, the evaluation system will have far reaching implications for the principal who evaluates the teaching staff. To effectively discuss the methods for navigating these complex hurdles, it is necessary to discuss some relevant theories on instructional leadership. This chapter will discuss Heifetz's (1994) practice of "adaptive leadership," Bolman and Deal's (2003) notion of a multiple lens perspective, and Schlechty's (2005) work on six critical systems. After providing a foundation for understanding each of these authors' leadership perspectives, the implementation implications of the *Performance Evaluation Reform Act of 2010* and *Senate Bill 7* for the areas of human resources, finances, legal liability, collective bargaining, culture, and student learning will be examined.

Relevant Theories

Heifetz—Adaptive Leadership

To understand Heifetz's foundational practices and perspectives, it is important to understand what sort of problem he is discussing. The author suggests that there are significant differences between technical problems and adaptive challenges. In a school district, a technical problem may be that a bus has a flat tire or a staff member is ill for a couple of days. There are relatively simple, tried and true methods for addressing these examples. A back-up bus should be ready to be sent out immediately, and substitutes can

be employed who can adequately cover classes.

Adaptive challenges are altogether different types of problems. Heifetz states:

> Adaptive work consists of the learning required to address conflicts in the values people hold, or to diminish the gap between the values people stand for and the reality they face. Adaptive work requires a change in values, beliefs, or behavior. (p. 22)

Heifetz suggests the following steps when faced with an adaptive challenge: identify the adaptive challenge and focus the attention on its specific issues, manage the stress by looking at the deeper issues in a strategic manner on a calculated pace, maintain focus on the adaptive challenge itself and not the symptoms, and place the ownership of the challenge on the constituents (p. 99).

Bolman and Deal—Multiple Lens Analysis

The work of Bolman and Deal, *Reframing Organizations: Artistry, Choice and Leadership* (2003), suggests that leaders must view their organizations through a variety of lenses in order to have a full and complete understanding of their operation. They describe the process in this manner:

> Reframing requires the ability to understand and use multiple perspectives, to think about the same thing in more than one way. We introduce four frames: structural, human resource, political and symbolic. Each is distinctive, each coherent and powerful in its own right. Together, they help capture a comprehensive picture of what's wrong and what might be done. (p. 5)

In its simplest form, the structural frame cries for rigidity: policies, procedures, memorandums, contracts, etc. In school districts, policies and procedures start with the federal government, are mandated by state regulation and local board policies, and finally are put into practice in building-level handbooks. While they are far from a blueprint for how staff members function every minute of every day, they attempt to identify the role

that each employee is responsible for to ensure that the ultimate goal of the organization is reached: children learn.

Six basic assumptions provide the foundation for the structural frame:

- Organizations exist to achieve established goals and objectives.
- Organizations increase efficiency and enhance performance through specialization and a clear division of labor.
- Appropriate forms of coordination and control ensure that diverse efforts of individual and units mesh.
- Organizations work best when rationality prevails over personal preferences and extraneous pressures.
- Structures must be designed to fit and organization's circumstances (including it goals, technology, workforce and environment.
- Problems and performance gaps arise from structural deficiencies and can be remedied through analysis and restructuring. (p. 45)

The human resource frame is completely about the development of people for the betterment of the organization. It focuses on the relationship that individuals have, not only between themselves and others, but with the organization as whole. People are generally most productive when they are satisfied in the work they are doing.

The human resource frame has a foundation based upon four basic premises:

- Organizations exist to serve human needs rather than the revers.
- People and organizations need each other. Organizations need ideas, energy, and talent; people need careers, salaries, and opportunities.
- When the fit between individual and system is poor, one or both suffer. Individuals are exploited or exploit the organization—or both become victims.
- A good fit benefits both. Individuals find meaningful and satisfying work, and organizations get the talent and energy they need to succeed. (p. 115)

Politics are everywhere. Pressure to make a decision for the interest of certain groups is not unique to education, but it certainly is a focal point. There are stakeholders in every corner of a school district, but, unfortunately, they all have different back-grounds, ideals and, most importantly, interests. In the analysis of the political frame, the number one interest of each stakeholder in the organization is the ability to gain power.

The political frame recognizes that organizations are more than just individuals. There are five pieces to the foundation of the political frame:

- Organizations are *coalitions* of diverse individuals and interest groups.
- There are *enduring differences* among coalition members in values, beliefs, in formation, interest, and perceptions of reality.
- Most important decisions involve allocating *scarce resources*—who gets what.
- Scarce resources and enduring differences make *conflict* central to organization dynamics and underline *power* as the most important asst.
- Goals and decisions emerge from *bargaining*, *negotiation*, and *jockeying for position* among competing stakeholders. (p. 186)

The final frame is called the symbolic or cultural frame. It deals with making meaning from the world around us, not simply being a part of it. The following quote from Schein (2004), who shared a similar perspective on culture, is presented to define culture:

> A pattern of shared basic assumptions that was learned by a group, as it solved its problems of external adaptation and intern integration, that has worked well enough to be considered valid and, therefore, to be taught to new members as the correct way to perceive, think, and feel in relation to those problems. (p. 17)

I believe that Schein has very accurately defined culture, especially in educational organizations. It is not only what the people in an organization deem to be the accurate way to behave, but so much so, that they pass it on to the new people that enter the organization. Culture is not espoused by some handbook or journal or, for that matter, even through formal discussion of such. Instead, it is the atmosphere that is portrayed and the set of values that is shared and accepted. If newcomers do not accept the culture, or are unable to adapt the culture to fit them, they are ultimately outcast from the organization.

Bolman and Deal (2003) suggest that meaning is much more important that actual events when discussing the symbolic framework. Meaning elicits responses when

decisions take on symbolic meaning. It is important to understand what meaning people will take from decisions before making them. The symbolic frame is comprised of five foundational beliefs:

- What is most important is not what happens but what it means.
- Activity and meaning are loosely coupled; events have multiple meanings because people interpret experience differently.
- In the face of widespread uncertainty and ambiguity, people create symbols to resolve confusion, increase predictability, find direction, and anchor hope and faith.
- Many events and processes are more important for what is expressed than what is produced. They form a cultural tapestry of secular myths, heroes and heroines, rituals, ceremonies, and stories that help people find purpose and passion in their personal and work lives.
- Culture is the glue that holds an organization together and unites people around shared values and beliefs.

Schlechty—Six Critical Systems

As is well documented in previous sections of this study, school reform efforts in Illinois and across the country are at an all-time high. School leaders are held accountable for their school data, including student achievement. While evaluation and tenure laws have changed, there still remains little opportunity for administrators and local boards of education to massively change staffs. They are forced to find ways to change the manner in which the organization operates from within.

Schlechty (2005) outlines six systems that are critical to address when attempting fundamental change in a school. They include,

> (1) The way new members are recruited and inducted, (2) the way knowledge is transmitted, (3) the way power and authority are distributed, (4) the way people and programs are evaluated, (5) the way directions and goals are set, and (6) the way boundaries that determine who is inside and who is outside the school are defined. (p. 65)

Any instructional renewal process needs to address all six systems in order to provide lasting improvement in the educational system.

Implications for Human Resources

The human resource department is the primary area responsible for the administration of teacher evaluation. In a large district, it will be the Director of Human Resources, or a similarly titled administrator's job to put a process in place to develop an evaluation system that is effective and meets the guidelines of the new laws. In a small district, it will likely be the superintendent who is responsible for human resources. According to Webb and Norton (2003), "A sound evaluation system can be characterized as one that has (1) established performance evaluation as a school district priority, (2) determined and disseminated clearly articulated evaluation purposes, and (3) adopted an evaluation plan that has a sound methodology" (p. 363).

Heifetz would clearly define the challenge of implementing the new Illinois evaluation system in an effective manner as an adaptive challenge. Illinois school districts are going to have to learn, grow, and change core values about what meaningful teaching looks like, how it is measured and what happens when it is ineffective. The inclusion of student growth and performance to the evaluation process will cause district teams to need to work together to examine current assessment practices, or even develop new ones that will provide a fair and accurate picture of student performance. Creating a new framework that identifies best practices will also need to be a part of the process. According to Danielson (1996), "Because teaching is complex, it is helpful to have a road map through the territory, structured around a shared understanding of teaching" (p. 2). This process is a prime example of what Heifetz suggests in his steps for dealing with adaptive challenges: mobilizing those affected to deal with the problem.

Bolman and Deal would have the school leader examine these new laws and requirements through the lens of each of their four frames. The frame analysis will provide certainty that implementation takes place in a manner that will be effective and sustainable while accounting for all potential pitfalls. The structural frame, with its roots in law and policy, would direct the administration to a careful analysis of the laws and their implications. Careful attention should be given to ensure that a district policy is drafted that meets all applicable timelines and requirements.

New policies will need to clearly document how the summative evaluation rating will be translated into a group for purposes of reduction in force (RIF). Prior to Senate Bill 7, teachers were RIF'd strictly in order of seniority. To comply with the change in the law, local policy should outline that RIF now occurs by first listing staff members by position and then grouping them based on summative evaluation score. Seniority dictates the order within the groups. Group 1 contains those non-tenured teachers who have not yet completed a summative evaluation. Group 2 contains any teacher who receives an "unsatisfactory" or "needs improvement" as their summative rating. Any teacher whose summative rating is "proficient" is categorized in Group 3. Group 4 contains only those teachers whose summative ratings are "excellent" in each of the two most recent evaluations.

The human resource frame would turn administrative attention to the people involved. The teachers will need to adapt to a completely different system of evaluation than what they have been accustomed to, some of them for decades. The administration would be wise to develop a clear communication plan that addresses the questions, concerns, and fears of teachers and assists them in understanding that this process will

actually be a benefit to good teachers. The goal of the new system will be to better identify and document good teaching and reward those teachers with enhanced protection, regardless of seniority. To help understand and address the teachers, a committee would be helpful in assisting to identify the issues and serve as a liaison between administration and the staff.

A factor that could have serious implications for student growth will be other people taking part in a teacher's class. Student teachers, teaching assistants, special education teachers, speech therapists, and other adults play an active role in classroom instruction. When the regular education teacher is evaluated on the students' growth in the class, it is important that there is a mechanism in place to document the amount of control other people have played in the different students' learning. The administrator and teacher should be able to determine which students and classes the teacher ultimately has enough time with to make a connection with the student data. For any student or class that the teacher does not have a high percentage of control over, the instruction should not be used in that teacher's student growth calculation.

The political frame is highly active in the passage of the new evaluation laws. The legislators that have passed these laws were not only seeking to access Race to the Top funds, they were seeking to improve student achievement in Illinois. The public is now watching to determine how schools respond to this challenge. It is critical that educators use this opportunity to provide data that suggests teaching is at the highest level it has ever been. It would be a political detriment to the organization for school districts to combat these mandates and attempt to circumvent these requirements, or present incomplete or false data.

Times of drastic change affect policies, people, and politics and, even more importantly, how meaning is determined. Viewing this issue through the symbolic frame would help the administration understand that teachers will be seeking to identify meaning amongst the chaos that will ensue in making all of the necessary changes. They will be asking questions in order to identify the true motivation of the administration. Are they trying to get rid of some of us? Are they trying to take advantage of us? Are they trying to get more out of us for less? These are some of the questions that teachers will be asking themselves and each other while, at the same time, attempting to identify the underlying meaning behind all of the changes for their learning organization. The committee that was suggested in the human resource frame will also be beneficial in addressing these questions. While they will not surface quite so apparently "at the table," they will be an undercurrent. It will be necessary to identify them and attempt to respond through appropriate action. Your words will only carry you so far. It will be important to demonstrate a respectful and caring nature as you craft these policies and procedures.

The symbolic frame is also referred to as the cultural frame. The culture of an organization equates to the heartbeat. The tempo of the organization is permeated through the walls and into each person that walks through the building. The culture is significantly impacted by the membership of the organization.

As principals and superintendents hire new staff members, they either change the culture or assimilate into the existing one. Marzano (2003) defines the culture of an organization as the "collegiality and professionalism" (pp. 60-67) that exists among the staff. As the administrator responsible for a building, it is critical to be continuously cognizant of the culture or climate to best be able to affect it. For example, when

negativity is pervasive in the environment, the leadership must work to foster a positive environment. This can be accomplished through staff activities, but also must be addressed when hiring new staff members.

Without question, the human resource is the most valuable in the organization. A negative climate will drive the best of the assets to other buildings or districts, thereby further damaging the culture of the organization. Fostering a climate of collegiality and professionalism, however, will result in higher levels of job satisfaction and will draw other quality candidates to the organization.

One of the six critical systems that Schlechty identified as important for addressing when going through organizational change is the evaluation system. Schlechty argued that there are two similar, yet very different, forces at play in the evaluation process;

> Looked at as a technical issue, the creation of standards and ways of assessing standards has mainly to do with validity and reliability. Looked at as a moral and aesthetic issue, evaluation has to do with believability and credibility. If men and women are to be persuaded to act on the results of evaluations that are conducted, those evaluations must be of moral authority as well as technical or rational authority; they must be credible and believable as well as valid and reliable. (p. 133)

Schlechty's observation here is similar to Bolman and Deal's. He has identified that while there is one issue, evaluation, it is rooted in two distinct realms, the technical (structural frame) and moral (human resource). It will be necessary to address issues of validity and reliability for the instrument to be usable for personnel decisions. To foster individual improvement, the tool will need to be credible and believable for the staff members it is administered to. For either of these to be ignored, or unsuccessfully addressed, will result in failure of one of the main objectives of the evaluation system.

Another of Schlechty's six systems that falls into the human resource area is the recruitment and induction system. According to Schlechty, school leaders can have a drastic effect on the success of the organization by developing a system that identifies talented individuals, brings them into the organization, and commits the proper support for their continued learning (p. 67). This process will result in employees who are highly motivated and very loyal to the organization, resulting in low turnover and continuous improvement.

Superintendents and principals need to be active in the recruitment and induction processes. While this is one of the most important aspects of school leadership, it is often one that is neglected. Administrators must remember that talented people do not always come looking for you. Many districts are fortunate to have an abundance of candidates for positions, especially in today's economic climate. However, there are always hard-to-fill positions that will not have as deep of a candidate pool as would be expected. It is the administration's job to seek out talent and draw it into the organization.

More important even than finding talented individuals is the process for retaining them. Turnover is a potential detriment to the organization when good people leave. It can cost resources to attract and hire new people, and it is expensive to train someone new. It is also important to provide a great deal of support to new employees in order to ensure that they are properly equipped to do their job and have the support to continue to grow and improve throughout their career.

Another frightening outcome of placing greater importance on the evaluation plan in the hiring, retention and release of staff members is competition among administrators. Most school districts have multiple buildings with a number of administrators evaluating

teachers. The evaluation plan must be highly objective to protect from administrator bias. With the increased potential for influence, administrators could significantly increase their chance of retaining their best teachers over other buildings' best teachers by inflating evaluation scores. Conversely, by deflating evaluation scores, the administrator increases the likelihood that their lowest performing teachers will be subject to RIF. While the goal of the new laws is to give "good" teachers greater protection over "bad" teachers, there must still be precautions built into the system for administrator bias, so teachers across the district are subject to the same measure of "good" or "bad" teacher.

Implications for Finance

The financial impact of the implementation of the new evaluation requirements will be what Heifetz refers to as a technical issue. Some cost will be associated with the development of the new evaluation system, as staff members may need to be paid additional time to serve on task forces and committees. With resources having been cut at the State and Federal level over the past few years, administrators are stretched to their capacity of instructional supervision in many buildings. In order to implement the new requirements, more time will need to be spent engaged in the evaluation process, observing teachers, analyzing data, and reporting results.

Some buildings will have one principal, with no other administrative support, who simply does not have the time to meet the new requirements. In this situation, district office administration will need to determine how to best support the principal, by either adding an assistant or by having some other administrator or teacher leader engage in the evaluation process. In either case, some pretty basic ratios can be examined and conversations between principals and district staff to determine capacity and address

shortcomings. Norton and Kelly (1997) discuss the process for resource allocation as complex analysis but, nonetheless, technical in nature (p. 36). This is not a situation that will require significant learning or changing of culture for the organization.

An examination through Bolman and Deal's structural frame is very similar to Heifetz's suggestion for addressing technical issues. We can look at policies, procedures, past practice, ratios, and even other schools' and districts' procedures to determine how many students and staff members an administrator can reasonably be responsible for. If buildings are found to be beyond capacity, examination of other options within the district is still a superficial process of examining resources and where support can be redistributed or added. Odden and Archibald (2001) examined the process of resource allocation throughout their book. While the process is, again, rather complex, it focused on attention to policies, procedures, laws, mandates, and other structural aspects of the organization.

The process becomes slightly more complex when examining it through the human resource frame. District-level administration must then begin to look at the people involved to determine, not simply from a numbers standpoint, but from a skill and talent aspect, how successful they can be with the new system. These evaluation changes should not be the impetus for replacing administrators, but it certainly could bring to the forefront the issue of an administrator who is lacking in evaluation skill. There could be a financial impact in properly training, or even eventually replacing principals who are unable to get to a level of proficiency with the new system.

The financial impacts of implementing this new evaluation system have a political impact on education statewide. The general assembly's legislation requires copious

amounts of training for anyone engaging in evaluating teachers or principals, but they have failed to provide new funding sources for this training. The Illinois State Board of Education has been forced to secure a vendor to develop and implement this training, but their overall budget was not increased. Of course there are line items dedicated to the new evaluation so that politicians can claim they have funded the system. With no new revenue for the State Board budget, this simply means that they have taken that money from other areas of education. This is the very definition of politics in that legislators have attempted to retain power by skewing the facts in their favor.

The symbolic nature of the financial impact on education from the new evaluation requirements is one of the current times. Everyone in education is being forced to do more with less, while being held to higher standards. These two issues seem to be viewed as two trains headed in opposite directions, directly at each other. Logic would lead one to believe that higher standards and accountability would be coupled with greater resource allocation. The inverse would also be logically expected; with fewer resources, standards and accountability measures would be relaxed. These two trains seemed to be headed for a collision. What will happen when the explosion occurs could be the end of the reform movement? What the result will be is simply a guess.

Schlechty identifies the directional system that must be addressed relative to the financial impact of the new evaluation system. The first goal of the directional system in today's school is to move our educational focus from one of compliance and attendance to one of engagement. Our school system was not originally designed to foster engagement: thought, creativity, and problem-solving. The system was designed for the simple transmission of knowledge and distribution of students along an achievement

continuum. Currently, all students are expected to achieve at high levels.

The directional system causes administration to consider how resources are allocated for the new direction of the organization. The new evaluation can be designed in a manner that fosters engagement rather than knowledge transmission. Resources will need to be provided to support not only the evaluation's development, but to improve staff members based upon its results. This will require careful analysis of resource allocation for professional development to be certain that teachers are supported and allowed to grow in the proper direction: toward engaging instructional practices.

The overall mission of the new legislative packages is to improve schools. While much of the legislation focuses on evaluation, it also includes mechanisms to hire and fire teachers based upon their overall effectiveness. Accountability is at an all-time high in the State of Illinois. However, the best and brightest are frequently the most expensive staff members. Regardless of the salary schedule design—be it a traditional column-and-lane approach, based upon experience and education, or a more modern approach, predicated on other measures of productivity—the better teachers are more expensive.

Many reformers have championed the notion that teachers should not be paid on a salary schedule, getting guaranteed increases each year due to increased experience, coupled with raises for continuing their education. With the addition of student growth data to the evaluation plan, there is a convenient transition available for changing the pay mechanism to align with the evaluation outcome.

Merit-based pay is a concept that has gained traction in recent years. While it has been talked about for decades, only recently have school boards attempted to find ways to quantify teachers in a way that leads to paying them based on their performance. With the

inclusion of student achievement in the evaluation system, coupled with a plan that uses research-based evidence of effectiveness, the design of a merit-based pay system is theoretically possible.

There is another trend in Illinois: less money. General State Aid payments, categorical State Aid, Federal grants, and even property taxes in many areas, are all significantly reduced over the past several years. Districts have reduced staff, increased class size, reduced support staff, and reduced or eliminated supplemental programs, services, and materials.

These two principals, greater performance requirements and decreased resources, may be the greatest conflict ever to face Illinois. As I discuss this premise with constituents, I liken it to two freight trains rapidly approaching each other on the same track. Each locomotive is rocketing down the track toward each other—one side expecting more, more, more, while the other side is demanding less, less, less. The result of these two freight trains is going to be catastrophic. The question that remains is: Is that the intent? Is the state, or federal government for that matter, setting up this collision to arrive at a point in time when the nation will cry out to fundamentally change how schools operate? At that point, schools could be privatized, vouchers could be ushered in, or the mechanism for funding could be completely re-written.

Marzano (2003) ended his book with a quote from Charles Dickens: "…it was the age of wisdom, it was the age of foolishness" (p. 178). He argued that the future of education will be "the best of times" or "the worst of times" (p. 178), also quoted from Dickens, depending on how educators and policy-makers use the data that we have available to us. An infinite amount of data is available to anyone interested in reviewing

it, not only on individual schools, but also on the state and national education system as a whole. If we do not use this information to continually improve our craft and to identify what the actual world-problems that are being manifested in the schools are, this train collision will be epic and catastrophic.

Implications for Legal Liability

Implementation of the new teacher evaluation laws will have a significant impact on a school district's legal liability. One of the first items that will change in the evaluation process is the assignment of summative ratings. According to Braun (2010), beginning in September of 2012, the categories of "excellent, proficient, needs improvement or unsatisfactory" must be used as the only accepted summative evaluation rankings (p. 209). It is important to note these new categories as they will be the basis for filling vacant positions and releasing teachers under the reduction in force (RIF) provisions.

As is the case when any significant change in law occurs, the final application of the law will be continually defined through court cases challenging many aspects of it. There is little confusion for school administration on whether this is a technical concern or adaptive challenge. Without question, it is the latter.

The simple passage and requirement for change opens up every school district in Illinois to greater legal liability. To protect their districts' interests, administrators must engage in complex processes of determining what is believed to be legal and defendable. Not only should this information be obtained from a number of legal sources, such as the district's retainer firm, school board association legal teams, and school administration association legal teams, but also by working closely with the local education association. As Heifetz suggests in his discussion on management of adaptive challenges, it is

important to put the responsibility for managing the changes on the shoulders of those who it affects; in this case, the shoulders of the teachers. The process must be facilitated in order to maximize universal involvement. This will reduce potential litigation from misunderstanding or misinterpretation of policies.

The Bolman and Deal structural issue is in the drafting of new evaluation policies. This frame is well accounted for in the process suggested in the previous paragraphs of engaging in a complete renewal process with the teachers. This process will need to begin immediately, as many of the requirements, such as the four new categories of summative evaluation, must be implemented in the fall of 2012. The outcome will be a new set of procedures for evaluation. If done cooperatively with the teachers, whether part of a collective bargaining unit or not, the resulting policies should be understood by all and implemented in a manner that will limit the risk of litigation. However, it is important that administration is careful to capitalize on all of the new rights that are being granted to them under the new law, while effectively cooperating with the teachers.

The human resource frame calls for the further development of relationships. While the goal of every administrator should be to have an excellent working relationship with the teachers, the legal liability issues that are now present make it critical to have trust in each other. Both sides must feel that while each will work to protect their own interests, the ultimate goal is to advance student achievement. If teachers feel the administration's true motive is to remove some of them, or in any other way take advantage of the situation, litigation is likely to occur as soon as implementation of the policy begins. However, if the teachers feel the administrative purpose is solely focused on what is best for students, they are much less likely to initiate legal action, unless there

is a clear violation of the law and policies.

Litigation over the new teacher evaluation laws will be the center of political debate for years to come. While it is the sworn duties of judges to uphold the law, they are human. They will be influenced by political pressures, on both sides, to make decisions that will impact how the laws are interpreted going forward.

Judges have the ultimate power in this situation. As litigation begins from teachers who are released, or are the object of other adverse action as a result of the new laws, the courts will decide how to interpret the language of the laws. They could weaken the impact of these laws to the point that they are non-effective. Likewise, they could strengthen the position of districts stalwartly, so that districts will see little risk of litigation and boldly evaluate and make personnel decisions. Regardless of the outcome, political pressure will be brought to bear on judges to favor the interested party's side.

The symbolic meaning of the legal aspects of the new laws could be many. The reform movement has had a steady increase in intensity going back to 1983 and *A Nation at Risk*. The passage of these laws in Illinois, and likely many other states in the coming months and years, signals a climax in the reform era. With the level of accountability and data analysis that will now be present, reformists should be able to begin to discern what the real state of education is and what the sources of both successes and failures are. A great deal of meaning should be gained, which will likely drive the next wave of reform.

Schlechty's boundary system is the aspect that must be addressed to protect the district's legal liability in the implementation of the new evaluation requirements. The boundary system is the complex balance that waivers back and forth as members of the organization attempt to control who can do what. This dynamic is best managed by an

effective effort, by both administration and staff, to continuously discuss and collaborate.

Many school superintendents use a monthly meeting with union leadership to address potential areas of concern before they get to a formal conflict. As has been presented previously, the formation of a teacher committee to be in continuous conversation with administration over the implementation process will serve a similar purpose. As Schein (2006) stated, "To function as a group, the individuals who come together must establish a system of communication and a language that permits interpretation of what is going on" (p. 111). The implementation of this committee and the process for carrying out further dialogue will most certainly assist in addressing "what is going on."

Implications for Collective Bargaining

Web and Norton define collective bargaining as, "the process whereby matters of employee relations are determined mutually by representatives of employee groups and their employer, within the limits of law or mutual agreement" (p. 212). This definition aids in determining whether, in Heifetz's eyes, the implementation of the new law would be a technical problem or an adaptive challenge. His latest work on the issue, which is designed to aid in the application of processes to make progress on adaptive challenges states, "Adaptive challenges can only be addressed through changes in people's priorities, beliefs, habits, and loyalties" (Heifetz, 2009, p. 19). Implementation of the new evaluation processes will be subject to collective bargaining and will require changes in beliefs, habits, and loyalties—an adaptive challenge.

Prior to bargaining the new evaluation language into the negotiated agreement, the formation of the teacher committee discussed previously would be wise. The formal

bargaining process could be shortened and improved if there is mutual understanding by both sides on the legal limitations and goals of the new process. Reeves (2009) discussed a concept of "weeding the garden" (p. 13) with the addition of new initiatives. Before we can begin to implement the new evaluation processes, we must pull out the old ones that are no longer legal or effective. The committee work ahead of bargaining can attempt to accomplish this in order to clear room for change.

Bolman and Deal's structural frame guides much of the collective bargaining aspect of the implementation of the new evaluation processes. This law is very prescriptive in directing the State Board and local districts to implement change in their current practices. There are also many policies, procedures, and laws that govern the collective bargaining process, as well. Many collective bargaining agreements are in place currently and will not be open for bargaining for another year or more. Until their completion, we must use the structural frame, the laws related to collective bargaining agreements, to determine what evaluation laws the district is subject to.

The human resource frame is often neglected at the collective bargaining table. Teachers and board members are often times friends in the community they both live in. However, they enter the bargaining room and many of those relationships are abandoned. The two sides interact in a much more formal manner, careful to keep information and goals tightly guarded. This is the staple of adversarial bargaining.

Marzano, Waters, and McNulty (2005) identify 21 aspects of leadership that are critical for an administrator to develop in order to be successful. One of these traits is referred to as "input" (p. 51). This technique involves seeking input from constituencies in the development and implementation of policies. The human resource frame presented

by Bolman and Deal suggests that input would be an important aspect of the policy formation.

To apply input to the collective bargaining process, the board/administration may attempt to transition from an adversarial style of bargaining to a mutual gains or win-win style of bargaining. This would allow the members of both sides of the bargaining process to engage in collaborative work with the focus on implementing the new legal changes in a meaningful and fair manner.

The politics of collective bargaining most likely remain from past bargaining experiences. To address the political frame, it is important to get a full understanding of the politics that exist between the union and board. The previous negotiations most likely had a political dynamic that demonstrated an agenda from both sides. Whether the agendas were authentic and aligned would assist in understanding what politics will be present as the bargaining room is entered to put into effect the changes in evaluation law.

When the collective bargaining process is on-going and long after its completion, both sides will attempt to find meaning in the new language. In the symbolic frame, the meaning is not simply what the language in the new contracts literally means, as that would be a characteristic of the structural frame. The meaning sought in the cultural frame is centered on what symbolism stems from the creation of the new policy.

To address the symbolism and associations with the new policy, administration should go to the collective bargaining table with a picture of what the new processes represent and try to demonstrate that image throughout the bargaining process. Teachers will have strong reactions to the policies when they begin to see them. They need to be given a picture from their bargaining representatives of the board/administration's

engagement in discussion and purposeful process that will foster student growth. It cannot be perceived as a struggle for power.

If past bargaining sessions have been adversarial and left a sense of distrust between the board and union, the site, format, and time should be changed in order to change the fears and assumptions that will pervade the staff and further the mistrust. Finding a comfortable room with a warm and inviting atmosphere and setting boundaries for bargaining times that honor family commitments will help to elicit a sense of caring and concern. If that tone is shared by the bargaining team with the full membership of teachers, the pictures and symbolism associated with the new contract will be of a board and administration that worked to implement the new legal requirements, but honored and respected their staff in the process.

The implementation of the laws will be discussed at the bargaining table, and trust will be an important element in that process. However, it will be a critical factor in the implementation of the new evaluation system between evaluators and staff members.

The new evaluation process will be more in-depth than it has ever been. To effectively assess the level of proficiency a staff member has in the Danielson model (or other district-selected model), the principal and teacher will have to engage in a process of open dialogue throughout the year. The process cannot be effectively completed in one week. This change in dynamic can actually have a positive influence on the level of trust between evaluator and teacher.

The new process begins with an in-depth discussion of the expectations, goals, and types of observations that will occur throughout the year. As informal observations occur, documentation occurs each time, and discussion is held at that time, to help

improve the staff member, or note current level of proficiency. This open dialogue process will help to provide a common understanding that both sides are working toward the same goal: improved instruction for greater student achievement. With an increased amount of data to support conclusions, teachers should have a higher level of trust that the administration is removing bias and focusing only on effectiveness.

As was the case with legal liability, Schlechty would argue that the bargaining process for implementing the new evaluation standards would fall under the boundary system. Bargaining the new evaluation system the first time will be an extensive exercise in boundary warfare. Both sides will be careful not to give up more than they absolutely have to, as it will represent a movement in the boundary line toward themselves.

The administration and board will need to be well-equipped with the full details of the laws and best practices for effective, reliable, valid, and fair performance appraisal to defend their position at the bargaining table. Although the term warfare was used, it is not meant that this must be completely adversarial. As documented in the Bolman and Deal symbolic section, the meaning that is made from the bargaining attitude will be important. Nonetheless, the district cannot afford to frivolously give back powers and authority to make personnel decisions that were recently granted to them by the Illinois General Assembly.

Implications for the Development of Culture and Climate

Bolman and Deal have one of their four frames completely devoted to understanding the impact of culture and climate on the organization as a whole, which documents the importance of this area of leadership. Their application of this frame has been covered in all of the other sections of this paper and will be woven throughout this

section. However, the focus of this part of the discussion will be on the implications of Heifetz' and Schlechty's theories on the development of culture.

Leading a process of vision/mission/goal development is a key role of the superintendent with the board of education. These statements become the banners that wave in the hallways. They become symbolic representations of the organization's purpose for being. Engaging the board in this process and disseminating it throughout the community is an example of Heifetz's definition of adaptive leadership.

As stated throughout the Bolman and Deal symbolic frame discussions in this paper, it has been noted that the underlying meaning or symbolism is what defines the culture of the organization. Using the board work with the staff allows the stakeholders to become active, engaged participants in developing the culture. Rather than the administration attempting to create it themselves, Heifetz would argue that the staff should drive its development. By using the board vision/mission/goal statements as a starting point, the culture of the organization will reflect the leadership's direction, yet will be the voice of the people. In developing the new evaluation system, it is critical that not just the laws and policies required are foundational, but also the vision, mission and goals of the board and community at large.

Schlechty would place the development of culture and symbolism into the knowledge transmission system. Simply put, this is the system that integrates how a school district as an organization interacts with the individuals who are a part of it. The traditions, ceremonies, and sacred values of the organization are what are passed through the knowledge transmission system.

A culture and climate are developed through the knowledge transmission system of all of the participants. The way new employees and guests are treated, what types and in what manner celebrations occur, and to what extent change is accepted all document the culture of the organization and are immortalized through the knowledge transmission system. Drath (2001) posited,

> Accomplishing the leadership tasks in such an organization can happen only to the extent that all the ways in which leadership is understood to happen are honored. Leadership, in this view, is far from being something that a person can offer independently, simply as an individual, and is seen as a complex construction of multiple levels of meaning. (p. 155)

An effective school leader must be an active participant in the knowledge transmission system to have a positive effect on the culture. As evidenced in the Drath quote, one cannot impose a "complex construction of multiple levels of meaning" on the participants of the organization. This can only be done when this aspect of leadership is understood to be a function of everyone in the organization.

In the development of the new evaluation system, it is critical to begin with very candid discussions with the teaching staff. They need to understand why this change is happening, what the new direction will be, and how it will help promote the vision of the organization. They need to have communication flowing to them on many occasions, such as faculty meetings, emails, and small group work sessions. As dialogue occurs, meanings are made and the symbolic nature of evaluation will develop as an opportunity to continually improve the organization.

Implications for Student Learning

Student learning is not reserved for the last implication in this study because it is of the least importance. In fact, it should always be the first and last question

administrators ask themselves as they deal with any issue. “What is best for kids?” is a very simple question, but the most important question that any administrator considers. It must continually be revisited as daily business is done and when any instructional renewal process is undertaken.

The first objective of the new evaluation system in every Illinois school district should be to increase student learning. Deciphering how that can be accomplished is, no doubt, an adaptive challenge. Administrators and teachers are going to need to spend a great deal of time identifying what types of instructional practices are best. The choice to use Daniel’s (1996) model may make identifying the behaviors themselves a technical problem. Putting the entire evaluation system together, however, is going to take adaptive work.

Putting students first in the structural frame of Bolman and Deal means clearly articulating the goals and objectives of the evaluation system to document that student learning is the first goal. This frame would require detailed rubrics that provide clarity as to what proficiency, and the lack thereof, looks like with a variety of teaching behaviors, as Daniels has done. The human resource frame would turn the administration’s attention to students as people. The new teacher appraisal system should encompass what students, as people, require, in order to learn and achieve at high levels.

The political frame will influence student achievement in the development of the new evaluation system as political forces battle over how much responsibility can be placed on a teacher for the students’ learning. Teachers will argue than many of the factors affecting student achievement are out of their control and will attempt to control the measurement of and analyzation of student achievement data. Administration, on the

other hand, understands that good teaching usually results in high levels of student achievement and will work to put as much weight on this metric as possible.

The symbolic nature of student achievement relative to the new evaluation system is one of servant leadership. Greenleaf (1977) presented questions that can be answered to determine if leadership is servant-based. "The best test, and difficult to administer, is this: Do those served grow as persons? Do they, while being served, become healthier, wiser, freer, more autonomous, more likely themselves to become servants?" (p. 27). If all members of the organization are responsible for the creation of culture, for students to achieve at high levels, everyone should be focused on serving the students and, as Greenleaf suggests, ensuring that all of his criteria are met for each and every child. The evaluation system should assess to what extent teachers are leading their students with a servant mentality.

Schlechty would place student learning in the directional system. According to Schlechty, the direction of schools must be toward engagement and away from compliance. For the evaluation system to foster engagement, it must be built from a foundation, Danielson's for one example, of a standards-based rubric that relies on engagement, rather than compliance. If a teacher can get a proficient, or even excellent rating without ever challenging their students to levels of engagement, the evaluation system is flawed and not serving the learning needs of the students under the teachers' direction.

Implications for Further Research

This study has been formulated as a predecision-making policy analysis due to its timing compared to implementation of student growth in the teacher evaluation plans in

Illinois. School districts will still have more than one full year to research, discuss, and adopt policies that will govern how student growth is incorporated into the evaluation plan.

This study has stated and documented how important the teacher is in relation to student learning. They are the key element to the recipe. Based upon that assumption, I have suggested that adopting research-based methods of evaluation, coupled with an effective plan for including student growth measures, a school district could expect to ultimately increase student achievement by 10-20%.

Further research should be done on the evaluation outcomes of all school districts to determine if student achievement does, in fact, increase after adopting the recommended evaluation protocols. While a post decision-making policy analysis certainly is one method that could be utilized for this process, a quantitative analysis of variance (ANOVA) may be a particularly useful tool in assessing the change in evaluation score.

Coupled with the ANOVA, a qualitative analysis of the variety of systems affected by the changed evaluation systems is necessary. Local school districts, along with field researchers, should seek to understand if the adaptive challenges discussed in this study have been successfully managed, and staff members' beliefs and principles have truly been changed to embrace the new culture of evaluation. This analysis will require interviews, reviews of policies, reviews of evidence to determine if the system was implemented with integrity and remains valid and reliable.

Summary

This chapter has identified the educational theories from Heifetz, Bolman and Deal, and Schlechty as three of the foundational theories from which to build a leadership perspective on any instructional issue. Teacher evaluation is certainly an important instructional issue. As this chapter has documented, the implementation of a new teacher evaluation system, including student growth and performance, is not an isolated task. It is a complex process that will have dramatic effects on many areas of the organization. An extraordinary attention to these effects will be required to ensure that the ultimate goal of the organization is promoted through the creation of the evaluation system: increased student achievement.

REFERENCES

Attinello, J. R., Lare, D., & Waters, F. (2006). The value of teacher portfolios for evaluation and professional growth. *NASSP Bulletin, 90*(2), 132-152.

Baker, P., Ashby, D., & Rau, W. (1997). *Local response to state mandated reform: A sStudy of the Illinois Public School Accreditation Process*. Illinois State University, Normal, IL.

Baker, P., Bakken, J., Blum, C., Cates, G., Swerdlik, M., Sylvester, B., & Thompson, J. (2007, March). *Developing learning-centered schools with a comprehensive system of student support*. Paper presented at the Fourth International Association for Positive Behavior Support Conference, Boston, MA.

Baratz-Snowden, J. (1993). Assessment of teachers: A view from the National Board for Professional Teaching Standards. *Theory Into Practice, 32*(2), 82-85.

Bill and Melinda Gates Foundation. (2010). *Empowering effective teachers: Strategies for implementing reforms*. Retrieved from http://www.gatesfoundation.org/united-states/Documents/empowering-effective-teachers-empowering-strategy.pdf

Blake, J., Bachman, J., Frys, M. K., Holbert, P., Ivan, T., & Sellitto, P. (1995). A portfolio-based assessment model for teachers: Encouraging professional growth. *NASSP Bulletin, 79*(37), 37-46.

Blankstein, A. M. (2004). *Failure is not an option: Six principles that guide student achievement in high-performing schools*. Thousand Oaks, CA: Corwin Press.

Bolman, L. G., & Deal, T. E. (2003). *Reframing organizations: Artistry, choice and leadership* (3rd Ed.). San Francisco, CA: Jossey-Bass.

Borman, G. D., & Kimball, S. M. (2005). Teacher quality and educational quality: Do teachers with higher standards-based evaluation ratings close student achievement gaps?. *The Elementary School Journal, 106*(1), 3-20.

Braun, B. A. (2010). *Illinois School Law Survey* (11th ed.). Springfield, IL: IASB Publishing.

Braun, H. (2005). *Using student progress to evaluate teachers: A primer on value-added models*. Princeton, NJ: Educational Testing System.

Bryk, A. S., & Schneider, B. L. (2002). *Trust in schools: A core resource for improvement.* New York, NY: Russell Sage Foundation.

Bush, G. W. (1990). *1990 State of the Union Address*. Retrieved from: http://millercenter.org/president/speeches/detail/3423.

Bushaw, W. J., & Lopez, S. J. (2012). Public education in the United States: A nation divided. *Phi Delta Kappan*, *94*(1), 9-25.

Clotfeltter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). *How and why do teacher credentials matter for student achievement?* (NBER Working Paper 12828). Retrieved from National Bureau of Economic Research: http://www.nber.org/papers/w12828.

Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Darling-Hammond, L. (1983). Teacher evaluation in the organizational context: A review of the literature. *Review of Educational Research, 53*(3), 285-328.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives, 8*(1).

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan, 93*(6), 8-15.

Darling-Hammond, L., Wise, A. E., & Peas, S. R. (1983). Teacher evaluation in the organizational context: A review of literature. *Review of Educational Research, 53*(3), 285-328.

Doran, G. T. (1981). There's a S.M.A.R.T. way to write management's goals and objectives. *Management Review, 70*(11), 35-36.

Drath, W. (2001). The deep blue sea: Rethinking the source of leadership. San Francisco, CA: Jossey-Bass.

Dunn, W. N. (1981). *Public policy analysis: An introduction*. Englewood Cliffs, NJ: Prentice-Hall.

Furtwengler, C. B. (1995). State actions for personnel evaluation: Analysis of reform policies, 1983-1992. *Education Policy Analysis Archives, 3*(4). Retrieved from http://epaa.asu.edu/epaa/v3n4.html.

Gallagher, H. A. (2004). Vaughn Elementary's innovative teacher evaluation system: Are teacher evaluation scores related to growth in student achievement? *Peabody Journal of Education, 79*(4), 79-107.

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Retrieved from http://www.tqsource.org/link.php.

Greenleaf, R. K. (1977). *Servant leadership: A journey into the nature of legitimate power and greatness.* Mahwah, NJ: Paulist Press.

Guthrie, J. W., & Springer, M. G. (2004). A nation at risk revisited: Did wrong reasoning result in right results? at what cost? *Peabody Journal of Education, 79*(1), 7-35.

Haefle, D. L. (1980). How to evaluate the teacher—let me count the ways. *Phi Delta Kappan, 61*(5), 349-352.

Halvorson, R., & Clifford, M. (2003, November). *Evaluation in the wild: A distributed cognition perspective on teacher assessment*. Paper presented at the meeting of University Council of Educational Administration, Kansas City, MO.

Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy, 100*(1), 84-117.

Hargreaves, A., & Fullan, M. (2012). *Professional capital*. New York, NY: Teachers College Press, Columbia University.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning.* New York, NY: Routledge.

Hazi, H. M., & Rucinski, D. A. (2009). Teacher evaluation as a policy target for improved student learning: A fifty-state review of statute and regulatory action since NCLB. *Education Policy Analysis Archives, 17*(5), 1-22.

Heifetz, R. A. (1994). *Leadership without easy answers*. Cambridge, MA: The Belknap Press of Harvard University Press.

Heifetz, R. A., Grashow, A., & Linsky, M. (2009). *The practice of adaptive leadership: Tools and tactics for changing your organization and the world*. Boston, MA: Harvard Business Press.

Heneman, H. G., Milanowski, A. T. Kimball, S. M., & Odden, A. (2006). Standards-based teacher evaluation as a foundation for knowledge- and skill-based pay (CPRE Policy Brief RB-45). Philadelphia, PA: Consortium for Policy Research in Education. Retrieved from http://eric.ed.gov/PDFS/ED493116.pdf.

Hoerandner, C. M., & Lemke, R. J. (2006). Can no child left behind close the gaps in pass rates on standardized tests?. *Contemporary Economic Policy, 24*(1), 1-17.

H.R. 1 – 111th Congress: American Recovery and Reinvestment Act of 2009. (2009). Retrieved from http://www.govtrack.us/congress/bills/111/hr1.

Illinois General Assembly. (2010). *Performance Evaluation Reform Act: Laws of Illinois, 2010.* (Illinois Public Act 096-0861). Springfield, IL. Retrieved from http://www.ilga.gov/legislation/fulltext.asp?DocName=09600SB0315ham001&GA=96&SessionId=76&DocTypeId=SB&LegID=41104&DocNum=0315&GAID=10&Session.

Illinois General Assembly. (2011). *School Code-Education Labor Relations-Pension*. (Illinois Public Act 097-0008). Springfield, IL. Retrieved from http://www.ilga.gov/legislation/BillStatus.asp?DocNum=7&GAID=11&DocTypeID=SB&SessionID=84&GA=97.

Illinois State Board of Education. (1985). *Legislative Package 1985*. Springfield, IL.

Illinois State Board of Education. (2012). www.isbe.net.

Kahlenberg, R. D. (2007). *Tough liberal: Albert Shanker and the battles over schools, unions, race, and democracy*. New York, NY: Columbia University Press.

Kimball, S. M., White, B., Milanowski, A. T., & Borman, G. (2004). Examining the relationship between teacher evaluation and student assessment results in Washoe County. *Peabody Journal of Education, 79*(4), 54-78.

Lewis, A. (1982). *Evaluating educational personnel*. Arlington, VA: American Association of School Administrators.

Marzano, R. J. (2003). *What works in schools: Translating research into action.* Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. J., Waters, T., & McNulty, B. A. (2005). *School leadership that works: From research to results*. Alexandria, VA: Association for Supervision and Curriculum Development.

Matsumura, L. C., Garnier, H., Pascal, J., & Valdes, R. (2002). Measuring instructional quality in accountability systems: Classroom assignments and student achievement. *Educational Assessment, 8*(3), 207-229.

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003) *Evaluating value-added models for teacher accountability*. Santa Monica, CA: RAND Corporation.

McGreal, T. L. (1982, March). *Developing a teacher evaluation system: Commonalities of those systems that function most effectively*. Paper presented at the annual meeting of the Association for Supervision and Curriculum Development, Anaheim, CA.

McGreal, T. L., Broderick, E., & Jones, J. (1984). Artifact collection. *Educational Leadership, 41*(7), 20-21.

Medley, D. M. (1982). *Teacher competency testing and the teacher educator*. Charlottesville, VA: Bureau of Educational Research.

Milanowski, A. T. (2004). The relationship between teacher performance and student achievement: Evidence from Cincinnati. *Peabody Journal of Education, 79*(4), 33-53.

Milanowski, A. T., Kimball, S. M., & White, B. (2004). *The relationship between standards-based teacher evaluation scores and student achievement: Replication and extension at three sites*. (CPRE Working Paper TC-04-01). Retrieved from University of Wisconsin-Madison, Consortium for Policy Research in Education: http://www.cpre.wceruw.org/papers/3site_long_TE_SA_AERA04TE.pdf

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: U.S. Government Printing Office.

National Comprehensive Center for Teacher Quality. (2009). *The American recovery and reinvestment act: Opportunities and strategies to advance teacher effectiveness, part 2*. Washington, DC: Author.

Newton, X., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Educational Policy Analysis Archives, 18*(23), Retrieved from http://redalyc.uaemex.mx/redalyc/pdf/2750/275019712023.pdf.

Nichols, S. L., & Berliner, D. C. (2007). *Collateral damage.* Cambridge, MA: Harvard Education Press.

No Child Left Behind Act, 20, U.S.C. (2002).

Norton, M. S., & Kelly, L. K. (1997). *Resource Allocation: Managing money and people*. Larchmont, NY: Eye on Education.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects?. *Educational Evaluation and Policy Analysis, 26*(3), 237-257.

Odden, A., Archibald, S. (2001). *Reallocating resources: How to boost student achievement without asking for more*. Thousand Oaks, CA: Corwin Press, Inc.

Odden, A., Borman, G., & Fermanich, M. (2004). Assessing teacher, classroom, and school effects, including fiscal effects. *Peabody Journal of Education, 79*(4), 4-32.

Performance Assessment Advisory Council. (2010). *Principal & teacher evaluation in Illinois: Past, present & future*. Retrieved from: http://www.nga.org/files/live/sites/NGA/files/pdf/1109ESPLIGHTFOOT.PDF

Peterson, K. (2004). Research on school teacher evaluation. *NASSP Bulletin, 88*(639), 60-79.

Peterson, K., & Kauchak, D. (1982). *Teacher evaluation: Perspectives, practices and promises*. Salt Lake City, UT: University of Utah, Center for educational Practice.

Reeves, D. B. (2009). *Leading change in your school*. Alexandria, VA: Association for Supervision and Curriculum Development.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review, 94*(2), 247-252.

Rockoff, J. E., & Speroni, C. (2010). *Subjective and objective evaluations of teacher effectiveness*. New York, NY: Columbia University.

Sappington, N. (2008). Don't ask teachers to teach to the test. *IASB Journal.* Retrieved from: http://www.iasb.com/journal/j111208_06.cfm.

Schein, E. H. (2004). *Organizational Culture and Leadership* (3rd ed.). San Francisco: CA: Jossey-Bass.

Schlechty, P. C. (2005). *Creating great schools: Six critical systems at the heart of educational innovation*. San Francisco, CA: Jossey-Bass.

Scriven, M. (1981). Summative teacher evaluation. In J. Millman (Ed.), Handbook of teacher evaluation (pp. 244-271). Beverly Hills, CA: Sage.

Smith, E., & Gorand, S. (2007). Improving teacher quality: Lessons from America's No Child Left Behind. *Cambridge Journal of Education, 37*(2), 191-206.

Taylor, E. & Tyler, J. (2011). The effect of evaluation on performance: Evidence of longitudinal student achievement data of mid-career teachers. (Working Paper 16877). Retrieved from National Bureau of Economic Research: http://4teachingexcellence.org/uploads/media_items/the-effect-of-evaluation-on-performance.original.pdf

United States Department of Education Task Force on Education for Economic Growth. (1984). *Action in the states: Progress toward educational renewal*. Denver, CO: Education Commission of the States.

United States Department of Education. (2009). *Race to the top program: Executive Summary*. Washington, DC: U.S. Government Printing Office.

Veir, C. A., & Dagley, D. L. (2002). Legal issues in teacher evaluation legislation: A study of state statutory provisions. *Brigham Young University Education & Law Journal, 1*(1), 1-16

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89-122.

Webb, L. D., & Norton, M. S. (2003). *Human resources administration: Personnel issues and needs in education* (4th ed.). Upper Saddle River, NJ: Merrill Prentice Hall.

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*(1), 57-67.

Xu, S., & Sinclair, R. L. (2002, April). *Improving teacher evaluation for increasing student learning*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

APPENDIX

STUDENT LEARNING OBJECTIVE (SLO) STATE TEMPLATE

Performance Evaluation Advisory Council
Model Student Learning Objective Template

This is an example Student Learning Objective (SLO) template that may be adapted to suit the needs of individual school districts. The example SLO template was designed to include guiding questions and statements that are important for both teachers and evaluators to reflect upon throughout the SLO process.

A Student Learning Objective (SLO) is a detailed process used to organize evidence of student growth over a specified period of time. The SLO process is appropriate for use in all grade levels and content areas and establishes meaningful goals aligning curriculum, instruction, and assessment. This template guides teachers and evaluators through a collaborative SLO process. Portions of this template were adapted from the Center for Assessment *SLO Toolkit*. In addition, domains and components that may align with each element of the template are included from the Danielson Group *Framework for Effective Teaching* to support discussion between teachers and evaluators.

Check boxes are included throughout the template to document the initial discussion and approval of each element. Evaluators may include written feedback concerning each element directly into the template using a different font color.

Educator Information

Academic Year	
Educator Name	
School Name	
District Name	

Planning Information

Course/Subject Name	
Brief Course Description	
Grade Level(s)	
Interval of Instruction	

Timeline and Sign-Off

Evaluator Name and Title	
Initial SLO Evaluator Sign-Off	
Midcourse Check-In Sign-Off	
Description of changes made during the Midcourse Check-In:	
Due Date of Final SLO	

Element #1: Learning Goal

A learning goal is a description of what students will be able to do at the end of a specified period of time aligned to appropriate learning standards. The development of a learning goal provides a solid foundation for meaningful, goal directed instruction and assessment. The learning goal encompasses a big idea that integrates multiple content standards.

Domain 1: Planning and Preparation
1a Demonstrating Knowledge of Content and Pedagogy
1c Setting Instructional Outcomes
1e Designing Coherent Instruction

Domain 3: Instruction
3c Engaging Students in Learning

☐ Describe the learning goal.	
☐ What big idea is supported by the learning goal?	
☐ Which content standards are associated with this big idea? *List all standards that apply, including the text of the standards (not just the code).*	
☐ Describe the student population.	
☐ Describe the instruction and strategies you will use to teach this learning goal. *Be specific to the different aspects of the learning goal.*	
☐ Identify the time span for teaching the learning goal (e.g., daily class-45 minutes for the entire school year).	
☐ Explain how this time span is appropriate and sufficient for teaching the learning goal.	

Questions to Guide Discussion

- Why is this learning goal important and meaningful for students to learn?
- In what ways does the learning goal require students to demonstrate deep understanding of the knowledge and skills of the standards or big idea being measured (e.g., cognitive complexity)?

Element #2: Assessments and Scoring

Assessments and evaluation procedures should be used to support and measure the learning goal. Consider how the assessment and evaluation procedures will be used to monitor student growth over multiple points in time in order to inform and differentiate instruction for all students.

Domain 1: Planning and Preparation 1d Demonstrating Knowledge of Resources 1f Designing Student Assessments	**Domain 3: Instruction** 3d Using Assessment in Instruction

☐ Describe the assessments and evaluation procedures (e.g., performance tasks, rubrics, teacher-created tests, portfolios, etc.) that measure students' understanding of the learning goal.	
☐ Describe how the assessments and evaluation procedures may be differentiated to meet the needs of all students described in the student population.	
☐ Explain how student performance is defined and evaluated using the assessments. Include the specific rubric and/or evaluation criteria to be used.	

Questions to Guide Discussion

- How often will you collect data to monitor student progress toward this learning goal?
- How will you use this information to monitor student progress and to differentiate instruction for all students toward this learning goal?

Element #3: Expected Growth Targets

In order to identify expected growth targets, educators must first identify students' actual performance through a review of available data reflecting students' starting points (i.e., baseline) concerning the learning goal. After the expected growth targets are identified, both the teacher and evaluator should reflect on whether the growth targets are ambitious, yet realistic for students to achieve in the specified period of time.

Domain 1: Planning and Preparation
1b Demonstrating Knowledge of Students
1c Setting Instructional Outcomes

☐ Identify the actual performance (e.g., test scores, performance tasks, etc.) to establish starting points (i.e., baseline) for students.	
☐ Using students' starting points (i.e., baseline) identify the number or percentage of students expected at each growth target based on their assessment performance(s) (i.e., expected growth). Be sure to include any appropriate subgroups.	

Questions to Guide Discussion

- Describe the courses, assessments, and/or experiences used to establish starting points and expected outcomes for students' understanding of the learning goal (i.e., baseline data).
- Explain how these expected growth targets demonstrate ambitious, yet realistic goals, for measuring students' understanding of the learning goal.

Element #4: Actual Outcomes

Domain 3: Instruction	**Domain 4: Professional Responsibilities**
3e Demonstrating Flexibility and Responsiveness	4a Reflecting on Teaching 4b Maintaining Accurate Records

☐ Record the actual number or percentage of students who achieved the student growth targets. Be sure to include any appropriate subgroups.	
Please provide any comments you wish to include about the actual outcomes:	

Required for Evaluator

☐ Explain how the actual number or percentage of students who achieved student growth targets translates into an appropriate teacher rating.	

Element #5: Teacher Rating

Unsatisfactory	**Needs Improvement**	**Proficient**	**Excellent**
Less than 25% of Students Met the Indicated Growth Target(s). ☐	25% - 50% of Students Met the Indicated Growth Target(s). ☐	51% - 75% of Students Met the Indicated Growth Target(s). ☐	76% - 100% of Students Met the Indicated Growth Target(s). ☐
Date:	Evaluator Signature:		
Date:	Teacher Signature:		