Improving Teacher Quality Through Merit-Based Evaluation: 
An Analysis of Indiana's New Teacher Appraisal System and its Ability to 
Promote Meaningful Learning in Schools

Jessica M. Kaplan  
Legal Studies  
Senior  
School of Public & Environmental Affairs

Stephen M. Harper  
Assistant General Counsel, Indiana University  
Adjunct Professor  
School of Public & Environmental Affairs  
Faculty Mentor
Abstract

The Obama administration’s 2009 launch of Race to the Top, a competitive grant program aimed at incentivizing states to improve teacher accountability at their schools, marked the start of an unprecedented trend toward the use of merit-based measurements in teacher evaluations. Teachers have long been regarded as the keys to student success, and as such, both education scholars and citizens have come to believe that reform efforts should focus first on improving teacher quality and effectiveness.

With the increased push for accountability, many states have designed new evaluation systems whereby teachers are assessed – and in some states, compensated – on the basis of merit, determined in part by students’ performance on standardized tests. Standardized testing has become increasingly popular in American schools, mainly because it provides conformity across states and districts and is fairly cheap to administer. It does, however, beg the question: If instructors are teaching for a test, are students actually learning?

Credential acquisition and the pursuit of high grades and test scores receive an emphasis equal to, if not more than, the actual acquisition of knowledge in K-12 schools. This thesis will examine how real learning and accomplishment are defined and measured, as well as how that might contrast with credential acquisition. It will then explore the characteristics and elements of Indiana’s new teacher evaluation system under the RISE rubric and evaluate them against established, research-based standards regarding what constitutes – and is likely to promote – meaningful learning in schools. In offering a critique of Indiana’s new teacher evaluation system, this thesis will compare methodologies that scholars suggest promote real learning with methodologies or activities that are incentivized under the new system. Finally, it will make recommendations for possible improvement of Indiana’s new program.
Introduction

In 2009, the Obama administration launched Race to the Top, a competitive grant program aimed at incentivizing states (and later school districts) to improve teacher accountability at their schools. Motivated by the prospect of receiving additional funding for their always-underfunded schools, many states began rewriting their laws and adopting new policies to satisfy the requirements under President Obama’s new program, thus marking the start of an unprecedented trend toward the use of merit-based measurements in teacher evaluations.

The primary reason for Race to the Top was to improve teacher effectiveness, which would in turn benefit students by promoting learning at school. In a 2010 speech, Secretary of Education Arne Duncan noted, “Talent matters tremendously in the classroom. It’s no surprise that the single biggest in-school influence on student academic growth is ... not socioeconomic status, not family background, but the quality of the teacher at the head of the class” (Duncan).

This is a well-known fact, but teacher evaluation reform has still been necessary because research demonstrates that traditional school systems have consistently failed to assess their instructors properly. In 2009, for example, The New Teacher Project released a study discussing teacher evaluation and dismissal in four states and twelve diverse districts revealing that all teachers are essentially rated “good” or “great” (Weisberg, Sexton, Mulhern & Keeling). The report concluded that on a national level, school districts treat all instructors as interchangeable “widgets” rather than professionals, meaning teacher excellence goes unrecognized, and poor performance goes unaddressed. As such,
teachers have little to no incentive to improve, and thus remain ineffective (Weisberg, Sexton, Mulhern & Keeling).

In an effort to fix this core problem and encourage professional development (and thereby promote meaningful student learning), education scholars, policymakers and citizens have come to believe that education reform efforts should focus first on improving teacher effectiveness. Lawmakers and local school administrators have struggled, however, to determine exactly how to measure a teacher’s effectiveness and subsequently with how to create a system whereby only the best and brightest candidates are chosen to instruct our nation’s students.

This thesis first seeks to define and explore how meaningful learning can be measured, with an eye toward teachers’ role in that process. It then explores the best and most effective ways of evaluating teachers based on thorough research of best practices and evaluates the characteristics and elements of Indiana’s new teacher evaluation system under the RISE rubric against standards from the literature to determine whether or not Indiana’s system promotes real learning in K-12 schools. The paper concludes by offering recommendations for the improvement of Indiana’s new program.

I. Defining Meaningful Learning in K-12 Schools

In recent years, society has become increasingly obsessed with quantifying the learning process and documenting tangible, positive outcomes from schooling, despite research demonstrating that such trends are actually detrimental to the very goals of education (Alexander). The pursuit of credits, high grades and test scores receives an emphasis equal to, if not more than, the actual acquisition of knowledge, skills and
competencies in K-12 schools. This is due in part to the fact that educational credentials are so closely related to occupational attainment in American society (Bills). Employers’ needs for skills have not increased as rapidly as their demand for credentials, and as a result, “what is learned in school has much more to do with conventional standards of sociability and propriety than with instrumental and cognitive skills” (Bills).

With the Obama administration’s increased push for teacher accountability, many states have opted to assess teachers in part by their contribution to student learning, which is determined heavily by students’ performance on standardized tests. Although the literature presents no formal consensus about what makes for meaningful learning in K-12 schools, there is certainly evidence to suggest that it involves more than the accumulation of facts, and more than a students’ average achievement gains on a standardized test (Bell, Goe, Little).

In the simplest of terms, meaningful learning is the ability to transfer acquired knowledge and use it in new situations. This means that students must move beyond mere memorization of academic subject content and truly understand concepts in a “flexible, connected and generalized” way (Shepard). Tony Wagner, author of Making the Grade, calls this the process of gaining “competencies.” According to Wagner, once a student has gained competency, he or she can effectively apply information in the search for solutions to a problem, or to create new knowledge (Wagner). It is important to remember, however, that content knowledge and information are still the foundations for meaningful learning, as students need something to research, hypothesize and write about (Wagner).

To engage student interest, school learning should be authentic and connected to the real world. In an era of “instant access and information overload” (Shepard), developing
the ability to master skills, think critically and problem-solve in real-world contexts is essential to defining meaningful student learning in K-12 schools. According to the Organization for Economic Cooperation and Development (OECD), the 2012 results from the Program for International Student Assessment (PISA) showed that American students had significant difficulty performing mathematics tasks with higher cognitive demands, “such as taking real-world situations, translating them into mathematical terms and interpreting mathematical aspects into real-world problems” (Results from PISA 2012: United States). These deeper learning skills are most in demand in today’s global economy, and yet many American students seem to lack them (Simon).

Beyond the essential components of acquiring and transferring knowledge, meaningful learning in modern society is no longer seen as a strictly cognitive phenomenon. As defined by educational psychologist Patricia Alexander, learning consists of “the systematic changes that students should undergo cognitively, motivationally and socio-culturally as they are becoming formally educated – not just for a year or two, but across the lifespan” (Alexander). The inclusion of the latter two components illustrates that in order to engage in meaningful learning, students should feel motivated and will hopefully even demonstrate a corresponding rise in interest in a particular academic domain as they develop intellectually, personally and socially in school (Alexander). Social and cognitive skills are, after all, inextricably related, as cognitive abilities are developed in part through socially supported interactions. For instance, the development of important dispositions, such as a student’s willingness to persist in trying to solve difficult problems, is fostered in part by the expectations and social norms taught in the classroom (Shepard).
The development of other important social skills, such as the art of having a face-to-face conversation, should also take place (at least to some extent) in the classroom during the learning process. According to a Kentucky teacher, whose insightful editorial was published in *The Atlantic* recently, “conversational competence might be the single-most overlooked skill we fail to teach students. Kids spend hours each day engaging with ideas and one another through screens – but rarely do they have an opportunity to truly hone their interpersonal communication skills” (Barnwell). Such skills are critical for students to learn and use during their lifetimes, as they are essential in job interviews, debates, negotiations and every-day-life conversations, among other things (Barnwell).

One frustrating issue that arises when attempting to define meaningful learning in schools is that researchers do not yet have a model that can differentiate between learning outcomes from formal schooling and those that result from other outside factors affecting the learning process (Alexander). Learning is a “complex, multi-dimensional, fallible and dynamic” (Alexander) progression of mental construction and sense making, and it is often difficult to measure it solely from a classroom perspective.

II. Measuring Meaningful Learning in K-12 Schools

With these characteristics and components of what creates meaningful learning in mind, another challenge that educational systems face is finding a method that properly assesses and measures growth in student learning. Schools today test for an abundance of factual information, even though “common sense, as well as countless studies, tell us that facts are quickly forgotten, easily looked up and not what’s most important to learn” (Wagner). Education scholar Ken Robinson argues that this current trend toward
conformity and standardization in all aspects of schooling and assessment – deemed the “production line mentality” of formal education – should actually be reversed (Robinson). Robinson contends that school corporations and teachers should encourage students to engage in creativity and “divergent thinking,” the latter of which he defines as “an essential capacity for creativity ... the ability to see lots of possible answers or ways of interpreting a question” (RISA Animate – Challenging Education Paradigms, Robinson). Researchers agree that new approaches to student assessment are needed but have struggled to map out a smart solution, as assessment still must have a component of standardization for measurement to be fair and equal across students.

The dominance of objective testing in classroom assessment has, in fact, provided many benefits to our educational society. Primarily, it has proven instrumental in shaping our beliefs about “the nature of evidence and principles of fairness” (Shepard) that are essential to any learning culture. Despite calls for the development of alternative ways of measuring student achievement, it is important to remember that objective testing is fundamentally fair. Its conformity, while criticized, allows for students to be assessed and measured equally and consistently. According to education scholar Lorrie Shepard, “any attempt to change the form and purpose of classroom assessment to make it more fundamentally a part of the learning process must acknowledge the power of these enduring and hidden beliefs” (Shepard) of basic fairness.

Decades ago, student assessments comprised of various multiple-choice and matching tests fit closely with what was deemed important to learn (Shepard). Today, however, that is no longer the case. The practice of testing mere subject content is outdated, and schools have instead turned their focus to assessing skills, competencies and
comprehension to gauge whether students are truly learning in the classroom. Recent studies have shown that classroom assessment works best as a dynamic, on-going process that supports and enhances student learning and addresses learning processes, as well as learning outcomes (Shepard). Schools should be concerned with how teachers impact students’ conceptions of, interests in and attitudes toward academic subjects, as well as with how such outcomes impact subsequent student development (Alexander). This emerging “constructivist paradigm” (Shepard) of schooling contends that student assessments should be a source of insight and help, rather than merely a means by which instructors can reward or punish students with a grade.

Because classroom evaluation should be used as an integral part of the learning process, assessment and instruction should match up in both time and purpose (Graue). In other words, assessments should be comprised of content that matches the teacher’s challenging subject matter standards and articulates the skills necessary for students to master a particular academic domain. In order to fully measure meaningful student learning, teachers must also provide ongoing feedback to students, include more challenging tasks to elicit higher order thinking and expand their methods of data gathering (Shepard). Because learning is a dynamic process, students would undoubtedly benefit from receiving feedback from assessments at the beginning or middle of the year, rather than at the very end (Shepard). Additionally, academic expectations should be transparent and explicit to all students.

As mentioned briefly above, the use of a broader range of assessment tools is necessary in order to truly “capture important learning goals and more directly connect assessment to ongoing instruction” (Shepard). Many researchers contend, for example, that
students should have a more active role in evaluating their own work. As supplements to a teacher’s own formal and informal assessments throughout the year, student evaluations can increase students’ responsibility for their own learning, make the relationship between teachers and students more collaborative and prepare students to defend their opinions with evidence (Shepard). Case studies have even shown that students who participate in regular self-evaluations are likely to find greater interest in criteria and feedback than in their grades (Shepard). Other methods of data gathering should measure whether students are learning to develop important social skills and dispositions and should include student observation, clinical interviews, reflective journals, projects, demonstrations and collections of student work, among others (Shepard). Teachers should aim to closely assess their students’ understandings from a mixture of these elements to ensure that they are engaging in meaningful learning.

The most obvious reform effort in this area has centered on devising more “open-ended performance tasks” to ensure that students can reason critically, solve complex problems and apply their knowledge in real-world situations (Shepard). Both American citizens and education scholars tend to believe these are the most important components by which to measure meaningful student learning. According to the 2013 Phi Delta Kappa/Gallup Poll of the Public’s Attitudes Toward the Public Schools, for example, Americans selected critical thinking skills as the most important of the 21st-century skills, followed by communication skills, setting meaningful goals, motivating students, increasing student creativity, and collaborating on projects (Bushaw & Lopez 45th).

In one particular case study from Making the Grade, author Tony Wagner evaluates the approach of Central Park East Secondary School, from which students could earn their
diplomas by “exhibiting mastery” rather than serving “seat time” (Wagner). The creators of
the project aimed to boost intellectual rigor by placing an emphasis on performance rather
than on credit accumulation; to graduate from the school, students needed to earn “merit
badges,” or “exhibitions of mastery in different intellectual and social domains” (Wagner).
While this is an extreme example that would fail to translate in its current form to
traditional or respectable primary and secondary schools, it does offer insight into
alternative ways of measuring student learning. Too many students today receive passing
course grades by simply attending class and thus graduate with very minimal relevant or
applicable skills.

III. Achieving Meaningful Learning Through Effective Teaching

Defining and measuring what makes for meaningful student learning in K-12
schools is an essential first step to improving America’s education system, but it is of even
greater importance to discuss how teachers inform and advance this learning process. For
the past few decades, research has illustrated that teachers represent the single most
important school-related factor affecting student learning (Cole, Eckes, Murphy & Rogan).
The public tends to agree. According to the 2010 PDK/Gallup Poll, the majority of citizens
felt that improving the quality of teachers should be the number one national education
priority. According to the results of the poll, American citizens believed “we should recruit
the best teacher candidates, provide professional development … and do whatever we can
to retain the best teachers while dismissing those who aren’t skilled or suited for the job”
(Bushaw & Lopez). Teachers provide students with the information and skills necessary for
them to achieve competency, and they offer the feedback essential for student development and progress (Shepard). Additionally, teachers help maintain student motivation.

Research demonstrates that learning is much more likely to transfer if students have the opportunity to practice with a variety of applications while learning; it is thus the teacher’s responsibility to ensure that students receive the necessary practice to achieve growth in meaningful learning (Shepard). According to the literature, teachers must give students the proper time and opportunity to get good at what exactly the standards require. Effective teachers therefore constantly ask about old understandings in new ways, call for new applications and draw new connections, all of which drive meaningful student learning (Shepard). Educators are required and expected to prepare all students to become college- and career-ready through the mastery of sophisticated 21st century skills, such as critical thinking, problem-solving and information literacy skills (Cole, Eckes, Murphy & Rogan). Finally, teachers are responsible in today’s society for fending off the negative effects of standardized testing and for developing a classroom culture that ultimately centers on learning (Shepard).

IV. Teacher Evaluation System Best Practices

In order to ensure that teachers are effective in driving meaningful student learning, it is essential for school districts to assess their teachers properly. From an amalgamation of research on the subject, this thesis argues that a combination of many different components, with none weighing too heavily on its own, is what makes for the best and most effective teacher evaluation system. Systems that place a “heavy weight on any one indicator will tend to be more correlated with the unique dimension of that indicator and
less correlated with the unique dimension captured by other indicators” (Lockwood, Mihaly, McCarey & Staiger). In other words, focusing on only one component of a teacher’s performance will fail to provide an accurate overall representation of a teacher’s true quality and effectiveness. Developing a teacher’s summative rating by equally weighing all combined elements of performance “will not be optimal for targeting any particular dimension of effective teaching, but will be close to optimal across many dimensions and more stable across years” (Cole, Eckes, Murphy & Rogan). This is thus what school corporations should strive toward.

According to a report compiled by The New Teacher Project, an effective teacher evaluation system should include at least an annual evaluation, clear and rigorous expectations, multiple measurements, multiple ratings, regular feedback and significance regarding employment or compensation decisions (Teacher Evaluation 2.0). The primary goal of a successful teacher evaluation system should be to explicitly differentiate performance across a range, from very effective to ineffective teaching (Daley & Kim). However, it should also aim to help instructors improve their teaching abilities through extensive feedback and professional development components. An effective evaluation system should encourage continuous teacher growth and improvement at an individualized level by “collecting and analyzing pertinent data and utilizing those data as the foundation for meaningful feedback and, by extension, as a foundation for flexible compensation” (Little, Gareis, Stronge). The best teacher evaluation systems conceptualize evaluation as a dynamic, ongoing process and aim to constantly measure an instructor’s performance against clearly stated goals. The system should be consistent enough to guide
effective practice for multiple teachers over time, but flexible enough to “encourage individual initiative and creativity” (Little, Gareis, Stronge).

Modern systems suggest evaluating instructors within three broad topics – what teachers bring to their positions, such as experience, content knowledge and certification (inputs), how they interact with students in the classroom (processes) and how they contribute to student growth and achievement (outputs) (Bell, Goe & Little). The latter of the three can also be categorized as an outcomes-based component, which typically uses students’ performance on standardized tests to annually validate teacher effectiveness.

Many education scholars today are in favor of a merit-based teacher compensation system that takes all of these elements – including student achievement – into account when determining teacher pay.

Before examining these ideas in greater detail, it is first interesting to note how American citizens, both from across the nation and from the state of Indiana exclusively, feel about the reform of teacher evaluation systems in K-12 schools. Does public opinion match up with what education scholars have deemed effective? When asked for their opinions on teacher compensation and the utilization of specific evaluation tools, the majority of American citizens have voiced opinions similar to those found in the literature.

In the Center for Evaluation and Education Policy’s (CEEP) 2010 Public Opinion Survey on K-12 Education Issues in Indiana, more than half of Indiana residents indicated that teacher evaluations should play a role in teacher salaries (Plucker, Spradlin & Whiteman). According to results from the 2010 PDK/Gallup Poll, the majority of Americans supported a change in the way teachers are paid so that quality of work, including improvements in student learning, determines how much teachers earn (Bushaw & Lopez). Both public
citizens and education scholars have thus voiced increasing support for merit-based teacher compensation systems, which will be discussed more extensively later in this section. With regard to the use of specific evaluation tools, CEEP’s poll revealed that Indiana residents supported the use of market forces, such as performance pay, in addition to traditional means, such as principals’ observations and years of experience, to measure teacher performance (Plucker, Spradlin & Whiteman). CEEP deducted from these results that Hoosiers supported “a variety of methods, measures and indicators in all areas of teacher evaluation and compensation” (Plucker, Spradlin & Whiteman) – a deduction that very closely matches the conclusion of this paper.

Public opinion on the role of standardized testing in teacher evaluation and compensation has, however, been slightly less congruent with scholarly opinion, particularly in recent years. According to CEEP’s 2010 poll, more than ¾ of Indiana residents believed student achievement in the classroom should influence teachers’ pay, while 58.5 percent felt that compensation should be influenced by students’ performance on standardized tests (Plucker, Spradlin & Whiteman). In the 2012 PDK/Gallup poll, an even slighter majority of citizens – 52 percent – favored including students’ standardized tests scores in teacher evaluations (Bushaw & Lopez 44th). One year later, in the most recent PDK/Gallup Poll, American opinion completely reversed, with 58 percent of citizens indicating that they opposed requiring teacher evaluations to include students’ standardized test scores (Bushaw & Lopez 45th). This suggests that recent trends in state and school district policy to assess teachers based partially on their contributions to student growth on standardized tests might not, in fact, have the support of most Americans.
Although we have clearly witnessed a growing emphasis on students’ standardized test scores in measuring teacher effectiveness, traditional means, such as a teacher’s years of experience and principals’ observations, are still exceedingly important to the evaluation process (Daley & Kim). The best teacher assessment systems have an observation component whereby multiple observers conduct both formal and informal observations multiple times a year. The well-respected TAP evaluation system recommends observing teachers four to six times a year with multiple trained raters using a research-based rubric (Daley & Kim). This qualitative element is meant to provide continuous feedback to teachers to help them improve and develop professionally, but in order for the assessment to be fair and useful, school corporations must also focus on proper observer training (Daley & Kim).

To further assess what teachers are doing in the classroom, teacher evaluation systems should include opportunities for instructors to submit student artifacts, such as written papers or projects, to demonstrate student learning, as well as teaching portfolios, teacher self-reports of practice (including surveys, teaching logs and interviews) and even student ratings of performance (Bell, Goe & Little). Additionally, school administrators should be encouraged to conduct both formal and informal evaluations of teachers. Formal evaluations should be used to gather information for a specific decision-making process, such as employment or compensation decisions; they thus carry substantial consequences. Informal observations, however, hold a formative purpose and are instead intended to gather information for feedback purposes to help teachers improve (Bell, Goe & Little).

As this paper has discussed, instructor effectiveness should be measured across all aspects of teaching. For example, an effective teacher should be successful at not only
fostering students’ mastery of formal academic subject matter, but also at socializing
students and promoting their personal development (Bell, Goe & Little). Evaluations should
thus measure a teacher’s effectiveness in “using diverse resources to plan and structure
engaging learning opportunities and in contributing to positive academic and attitudinal
outcomes for students such as regular attendance, on-time promotion to the next grade and
on-time graduation” (Bell, Goe & Little). Additionally, teachers should be responsible for
effectively collaborating with other teachers, administrators and parents to ensure that
students succeed in the stated learning objectives (Bell, Goe & Little).

One of the most hotly debated issues in teacher evaluation reform today is whether
a teacher’s pay should depend on his or her summative rating from the many components
of performance detailed above. The controversy has arisen primarily because many
teachers do not want their compensation to depend, even partly, on students’ performance
on standardized tests. Many teachers argue that they should not be held entirely
accountable for their students’ growth and achievement, because so many other factors –
such as a student’s home life or financial situation, or the fact that a student is a second-
language English speaker – affect performance but are outside of the teacher’s control.
Despite the many drawbacks and criticisms, however, research – as well as many
developing real-world examples – has demonstrated that teacher evaluation systems that
incorporate merit-based pay are ultimately beneficial in encouraging both teacher and
school-wide improvement.

The increasingly widespread support for merit-based compensation is grounded in
the idea that replacing single-salary schedules with financial incentives based at least
partly on performance will motivate teachers to improve instruction (Barnett & Ritter),
which will in turn promote meaningful student learning. The hope is that such a system will also attract better teachers to the profession. Although the teaching field is not necessarily desired by those motivated primarily by money, research has shown that many who leave the profession dissatisfied cite the low salary as one of their reasons for leaving (Little, Gareis, Stronge). Financial incentives are, in any career, beneficial to both productivity and professional growth and accomplishment (Little, Gareis, Stronge).

The single-salary schedules that have ruled the teaching profession for years were originally intended to ensure equal pay for teachers with the same qualifications, regardless of such things as race or gender. According to the literature, however, such systems are now severely outdated, as they fail to differentiate between superior and inferior performance and result in compensating effective and ineffective teachers equally (Barnett & Ritter). This thesis contends that teachers should be compensated based on their performance, so long as their evaluations include the variety of components detailed above. Performance pay systems must be clear and understandable and should actively encourage collaboration while discouraging counterproductive competition amongst teachers (Ritter & Barnett). Furthermore, merit pay bonuses should be substantial and meaningful if the goal is to truly motivate teachers to improve in the classroom (Ritter & Barnett).

Although opponents of performance pay contend that such a system will reduce teachers’ intrinsic motivation and lead to perverse incentives, little evidence exists to support this claim. On the contrary, studies of existing performance pay programs illustrate that “teachers see both student learning and bonuses as positive outcomes of programs and do not perceive the two as antithetical” (Little, Gareis, Stronge). Additionally, numerous
case studies of schools that have successfully implemented merit-based teacher compensation systems are beginning to emerge all over the country. Take, for example, a program called IMPACT, which was introduced to Washington, DC public schools in 2009. IMPACT was one of the first systems in the nation to not only evaluate, but also compensate teachers based on student achievement, instructional expertise, collaboration and professionalism (An Overview of IMPACT). Under the IMPACT performance assessment, instructors are rated as “highly effective,” “effective,” “developing,” “minimally effective” or “ineffective,” and there are immediate consequences for teachers based on their performance. For instance, teachers rated “ineffective” for one year are immediately fired from the school system. Teachers rated “highly effective” get a substantial bonus, and the second time they earn that rating, they get a significant base-salary increase (Brown).

IMPACT was met with many mixed reviews at its introduction four years ago, but a recent study indicates that the imminent consequences of the evaluation system have actually inspired “low-scoring teachers facing the prospect of being fired, and high-scoring teachers within striking distance of a substantial merit raise” to improve significantly more than others (Brown). It also found that teachers rated “minimally effective,” the second to lowest rating, were more than twice as likely to leave their jobs voluntarily than teachers with higher ratings (Brown). This evidence suggests that merit-based compensation systems actually motivate teachers to get better. Opponents of IMPACT do not necessarily agree, however, as some have argued that a handful of excellent teachers in impoverished schools in the DC area quit after low test scores pushed their rating down to “minimally effective” (Brown).
To education scholars, real-world examples like IMPACT illustrate that merit-based teacher compensation systems can be successful. Systems such as IMPACT are still in their early stages, and researchers need to collect more evidence before they can comprehensively assess and understand whether performance pay systems are truly effective in both the short- and long-term. According to what research currently exists, however, this thesis argues that teacher merit pay encourages and rewards teacher and school-wide improvement.

V. Indiana’s New Teacher Evaluation System

Three years ago, the Indiana General Assembly passed a reform package, the Senate Enrolled Act 001 (SEA 1), which significantly altered how teachers are evaluated in the state of Indiana (Cole, Eckes, Murphy & Rogan). Among other measures, SEA 1 requires schools to evaluate teachers annually, to use objective measures of student achievement and growth, to provide an explanation of the evaluator’s recommendation for improvement and the time in which improvement is expected, and to use rigorous measures of teacher effectiveness across four rating categories – “highly effective,” “effective,” “needs improvement,” and “ineffective” (Cole, Eckes, Murphy & Rogan).

According to SEA 1, a teacher who negatively affects student achievement and growth cannot receive a rating of “effective” or “highly effective” (Cole, Eckes, Murphy & Rogan). Indiana state law also requires schools to tie their assessments to pay but does not necessarily specify how individual districts should award effective educators. It is thus up to the districts to design a merit pay system that “complies with state law but also satisfies
Local districts have the freedom to choose which teacher evaluation tools to use, as well as how to train their evaluators.

Although not required, many school districts have opted to assess their teachers using RISE Indiana, the teacher evaluation rubric developed by the state’s Department of Education. Other school corporations have chosen to use an adaptation of RISE, or to select tools from other established teacher evaluation systems (Cole, Eckes, Murphy & Rogan). RISE uses two components – professional practice and evidence of student learning – to provide teachers with a summative evaluation rating. Professional practice under RISE is an assessment of a teacher’s instructional knowledge and skills, and it is measured through frequent classroom observations and teacher-principal conferences. The observations are conducted by one primary evaluator and many secondary evaluators, who are responsible for conducting observations as objectively as possible, providing feedback and gathering evidence and artifacts of student learning (RISE: The Professional Practice Component).

The student-learning component of RISE is intended to “shine a spotlight on teachers’ contributions to student academic progress” (RISE: Measures of Student Learning Component). To measure student learning, RISE uses the mathematical Individual Growth Model data (IGM) for fourth to eighth grade teachers in math and the language arts, school-wide learning measures and student learning objectives for non-growth model grades and subjects (RISE: Measures of Student Learning Component). Both components are then combined to create a teacher’s summative evaluative rating, from which compensation can be based. Under new Indiana law, teachers rated “ineffective” or “needs improvement” are not eligible to receive any salary increase for the following year, and funds allocated for their compensation are redistributed to teachers rated “effective” or “highly effective.”
(RISE: Scoring and Rating). As one might expect, such a policy has been somewhat controversial in the state.

VI. Evaluating the Effectiveness of RISE

Overall, RISE is a well rounded and well-researched teacher evaluation system, and each of its components has an explicit and specific purpose with ties to what the literature has deemed best assessment practices. As mentioned above, the best teacher evaluation systems utilize a combination of methodologies and measurements, and RISE’s inclusion of both instructional quality and student learning components gives teachers the opportunity to prove themselves in many different ways and settings. Indiana’s new evaluation system, which bases teachers’ annual summative ratings heavily on frequent observations and conferences, should be commended for placing such an emphasis on what teachers are actually doing in the classroom, rather than focusing solely on students’ growth on standardized tests. Indiana’s system adheres to best practices in using a variety of evaluators and multiple observations throughout the year, which ensures that teachers will be rated on their performance more fairly and comprehensively across an extended period of time. This element also emphasizes the importance of providing continuous evaluation and feedback, which is the best way for administrators to help their instructors improve professionally.

The first numbers for RISE were released less than a month ago, however, and results found that only 2 percent of the states’ educators are in need of improvement, with less then 1 percent deemed ineffective (Associated Press, Wildeman). According to members of the Indiana State Board of Education, this provided clear evidence that the new
performance evaluation system failed parents, students and teachers, because the results were too good to be true – and thus indicated that the new system is flawed (Associated Press).

As such, this thesis offers a few recommendations for improvement. For one, RISE should place more of an emphasis on the collection of additional evidence of student learning to fully evaluate a teacher’s effectiveness in the classroom. Specifically, Indiana’s new evaluation system should more vigorously encourage teachers to submit student artifacts, such as creative writing samples, papers or projects, to demonstrate outcomes of their instruction. Research shows that such artifacts can be very representative of what students are learning in the classroom on a day-to-day basis.

Additionally, there is an issue of equity in evaluation under Indiana’s new assessment system that needs attention. Reading and math teachers, for example, are likely to be held to a more rigorous standard than teachers in the fine arts and other non-contested areas (Cole, Eckes, Murphy & Rogan). One solution to this problem is to encourage all teachers, including those in the “accountability areas,” to submit qualitative student artifacts in order to validate the results of standardized measures in determining student growth ratings (Cole, Eckes, Murphy & Rogan).

Lastly, but most importantly, Indiana’s new teacher evaluation system lacks consistency at the local level. The fact that the system leaves many important decisions – such as which evaluation tools to use and how to properly train evaluators – up to local school corporations and districts is both a benefit and a flaw. On one hand, such a policy allows school corporations to “select a tool that may already be in use or is aligned with the corporation goals, or to develop one that is more aligned to the corporation’s instructional
focus” (Cole, Eckes, Murphy & Rogan). To some extent, local schools probably do know what works best for their teachers and students. However, such a system makes it difficult to standardize evaluations across the state, which likely contributed to RISE’s faulty first numbers. Although doing so would undoubtedly elicit criticism and garner some resistance, this thesis argues that Indiana needs to take away some local control over evaluator training, since there is currently no consistency or clear training standards, and thus no way to guarantee that evaluators are highly trained and skilled (Cole, Eckes, Murphy & Rogan). This variability can be extremely unfair to teachers, especially if their compensation depends on it.

VII. Conclusion

The question thus remains: Does Indiana’s new teacher evaluation system promote meaningful student learning in K-12 schools? The answer to this complex inquiry depends on whether or not Indiana’s new program can effectively differentiate teachers across a range and make hiring and compensation decisions accordingly. If this were true, then Indiana’s new instructor appraisal system would stimulate effective teaching, which would in turn promote meaningful student learning in the classroom.

Based on the recently released statistics cited above, however, which deemed only 3 percent of the state’s teachers either ineffective or in need of improvement, this is not the case. Although Indiana’s new teacher evaluation system has numerous positive attributes and a solid foundation in research of best practices, it has seemingly failed to reward high-performing instructors, aid the professional development of those in the middle performance categories and encourage those completely lacking the appropriate skills to
leave the profession entirely. As such, this thesis concludes that Indiana’s teacher
evaluation system in its current state is not entirely effective in promoting meaningful
student learning – which is not to say that it will be incapable of doing so successfully in the
future.

Ultimately, Indiana’s new teacher appraisal system under the RISE rubric needs
time to improve and adapt, and additional studies on its effectiveness are needed before
comprehensive conclusions can be drawn regarding its ability to drive effective teaching
and promote meaningful student learning. Indiana’s new program should be commended
for its utilization of multiple evaluation components and methodologies. The system is
clearly based on what education scholars have deemed best evaluation practices, and it has
the potential to motivate effective teaching (and thereby drive meaningful student
learning) sometime in the future.
Works Cited


