

2014

# Teacher Evaluation, Development and Improvement

Mary K. Barden  
*University of Rhode Island*

Follow this and additional works at: [http://digitalcommons.uri.edu/lrc\\_paper\\_series](http://digitalcommons.uri.edu/lrc_paper_series)

---

## Recommended Citation

Barden, Mary K., "Teacher Evaluation, Development and Improvement" (2014). *Seminar Research Paper Series*. Paper 27.  
[http://digitalcommons.uri.edu/lrc\\_paper\\_series/27](http://digitalcommons.uri.edu/lrc_paper_series/27)[http://digitalcommons.uri.edu/lrc\\_paper\\_series/27](http://digitalcommons.uri.edu/lrc_paper_series/27)

This Seminar Paper is brought to you for free and open access by the Schmidt Labor Research Center at DigitalCommons@URI. It has been accepted for inclusion in Seminar Research Paper Series by an authorized administrator of DigitalCommons@URI. For more information, please contact [digitalcommons@etal.uri.edu](mailto:digitalcommons@etal.uri.edu).

## TEACHER EVALUATION, DEVELOPMENT AND IMPROVEMENT

MARY K. BARDEN

University of Rhode Island

Teacher evaluation has become the focal point for education reform and improvement of U.S. schools over the last several years. At the center of the debate is the primary function of teacher performance management systems. This paper argues that teacher performance management systems which focus on the function of development and improvement will be most effective. Current teacher evaluation systems are examined using the transtheoretical model for behavioral change and examine which evaluation systems create the environment for teacher behavioral change that will be most successful. Ultimately, an organizational and educational culture of trust and support is necessary for the implementation of an effective teacher performance management system.

“Every classroom should have a well-educated, professional teacher, and school systems should recruit, prepare, and retain teachers who are qualified to do the job. Yet in practice, American public schools generally do a poor job of systematically developing and evaluating teachers” (Baker, et al., 2010: 1). Historically, teacher performance appraisal and management has been viewed as ineffective by multiple stakeholders. These stakeholders include school administrators, school districts, state departments of education and state political leaders, public office holders, teachers, unions, students and parents. Almost all these stakeholder want increased student learning and achievement as a result of teacher performance management systems and there is research evidence that improvement in teacher performance leads to improvement in student achievement (Borman, 2005; MET Project, 2013; Milanowski, 2004; Milanowski, Kimball and White, 2004; Odden, 2004; Rockoff and Speroni, 2010). However, the importance of teacher performance management also varies depending on the perspective of these multiple stake holders. School committees and school districts are interested in cost management as teachers are the largest portion of school budgets; they want to ensure they are employing effective teachers and change salary based on performance. Principals and superintendents want to correctly identify effective teachers to make decisions about retention, tenure and promotion. Teachers want to improve their performance and have a valid system that identifies strengths and weaknesses. Unions want a system that is valid to ensure there is a fair process that protects teachers’ rights.

Ultimately, the system must be seen as valid by all stakeholders. However, the varying stakeholders do not view teacher evaluation in the same way. “Stakeholders may have divergent views of the primary purpose of teacher evaluation and hence, of what constitutes a successful evaluation system” (Darling-Hammond, 1990:20). This makes creating and implementing teacher evaluation extremely difficult as often teachers views on evaluation differ significantly from external stakeholders, creating an inherent conflict. Most importantly, teachers and administrators must have faith in the performance management system for it to be effective. In Rhode Island, for example, a “lack of faith in the accuracy of the (evaluation) scores undermines the ability of evaluators to help teachers improve” (Borg, 2013).

Furthermore, the recent education reform efforts in the United States have spurred a national debate on how best to evaluate teacher performance. Therefore, today’s current education reform and accountability policies include reforms to educator performance evaluation. “ Every aspect of school reform - the creation of more challenging curriculum, the use of ambitious assessments, the implementation of decentralized management, the invention of new model schools and programs - depends on highly-skilled teachers” (Darling-Hammond, 2009:1). These reforms include federal legislation such as No Child Left Behind and federal grant funds under Race to the Top. These federal programs

emphasize student achievement based on standards and the key role of teacher quality in improving student achievement. Moreover, these teacher evaluations have been tied to high stakes decisions about teacher tenure, employment and ultimately certification.

### **WHY TEACHER EVALUATION IS IMPORTANT**

#### **Problems: Historically and Today**

Clearly there has been a historical problem with teacher evaluation that is recognized by all stakeholders. Teacher evaluations were generally done as a check off during a one time observation and teachers were rated either satisfactory or unsatisfactory. Teachers were often not given performance criteria, did not know what the evaluator was looking for and did not receive any meaningful feedback from the process. Awarding teacher tenure was not seen as a rigorous process. As a result, few teachers were rated ineffective or were terminated for performance. Also, the same evaluation process was used for beginning teachers as well as veteran teachers. Moreover, there was no value in the process for teachers. It did not help them improve their teaching. Teacher evaluation was also not linked to professional development and improvement. Weisberg, Sexton, Mulhern and Keeling (2009) document the inadequacies in the traditional teacher evaluation. In the 12 districts in 4 states they surveyed, few teachers were given information about areas for improvement, had informal conversations about practice with their principal or received support. "... 73 percent of teachers surveyed said their most recent evaluation did not identify any development areas, and only 45 percent of teachers who did have development areas identified said they received useful support to improve" (Weisberg, Sexton, Mulhern & Keeling, 2009:6). This process is often referred to as a "drive by evaluation". This typical process was also a very top down performance management system and the sole responsibility for it rested with administrators. Teachers had little investment or control over the process. Overall, "virtually everyone agrees that teacher evaluation in the United States needs overhaul. Existing systems rarely help teachers improve or clearly distinguish those who are succeeding from those who are struggling" (Darling-Hammond, 2013:1). Top performers were not recognized and poor performers were not identified or given support to improve.

However, the historical problem with teacher evaluation has now been magnified and entangled by the magnitude of educational reform. States have attempted to implement rigorous teacher evaluation while also implementing the shift in curriculum to the Common Core State Standards which are also controversial. The new forms of teacher evaluation which have been implemented across the country do not create a culture of support to help teachers develop. Rather, what has been created is a system which sorts, blames and shames teachers. "Schools need a system that ties performance to school goals, monitors performance, and provides feedback, support and consequences on whether growth goals are met. These processes should be linked to the school induction and professional development system" (Kimball, 2011:16). Yet, the development aspect of teacher evaluation has been lost in the drive to fire or deny tenure to low performing teachers. With the current torrent of education reform which has been "too much all at once", teacher evaluation has fallen victim to the accountability movement rather than being seen as a tool to help teachers grow and develop. This can be seen in an example from Tennessee with a principal from a school in Nashville. Although he saw what he deemed to be a good lesson, he could not rate it as such. Although he had seen this veteran teacher group students effectively in other lessons, since he did not see it in this one lesson, he had to give the teacher the lowest rating (Anderson, 2012). All across the country, districts which had little or no teacher evaluation have replaced it with a system that is mechanical and strict. "As states and districts embark on these reforms, it is crucial for schools, teachers and students that we move forward to improve the quality of instruction while avoiding pitfalls that could damage education. It is imperative that we not substitute new problems for familiar ones, but that we instead use this moment of transformation to get teacher evaluation right" (Darling-Hammond,

2013:1). One key to getting teacher evaluation right is its underlying purpose. The main purpose of teacher evaluation can't just be to fire or rank teachers. The purpose of any teacher performance management system should be to "keep talented teachers in the profession and to identify those who can take on roles as mentors, coaches, and teacher leaders who develop curriculum and professional learning opportunities, who redesign schools, and who, in some cases become principals" (Darling-Hammond, 2012:2).

### **PERFORMANCE MANAGEMENT SYSTEMS**

Performance management systems are integral parts of human resource strategies. "Performance management can be defined as all those processes, led by managers, that help employees perform as effectively as possible (Curry, 2000:15.5). Performance can be measured using processes and behaviors or goals and outcomes. Performance that is measured in terms of processes and/or behaviors looks at how a person carries out their responsibilities, methods, procedures and how they get their work done. Typically these performance management systems use behaviorally anchored rating scales or behavioral observation scales. These types of performance management systems are best in situations that are stable, where an unfavorable outcome is costly and situations where clear, measureable objectives are difficult. In contrast, performance measured by goals and outcomes looks at the results. These might be assessment of goal accomplishment, sales quotas or, in the case of teaching, measures of student achievement. The use of outcomes to measure performance should be used when there are clear, measureable objectives and there are multiple acceptable ways to achieve these goals. In teaching, the outcome is student learning or student achievement but a clear definition of what this is or how to measure it is most certainly controversial in education.

In general, there are four variables that affect employee performance. These are effort/motivation; knowledge, skills and abilities (KSAs); role perception and resources. In terms of effort as a factor, higher effort would mean higher performance. Looking at KSAs in performance examines whether the person has the requisite skills to perform the job. Role perception means the employee knows and understands the expectations for performance. Finally, resources are the tools, equipment, people and information necessary to perform at an acceptable level. Using these four variables, an effective performance management system, examines employee behaviors and outcomes against standards, identifies performance gaps and their causes, provides feedback to employees and assists in correcting performance issues. In teaching, performance management focuses on the evaluation of teachers. "Evaluation systems are a critical component of performance management because evaluations provide the signals and underlying information that drive other aspects of performance management" (Weiner & Jacobs, 2011:4). Yet, there is no consensus on what makes the best teacher performance management system. Defining teacher effectiveness is complex; it combines a set of knowledge, skills, abilities, traits, and behaviors as well as outcomes. "Compared to other professionals, teacher performance is more difficult to measure in valid, reliable and fair ways" (Springer & Gardner, 2010: 11). Much of what makes a good teacher performance management system hinges on the perceived function of teacher performance management.

### **Functions of Performance Appraisal**

Performance management is a process which has multiple functions. The four main functions of performance evaluation are performance improvement, employee development, determining relative pay and evaluating the effectiveness of human resource functions. It is difficult to find a performance management system that addresses all these functions and one system that may be effective at determining pay may not be effective at employee development.

The function of teacher evaluation in determining relative pay has had mixed results related to increased teacher performance, changing teacher practices and teacher development (Darling-Hammond, 2009; Yuan, et al., 2012). In some instances, early teacher pay for performance programs were poorly designed and implemented resulting in compensation structures that actually reduced teacher collaboration and productivity and disincentivized teachers working with needy student populations (Darling-Hammond, 2009; Springer & Gardner, 2010).

### **Outcomes: Student Test Scores**

Much of the literature also cautions against basing teacher evaluation substantially on student test scores (Darling-Hammond, 2009; Darling-Hammond, 2013; Danielson, 2000; Looney, 2011; Weiner & Jacobs, 2011). Using student test scores is the outcome model of performance management while looking at a teacher's knowledge, skills and abilities assesses inputs while examining teaching practice is in line with the process model of performance. An effective teacher performance management system will include all these elements but which piece makes up the substantial portion of the evaluation is highly debated.

The current rage in teacher evaluation is the incorporation of Value Added Measures (VAM) of student learning. This newer way of using test scores is an improvement over using just a flat test score of a student as an outcome. However, even the use of VAM is still controversial. " 'value added models' refer to a variety of sophisticated statistical techniques that measure student growth and use one or more years of prior student test scores, as well as other background data, to adjust for pre-existing differences among students when calculating contributions to student test performance" (Fuhrman, 2010). These VAM are supposed to take into account how similar students in similar circumstances perform and ultimately measure the gains a student made from one year to the next on a standardized test. They are currently being used in teacher evaluations across the country as a result of the Race to the Top requirement that some portion of teacher evaluation be based on student achievement. In many states, including Rhode Island, they are being melded with measures of teacher behavior to create a final effectiveness rating. Yet, Milanowski (2011) argues that value added student measures and measures of teacher behaviors assess two very different pieces of teacher performance. "Value added and instructional practice measures two different constructs and have different measurement properties" (Milanowski, 2011: 23). He advises using both but not to average them together. Weiner and Jacobs suggest that "value added data... can help identify the most and least effective teachers in terms of student test score gains, but these data are not very helpful in elucidating why certain teachers excelled or struggled and what teachers should do to improve" (2011:6). Moreover, there may be negative or unintended effects of using student test scores to assess teacher performance. If only the outcome of student's achievement on tests is used in teacher evaluation, it sends the message that test scores are the only thing that is important and it doesn't matter the process of how they are achieved. Often, this results in teachers teaching to the test which significantly narrows curriculum and instruction as teachers only focus on tested content. Also, there are many instructional areas and grade levels where there are no standardized tests or no tests with multiple years of data from which to create a value added score. In states like Tennessee those teachers with no standardized tests have to choose another teacher's scores to use in their evaluation. These impacts include disincentives to teach Special Education, English Language Learners and other high needs students.

Moreover, multiple researchers (Baker et al., 2010; Fuhrman, 2010; Darling-Hammond, 2009; Darling-Hammond, Amerien-Beardsley, Haertel & Rothstein, 2012; Darling-Hammond, 2013) have seriously questioned the validity and reliability of Value Added Measures or models of student achievement as accurate and valid indicators of teacher performance. Most recently, the American Statistical Association has taken a stance against using VAM in evaluating teachers. They state that "VAMs typically measure

correlation, not causation: Effects – positive or negative – attributed to a teacher may actually be caused by other factors that are not captured in the model’ ” (Straus, 2014). Despite the fact, that VAM are supposed to take into account the amount of growth a student made and ascribe that growth to a specific teacher, it is difficult to do so. Darling Hammond et al. (2012) illustrate how a teacher’s VAM scores vary widely depending on the model used to calculate the score. Others (Rockoff & Speroni, 2010) caution against using VAM due to the many challenges they pose. “... ‘Value added measures’ of effectiveness are noisy and can be biased if some teachers are persistently given students that are difficult to teach in ways that hard to observe. Thus, using other information may achieve more stability and accuracy in teacher evaluation” (Rockoff & Speroni, 2010:261). VAM are only accurate if teachers are randomly assigned students. However, this is often not the case. Certain teachers may work well with English language learners or students with learning disabilities. Overall, these Value Added Measures of student achievement may be more useful for research or to examine schools, teams of teachers or programs rather than individual teachers. Finally, value added measures do not identify why certain teachers were effective and others were not and more importantly do not explain how struggling teachers can improve. This is not to say that student test scores should have no place in teacher evaluation. VAM could be used by principals or evaluators as sources of information when conducting evaluations. “Objective performance data provides useful information to principals in constructing employee evaluations and using these evaluations to improve productivity” (Rockoff, Staiger, Kane & Taylor, 2010: 1). This suggests that rather than use student test scores as the sole or major component of teacher evaluation, they can be used to inform evaluators and as a way to verify evaluation results. (Rockoff et al., 2010:2).

Although outcomes in student learning are what most stakeholders want, there are many other influences on student learning in addition to the teacher. These influences include socio-economic status, if the students are English language learners or have learning disabilities, student attendance, student health, the education level of parents, the amount of support students receive outside school, and other teachers the student may have or has had in the past. Student growth can also be influence by class size, materials, instruction time, availability of support and resources in school, or the availability of technology. “It is therefore critical, in order to ensure fairness to teachers, that any plans to reward or punish them for gains their students have or have not made control for differences among students in their family situations and other factors that are beyond the teachers’ control” (Fuhrman, 2010).

Using performance management systems to evaluate the effectiveness of human resource strategy is also an important function of performance management systems but many school systems are still in the early stages of developing performance management systems linked to human capital policies. In fact, many school systems emphasize improvements in curriculum and instruction as part of school reform but neglect to reform their human resource practices. “ ... they often overlook the need to support these changes by changing district human resource (HR) management practice... in the private sector, research has shown that there are clear links between the nature and quality of HR management practices and various indications of organizational performance” (Heneman & Milanowski, 2004:109). The emphasis on educational aspects as well as human resource aspects, including performance management systems, is important for schools to improve.

Teacher evaluation and performance management’s sole objective can also not be to just get rid of bad teachers. Typically these systems use a ranking system and those ranked lowest are terminated. Yet this will not increase the effectiveness of the teachers still employed. “Presumably part of the theory in using an aggressive ‘de-selection’ strategy is that employees will exert greater effort to avoid the sanctions and reap the rewards. This could have the opposite effect if teachers don’t know how to meet the expectations. If new evaluation systems are focused inordinately on removing low performers without commensurate attention to developing the talents of teachers in the middle range of effectiveness, teachers are less likely to improve” (Weiner & Jacobs, 2011:4). Therefore, teacher performance

management systems whose only function is to terminate poor performing teachers will serve no useful purpose for the majority of teachers who are not in that category.

### **TEACHER EVALUATION: CREATING A NEED FOR CHANGE**

Teacher evaluation that would be most effective should improve teacher behavior and practice. The majority of teachers are not the lowest or highest performers. They will not be terminated for poor performance but need to improve their teaching practice and student achievement. When teacher evaluation is based solely on accountability and is the basis for high stakes decisions such as tenure or retention, it may not also be able to serve the function of improving performance. Often, performance management systems that are high stakes are seen as punitive and instill fear. Yet, risk taking is needed to change behavior. This illustrates the need for more than one type of evaluation that would be differentiated depending on the function. One evaluation could be used for high stakes decisions such as tenure and retention but the other type of evaluation would focus on developing and improving teaching practice.

#### **Transtheoretical Model**

Therefore an effective teacher performance management system should promote behavioral change. This type of teacher evaluation emphasizes the developmental function of performance management. The Transtheoretical Model of Change illustrates how and why behavioral change occurs. “The model describes how people modify a problem behavior or acquire a positive behavior” (Cancer Prevention Research Center, 1998). An effective teacher performance management system would bring teachers through the stages of the Transtheoretical Model (TTM) to promote professional development and get them to improved practice. “The Transtheoretical Model of Change has been recognized as the most influential approach to the integration of behavior change theories and practice” (Prochaska, Prochaska & Levesque 2001: 248). The TTM is focused on the individual and designed to match the needs of the individual. Teaching can be a very individualistic profession and day to day decisions are made by teachers in regards to an array of behaviors in relation to teaching practice such as instructional practice, involvement in school district initiatives, classroom management, interactions with other faculty, staff, parents and students. The Transtheoretical Model is a good model for improvement and change in behavior as it emphasizes building self-efficacy and a sense of mastery which is needed to improve. Overall, an effective teacher performance management system will help teachers improve and develop their teaching practice by changing teachers’ behavior.

In the Transtheoretical Model there are five stages to behavioral change. Although used mainly to describe changes in health behaviors like smoking or weight loss, it has also been used to explain organizational change (Prochaska et al., 2001). The five stages are precontemplation, contemplation, preparation, action and maintenance.

In the precontemplation stage, teachers would see their performance as acceptable and would not see any reason to change. “The individual may be at this stage because he or she is uninformed or under informed about the consequences of the given behavior” (Prochaska, 2008:845). There would then need to be some catalyst that would move the person into the next stage. There would need to be something that got the teacher’s attention to create a belief that some part of their teaching practice needs improvement. This might be observation feedback, a conversation with a colleague, student feedback, or attending a conference. When teachers develop this belief that some improvement is needed is when teachers would move into the contemplation stage. They have identified an area to improve but have not yet decided to take action. Here the teacher might realize they need to improve classroom management but they are also weighing that decision. “People in Precontemplation and Contemplation Stages are likely to see change as imposed and can become resistant if forced to take action before they are

prepared. When a majority of staff are in Precontemplation and Contemplation, organizations need to prepare their employees by creating the conditions for change” (Prochaska et al., 2001: 249). This is often what is significantly lacking in teacher evaluation. Evaluation for high stakes is often forced upon teachers who then respond with fear and resistance. Instead, the evaluation system should recognize the stages of change and work with teachers to create conditions where they are willing to change their teaching practice. In the preparation stage, the teacher would decide there is a need for action to improve some teaching practice and then a plan of action is developed. In a performance management system for teachers this would be a formal plan to learn a new skill, attend a training session, observe a colleague or watch a webinar. At this stage the evaluator needs to encourage the plan of action and keep the individual progressing. In the action stage, the teacher would implement the plan and change some behavior like using a new assessment for students. In the maintenance stage, the teacher has now consistently incorporated the new behavior into their teaching practice and it is now part of how or what they do.

This model of behavioral change as described in the Transtheoretical Model is cyclical and not linear. A person may move back and forth between stages and people don’t always maintain the change. “The middle stages of preparation and action are the most changeable, in which individuals are very likely to progress or regress, depending in part on the help they receive. Effective decision making is an important determiner of how people can progress through the stage” (Prochaska, 2008: 846). In terms of teacher evaluation, this indicates that the evaluator and/or school district needs to provide the help teachers need to enact a behavioral change. In a developmental model of teacher evaluation, the teachers are not on their own, but the evaluator is someone who can assist them in the process. This is not a performance management model in which the evaluator just checks boxes on a rubric and hands out a rating. The evaluator must identify ways to assist the teacher in improving. This model reflects a more realistic and accurate view of how people change their behavior. As a model that is based on the individual, it is important for the evaluator to know which stage the teacher is at before change can begin. The evaluator has to know where the teacher is along the continuum and adjust their behaviors to meet the needs of the individual teacher to help move them forward to change. Prochaska et al. argue that “stage matched interventions can have a greater impact than one-size-fits all programs by increasing the likelihood that individuals will take action (2001:251). In the performance management system for teachers, this indicates that professional development opportunities should be specific to the needs of the individual teacher rather than broad based programs. “For example, with employees who are in the precontemplation stage, organizations can facilitate consciousness-raising... to help employees progress to Contemplation (Prochaska et al., 2001: 253).

During the movement between stages in the Transtheoretical Model, people weigh the pros and cons of changing their behavior. This is known as decisional balance (Prochaska, 2008). As a person moves through the stages to changing behavior, the pros begin to outweigh the cons. In teacher evaluation, the teacher will need to see the benefits of changing their practice before taking planning or taking action. Another key component of the Transtheoretical Model is processes of change which represent a set of independent variables which affect the movement between the stages (Prochaska, 2001). “These processes also describe covert and over activities that leaders can encourage or elicit in staff to help them change work behaviors, affects, cognitions, or interpersonal relationships” (Prochaska et al., 2001:250). Evaluators and the performance management system should be helping teachers to move to next stage of behavioral change.

Moreover, Prochaska et al. (2001) caution against forcing behavioral change on employees as this creates resistance. For teacher evaluation, this means that the needed changes in teaching practice cannot just be demonstrated and evaluators then expect that teachers are implementing the changes. “Imposing action with the force of authority on organizations that are not prepared is bound to produce conflict” (Prochaska et al., 2001: 253). Unfortunately these imposed actions seem to be prevalent in

current teacher evaluation specifically and in education reform in general. Rather than imposing change, the TTM suggests assessing where individual teachers are along the continuum of change and matching professional development programs to individual needs. “Planned interventions and interactions are then matched to the employee’s stage. A predictable consequence is greater participation, less resistance, and more progress (or change) toward the desired goal” (Prochaska et al., 2001:255). This type of teacher evaluation is geared toward the needs of the individual teacher and therefore will create more lasting change and increase in performance. Teachers would be given individualized feedback on their progress toward goals.

Using the transtheoretical model, an effective performance management system for teachers would have several characteristics. First, this performance management system will be based on development rather than high stakes decision such as termination or tenure. A separate performance management system should be developed to make these decisions. This performance management system would move teachers through the stages of change. Initially, there should be training for evaluators to understand the developmental purpose a performance management system and how to help teachers move through the stages. As previously mentioned, many principals are only now being held accountable to conduct evaluations so therefore training in how to help teachers recognize areas for improvement is necessary. This is not a “gotcha” system so principals will need to learn how to speak to teachers to emphasize the developmental aspects of the evaluation. Much of this will be organizational change and a shift in focus. However, Prochaska et al. argues that “a more promising approach is social influence in which leaders scientifically and sensitively assess the stages of change in individual employees” (2001:253). Effective performance management systems for teachers would also first help evaluators identify what stage teachers are in. This could be accomplished by a quarterly self-assessment or feedback form. Moreover, teachers also need to be prepared to change or they will be resistant. Again, conditions which emphasize the developmental nature of the evaluation will be important. If teachers know the evaluation’s main purpose is for high stakes, they will not be open to change. This might allow teachers time for reflection on performance and a mechanism to do so. Ultimately, the evaluation system must have multiple components which would be matched to teachers at varying stages. “TTM research has shown that stage matched interventions can have a far greater impact than action-oriented, one-size-fits all programs by increasing participation and increasing the likelihood that individuals will progress to action” (Prochaska et al, 2001:251). This means there is not one professional development program for all teachers and that the training and development is developed on an individual basis.

Moreover, an effective teacher performance management system would move teachers from precontemplation to contemplation. Here, teachers will need some information, data or feedback, which will help create the belief that improvement is needed. This requires mentor teachers, principals and colleagues reaching out to those in this stage. At this stage, a teacher might be invited to join a group or committee that works on certain teaching practice. This might look like monthly data meetings to look at student data. It could be feedback from an observation. It could come out of monthly reflections on instructional practices or classroom management. Observing colleagues on a regular, planned basis might help move someone into the precontemplation stage as they are sparked by a new idea or practice. The importance here is that teachers are given time and mechanisms to do this in an authentic way that is not seen as punitive, negative or overly critical. This cannot be done in lieu of other teaching responsibilities or on top of current responsibilities with no additional time. This requires flexibility on the part of administrators as well as funding. As much as an effective teacher evaluation will create behavioral change in teachers, it will also create institutional change.

Next, an effective teacher performance management system will have characteristics that encourage contemplation. These would encourage teachers in the decisional balance, weighing the pros and cons. At this stage, it is important to reduce the negativity associated with behavioral change.

Evaluators and school districts should help to eliminate the “cons”. The characteristics would be similar to those previously mentioned for movement from pre-contemplation to contemplation.

An effective teacher performance management system will again support teachers in the preparation phase. Here mentor or master teachers would help teachers develop an individual plan of action. This plan would include various types of actions to learn new skills. These include attending professional development training, reading a particular book, participating in a group that meets regularly on a particular practice, watching a webinar and multiple other methods. Districts have to support the plans with time, resources and funding. Then the plan will address the implementation of the new skill or strategy with an opportunity to reflect with someone. The action stage would be the implementation of the plan. In both the action and maintenance stage, a master teacher or mentor would help teachers carry out the plan, reflect on the changes, and then help them adapt and maintain.

Therefore, teacher performance management systems may be most effective in the function of development and performance improvement. The teacher performance management system must be able to identify good teachers and then change teacher behavior. Weiner and Jacobs (2011) refer to this as building teacher effectiveness, not just identifying it. However, historically, “teacher evaluation has generally been defined as a mechanism for appraisal in order to determine fitness for employment rather than a means for improving performance” (Goldstein, 2007:487). This leads to the question of what performance appraisal systems for teachers have the greatest likelihood of changing behavior and increasing performance? Can these performance appraisals measure performance, increase learning and change teacher behavior? What is the effectiveness of these performance appraisal systems in bringing about behavioral change? How can performance management systems identify strengths and weaknesses in teacher performance and then develop teachers to more effective teaching and ultimately an improvement in student learning.

At this current stage in US education reform, it would seem that evaluation that focuses on teacher development, improvement and behavioral change is rare. There are some standards based teacher evaluation systems which can serve the purpose of behavioral change. There may be some teacher evaluations that have behavioral change as an element but they are not being used in that manner. Unfortunately, with the emphasis on student outcomes, much of the conversation around teacher evaluation has focused almost solely on student outcomes (particularly on standardized tests) as the criteria for teacher effectiveness and therefore the basis for teacher evaluation. In fact, much of the focus of the literature on teacher evaluation equates teacher performance (and student achievement and learning) to student test score gains. As this is also the focus of the federal Race to the Top money, there are probably few teacher evaluations that incorporate the characteristics needed for behavioral change as outlined in the transtheoretical model. When they do exist, they are then coupled with heavily weighted student outcomes and therefore their effectiveness at behavioral change is reduced.

## **RESEARCH OUTCOMES: TYPES OF TEACHER EVALUATION**

### **Standards Based Teacher Evaluation**

The first step in bringing about behavioral change in teachers would be a set of criteria that defines good teaching and then for teachers to understand these criteria. It is important to look at why teachers are evaluated. Teachers are evaluated to ensure there are good teachers in the classroom and this can be done by having a common understanding of the definition of a ‘good teacher’. “Everyone in the system – teachers, mentors, coaches, and supervisors – must possess a shared understanding of this definition. Having a common language to describe practice increases the value of the conversations that ensue from classroom observations” (Danielson, 2011:35). This is one area where the literature on teacher evaluation

is in agreement. The basis of any effective performance management system for teachers is a clear set of standards and definitions of expected behaviors that teachers can access and understand.

In recent years, several standards based teacher evaluation systems have been developed which outline the knowledge, skills, abilities and behaviors which characterize good teaching. “Building an HR management system to support the teacher performance competencies that define teacher quality requires developing or adapting a model that specifies these competencies. Performance competencies are actual behaviors engaged in by teachers that theory and research has suggested are linked to student achievement” (Heneman & Milanowski, 2004:110). The literature (Borman & Kimball, 2005; Danielson, 2000; Darling Hammond, 2013; Odden, 2004) is generally in agreement about what a standards based teacher evaluation system should include. Typically there is “a set of standards which describe what teachers should know and be able to do” (Odden, 2004:127). These standards are linked to curriculum, assessment and student learning standards. Teachers are evaluated based on classroom observations but also on other evidence. Standards based teacher evaluations include multiple forms of data in the form of artifacts which are evidence of a teacher’s performance. This evidence might include reflections on teaching, planning materials, assignments, assessments and parent communication. There is also a rubric which can be used to score both the behaviors and the artifacts in terms of level of performance.

With the need for a set of standards to define good teaching, several have been developed. Examples of standards for teaching are the National Board for Professional Teaching Standards (NBPTS) and the Interstate New Teacher Assessment and Support Consortium (INTASC). These standards would be the first step in an effective teacher performance management system that focuses on behavioral change. In order to know that behavior needs to be changed, the target behaviors have to be known. In terms of the transtheoretical model, teaching standards can help move a teacher into precontemplation or contemplation. “One of the reasons these standards seem to promote productive learning through the evaluation process is that they are expressed in performance terms- that is, they describe what teachers should know, be like, and be able to do” (Darling-Hammond, 2013:23)

“One of the most widely used systems that defines good teaching is the Framework for Teaching, which describes not only the teaching that occurs in the classroom, but also the behind the scenes work of planning and other professional work, such as communicating with families and participating in a professional learning community” (Danielson, 2011). The Framework for teaching contains 4 domains of: planning and preparation, the classroom environment, instruction and professional responsibilities. Within each of the domains are 5 or 6 standards. For each aspect, there are four levels of performance: unsatisfactory, basic, proficient and distinguished. One criticism of the Framework for Teaching is that there are too many standards with 22 components. In this evaluation process, using the standards, teachers are then evaluated through observations, walkthroughs and artifacts. Evidence of student learning is included but is not given a specific score. Evidence of student learning also takes multiple forms.

Multiple researchers (Borman, 2005; MET Project, 2013; Milanowski, 2004; Milanowski, Kimball & White, 2004; Odden, 2004; Rockoff & Speroni, 2010) indicate there is a strong correlation between teacher evaluation scores using standards based teacher evaluation and student learning gains. “Districts and schools can design and implement ambitious, performance based teacher evaluation systems that have a substantial degree of criterion validity” (Odden, 2004: 128). An important aspect of any effective performance management system is that it is valid and reliable. Although there are issues with only measuring student performance in relation to teacher effectiveness through standardized tests, the research nonetheless shows a link between teachers’ performance ratings in a standards based evaluation system and student achievement.

Milanowski (2004) conducted a study of teachers in Cincinnati, Ohio, who were evaluated using standards from the Framework for Teaching. Teachers were evaluated on six observations and a portfolio. “The empirical results show that evaluations produced by a relatively rigorous, standards based system

are related to an accepted measure of student learning.” (Milanowski, 2004:49). This study was then incorporated with a study of two other schools/districts in Los Angeles and Washoe County Nevada by Milanowski, Kimball and White (2004) and achieved similar results. Both Cincinnati and Washoe County used adapted versions of the Framework for Teaching. Their research concluded that “... the scores produced by these standards-based teacher evaluation systems have a substantial positive relationship with the achievement of the evaluated teachers’ students” (Milanowski et al., 2004:18).

Rockoff and Speroni (2010) looked at how well subjective and objective evaluation scores of teachers could predict future performance of students. They conducted their study in New York City in grades 3 through 8 with first year teachers. They examined subjective teacher evaluation data from a number of sources. These scores were then compared against student achievement data. Their research concluded that “... teachers who receive higher subjective evaluations either prior to hire or in their first year of teaching produce greater average gains in achievement with their future students” (Rockoff & Speroni, 2010:264).

The Measures of Effective Teaching Project also found similar results in their study of teacher effectiveness rating and student achievement. They looked at teachers observation ratings and student surveys as measures of teacher effectiveness with student achievement gains. “The research confirmed that, as a group, teacher previously identified as more effective caused students to learn more” (2013:6)

Most importantly though, teacher performance management system should promote professional development and help teachers improve their teaching practice. “A commitment to professional learning is important, not because teaching is of poor quality and must be ‘fixed’, but rather because teaching is so hard that we can always improve it. No matter how good a lesson is, we can always make it better” (Danielson, 2011:37). Standards based teacher evaluation focuses on the key elements which might move a teacher between the stages in the transtheoretical model. Teachers will most likely not change if left on their own with no/poor evaluation, no feedback and no collaboration with colleagues. Duke and Stiggins argue that evaluation is crucial to change and growth in teacher behavior. “Once individuals have exhausted their own mental and emotional resources, they are unlikely to be motivated to grow without the intervention of some external impetus... evaluation feedback can provide the challenge found to be vital to stage growth...” (1990:119). With a set of standards, teachers, just by being made aware of them, may move into pre-contemplation. In gathering and examining evidence with colleagues, teachers might realize that certain students are struggling and therefore realize there is a problem, contemplation. Standards based teacher evaluation also emphasizes working with other colleagues to create plans. This is emphasized in the preparation stage of the TTM. Milanowski, Kimball and White argue that “in essence, standards based teacher evaluation systems provide both incentives and guidance for teachers to change their practice toward the model embodied in the standards (2004:2). Standards based teacher evaluation also allows time to reflect after plans are enacted which helps in the maintenance stage of behavioral change. “If we want teacher evaluation systems that teachers find meaningful and from which they can learn, we must use processes that not only are rigorous, valid and reliable, but also engage teachers in those activities that promote learning – namely self-assessment, reflection on practice, and professional conversation” (Danielson, 2011). These activities can serve to bring teachers through the stages of behavioral change. Self-assessment could bring a teacher from precontemplation to contemplation. Reflection on practice could help a teacher move from contemplation to action. Professional conversations with colleagues could help teachers move through all the stages.

## **Portfolios**

Another teacher performance management system that has been found to be effective is the use of portfolios in assessing teachers. Portfolios are often used in conjunction with standards based teaching assessments. Portfolios are an authentic performance assessment as it represents a collection of evidence

of a teacher's instruction. This evidence might include video tapes of instructional lessons, samples of student work, lesson and curriculum plans. The evidence also includes teacher reflections on instruction and assessment. National Board Teacher Certification is a rigorous national certification process for teachers which includes the use of portfolios. Teachers are assessed on their achievement of the National Board standards; this assessment process uses portfolios. "... a number of studies have found that the National Board Certification assessment process distinguishes teachers who are more effective in improving student achievement from other who do not achieve certification" (Darling-Hammond, 2013:27). Through the process of developing the portfolio, a teacher can be brought the stages needed for behavioral change. The process of creating the portfolio can move a teacher from the pre-contemplation to the contemplation stage. "Many studies have found that teachers' participation in the National Board process supports their professional learning and stimulates changes in their practice" (Darling-Hammond, 2013:39). Therefore, portfolios as used in National Board Certification can help create the need for change.

Tucker, Stronge, Gareis and Beers (2003) conducted research on the use and efficacy of portfolios in a school district that adopted portfolios for accountability and to improve instruction. Their research concluded the portfolio did present a valid representation of a teacher's responsibilities and it helped administrators distinguish between varying teacher performance. However, "teachers and administrators gave mixed response about the contribution that portfolios made to the professional growth of teachers" and "some teachers feel the portfolio is a 'bureaucratic exercise'" (Tucker et al., 2003:591). This emphasizes the importance in how evaluation artifacts and evidence are used. If they are merely used as a collection of papers, then they will not aid in helping teachers improve practice. Tucker et al. add that "one possible explanation may be that mechanisms need to be developed to help teachers connect reflections on their work with future action that will affect instructional practice" (2005:591). This research emphasizes that portfolios should be used in a process and not as the final means for evaluation. Self-reflection is important in the behavioral change process and may lead a teacher through pre-contemplation and contemplation but there then needs to be preparation and action for the self-reflection to have value in changing behavior. Clear guidelines about what should be included in the portfolio are also important so that a teacher focuses on more than just the positive aspects of her practice. "For a portfolio to be useful in planning and professional growth activities, it needs to be focused on instructional data about student learning, instructional challenges and how they were addressed, and reflections on practice" (Goe, Biggers, & Croft, 2012:10). Portfolios if used properly can help in the planning stages of behavioral change and could also move teachers into precontemplation or contemplation just from the collection of evidence.

### **Peer Assistance and Review Programs**

Another evaluation system which research has shown to be effective is the Peer Assistance and Review (PAR) process. Goldstein (2007) describes the key components and successes of the program based on a longitudinal study of one district in California. In this system, highly trained mentor teachers conduct evaluations and provide support for teachers who need it. These teachers report to a district wide committee which is overseen by the union president and director of human resources. Teachers are highlighted for PAR based on their poor performance rating from a principal. PAR deals solely with instructional performance issues. Overall, Goldstein (2007) highlights how PAR differs from traditional teacher evaluation in that it requires more time, is transparent, involves the teachers' union and provides a degree of accountability. In relation to helping teachers change behavior, the system focuses on helping poor performing teachers improve because of "the relationship that professional development has to evaluation, where evaluation is linked to support and professional development, including matching evaluators and evaluate by grade and subject and using performance standards" (Goldstein, 2007:484).

Trust is important in behavioral change and teachers will be more likely to trust someone who has taught the same grade level and content when discussing teaching practice. The outcome of the PAR is used to make personnel decisions such as awarding tenure and termination. Overall, Goldstein concluded that “the PAR program... addressed structural barriers to the system of teacher evaluation that allowed the CTs (master teachers) to achieve results that principals are typically unable to achieve” (Goldstein, 2007:497). Darling Hammond (2013) also details the implementation of PAR in two districts in California in research done by Stanford Research International. Although the system is used to make high stakes decisions, it is also effective at helping teachers improve. “Historically about two-thirds of veterans identified for intervention have improved substantially and successfully completed the program; about one-third in each case were resigned or dismissed” (Darling-Hammond, 2013:127). Because the union is involved and there is due process, dismissal does not usually result in a grievance.

### **Professional Development**

The ultimate goal of the teacher performance management system should be to improve practice which occurs through professional development and professional learning opportunities. Standards based teacher evaluation emphasizes continuous improvement and development which is embedded in the evaluation process. However, these must be tailored to the individual needs of the teacher rather than the typical, drive by professional development offered as an after school workshop or one day conference. Teachers need continuous intensive professional development and professional learning opportunities in order to be brought through the stages of change. Behavioral change is cyclical and therefore the supports and impetus for change must also be cyclical. A one shot, one time offering of a workshop will not create behavioral change. This type of professional development might be focused on specific content and curriculum, is connected to what teachers do on a daily basis in the classroom, and helps improve student achievement. Furthermore, professional learning is also important for behavioral change. These are opportunities for teachers to learn from each other through coaching, discussion, models of lessons and time for collaboration. “However, this kind of professional development is relatively rare in the United States. National data show that although most teachers participate in some kind of professional development each year, very few have the change to study an aspect of teaching for more than a day or two” (Darling-Hammond, 2013:101). Most often there is no connection between the evaluation system and the professional development offered to teachers. While professional development could be part of an action plan for behavioral change it will not be effective if it is not tailored to the individual teacher’s needs. “... It is important to link both formal professional development and job-embedded learning opportunities to the evaluation system. Evaluation alone will not improve practice. Productive feedback must be accompanied by opportunities to learn. Evaluations should trigger continuous goal setting for areas teachers want to work on, specific professional development supports and coaching, and opportunities to share expertise, as part of recognizing teachers’ strengths and needs” (Darling-Hammond, 2013:99). Here again the elements of a standards based evaluation system would help move a teacher through the stages of change. Goal setting is part of preparation. The supports and coaching could help someone move from preparation to action but also with maintenance of behaviors. Moreover research has shown that these types of professional development improve student learning. “ For example, a review of experimental studies found that whereas professional development offering of less than 14 hours per year on a given topic had no effect on student learning, higher quality professional development programs averaging about 50 hours over a 6-to12 month period increased student achievement by 21 percentile points on average” (Darling Hammond, 2013:100). Moreover the professional development should be directed by the teacher. Rather than having the evaluation and accompanying professional development be “done to” them, teachers should be at the center of the process. Danielson argues that “...to the extent possible, the teacher (rather than the administrator) should direct the evaluation activities. Evaluation should provide maximum opportunities for self-directed

inquiry. The teacher, in other words, should play as active a role as possible” (2000:61). This gets to the heart of behavioral change under the transtheoretical model; individuals change when they are invested and engaged in the process and see that the pros outweigh the cons.

The evaluation which highlights areas for development alone does not help the teacher change behavior and improve practice. It is the support, feedback and reflection that are inherently built into the evaluation process that does so. “The process of evaluation provides the evidence of areas where teachers need help, but that process alone does not change teaching practice. Rather, using the evidence for professional growth opportunities and coaching sessions is where it will have an impact on instruction and student outcomes. Observation alone without the opportunity for feedback and discussion may serve accountability purposes but will have little or no impact on teaching and learning” (Goe et al., 2012:15). This linking of evaluation and professional development will be key to support teachers in behavioral change.

### **Rhode Island Model Teacher Evaluation System**

Overall, the research suggests that there are effective teacher performance management systems which could promote development through the phases of transtheoretical model, but why haven’t they been used. The literature shows that these standards based teacher evaluation systems do lead to student learning gains and can be effective, yet their implementation has been slow. “Despite these advances, there has been remarkably little effort to connect these standards to district’s on-the-job evaluations of “(Darling-Hammond, 2013:24). Moreover, even when teaching standards are used, they take a back seat to heavily weighted measures of student achievement and/or value added measures. For example, the Rhode Island Model Teacher Evaluation uses elements from Danielson’s Framework for Teaching but the majority of the final score (over 51%) for teaching effectiveness is based on student achievement gains from student learning objectives and/or value added measures. This mingling of two very different measures and constructs of teacher effectiveness has not been successful. Duke and Stiggins argue that teacher “growth can be inhibited as a result of evaluation that is overly threatening, poorly conducted, or inadequately communicated” (1990:119). Unfortunately, the Rhode Island Model has all these characteristics.

Most recently, the Community Training and Assistance Center (Slotnik, Smith & Guodong, 2013) conducted a study of the Rhode Island Model Teacher Evaluation System. The study was commissioned by the Rhode Island Innovation Consortium and the Rhode Island Federation of Teachers and Health Professionals. Although the study focused on the urban districts in the Consortium along with 4 suburban districts, the results are indicative for the whole state, as the other districts implemented an almost similar evaluation model. The study was based on surveys and interviews with teachers, principals, superintendents, union leaders at the local and state level as well as examination of evaluation documents. In its Summary of Recommendations, they find that in Rhode Island “there is substantial agreement that teacher and principal evaluation need to improve. The concern is that districts have gaps in readiness and internal capacity needed to implement a student learning and growth-based approach with fidelity. These influences together contribute to the fear, anxiety, and distrust that are prevalent in a significant number of districts within the state” (Slotnik et al., 2013:22). This is exactly the type of environment that the TTM suggests will prevent any behavioral change.

Moreover, high stakes decisions attached to the evaluation model not only include, tenure and termination but also potential revocation of a teacher’s license. In Rhode Island, “ teachers and principals feel the link of evaluation to certification is punitive, thereby shaping their perceptions of the overall evaluation system and leading them to question its intent” (Slotnik et al., 2013:12). Again, this is seen as threatening and intimidating, leading to a culture in which teachers will not be able to take risks and change behavior. “...teachers... should not fact license revocation unless they are a danger to children.

Just as in other professions, those who fail to meet performance standards of a particular employer should not be barred from the profession, because ‘fit’ matters and an effective match with a new school may lead to improved instructional performance” (Weisberg, et al., 2009:30).

There are lessons to be learned from Rhode Island and other states who have too quickly adopted and implemented teacher evaluation systems that are top down models and rely too heavily on student achievement. These systems ultimately do not lead to improved instruction because teachers do not trust the system. “If there is an irreducible truth of teacher evaluation, it is: Teacher evaluation will be no more effective than the extent to which teachers support it” (McLaughlin, 1990:404). An effective performance management system for teachers will take the time to create support and buy in from teachers. Moreover, “Top down management and a ‘take it or leave it’ approach to new initiatives are not hallmarks of professions that encourage either innovation or outstanding performance” (Weiner & Jacobs, 2011:11). Although, changes have been made to the model as a result of feedback from teachers and principals, the damage in terms of trust has already been done and teachers are unlikely to see the evaluation process as one they can take risks in. In fact, the most recent ranges to the Rhode Island model include less classroom observations, which mean less feedback based on the standards but continues the use of student learning objectives. Going forward, teachers will have to complete student learning objectives every year but principals will only have to observe teachers using the standards every two or three years.

## CONCLUSION

### Teacher Evaluation for Development

Although there has been considerable research on existing teacher evaluation systems which could lead to behavioral change in teachers, there is a lack of research in if or how teacher performance management systems have actually done so. The focus has been on creating common understandings of behaviors and that these behaviors lead to improved student learning as evidenced on standardized tests. This assumption that student learning can be shown through standardized tests is in and of itself a controversial topic but one outside the scope of this paper. The research shows that there are teacher performance management systems that could be used in such a way to help bring about changes in teacher behavior using the transtheoretical model but they are not being used in that way. Ultimately, the purpose of the teacher evaluation needs to be reexamined. If it is for development purposes then it has to be implemented in a different way. There is a fundamental difference in the perception of the purpose of teacher evaluation by the multiple stakeholders involved. Many state departments of education, local school committee and superintendents want teacher evaluation to be solely for the purpose of high stakes decisions on tenure, termination and salary increases. However, teacher development, behavioral change and improvement of practice should be at the core of teacher evaluation. “Teacher evaluation systems need to be designed and implemented with teacher learning and development at their core, rather than appended later as an afterthought. Professional development is regularly associated with the ‘results’ of evaluation, instead of recognized as an integral part of the evaluation process itself. Thus, the power of evaluation to generate greater teaching effectiveness is diminished” (Coggshall, Rasmussen, Colton, Milton & Jacques, 2012:1). For the majority of teachers, an evaluation process that focuses on development will be key. “Educators, on the other hand, tend to think that teacher evaluation should be designed for the purpose of professional development and the improvement of teaching. Experienced practitioners argue that professional dialogue about teaching, in a safe environment, managed and led by teachers, is the only means by which teachers will improve practice” (Danielson, 2000:9).

From the research, there are several aspects of a teacher performance management system which could lead to behavioral change.

**Cultural change.** The single most important element of creating an effective teacher performance management system is the organizational and educational culture in which evaluations are implemented. According to the transtheoretical model, individuals will not change their behavior if they see the behavioral change as risky. All the other recommendations for teacher evaluation will not be effective unless they are implemented in a culture which emphasizes the development and support of teachers in their practice. Teachers must feel supported and competent in changing the behavior in order to do so. Danielson argues that “the environment for such conversation...must be safe for taking professional risks...the discussions must be genuine professional conversations, without undercurrents of point scoring, or posturing, that occasionally characterize such debates” (2013:26-27). Sadly the new models of teacher evaluation have not created this culture. Darling-Hammond (2013) details the concerns of a principal in Tennessee using the new teacher evaluation. “Shelton argues that between the test-based requirements and the excessive observation requirements, the new state evaluation policies ‘put everyone under stress, are divisive, and suck the joy out of a building’ “ (Darling-Hammond, 2013:135). This type of culture cannot implement a model of teacher performance management based on behavioral change. Moreover, the culture cannot be a top down, control approach to education or management. Teachers need to trust evaluators and administrators to take the risk necessary to try new methods of instruction or behavior. However, in a system where they are not part of decisions made about instruction or evaluation, they do not feel value or invested. Prochaska et al. argue that imposing change is detrimental to behavioral change. “Imposing action with the force of authority on organizations that are not prepared is bound to produce conflict” (2001:253). With this also comes the recognition, that change occurs slowly and that leaders must prepare teachers for change. Principals and evaluators must create culture that eliminates the us versus them mentality in which leaders and teachers work collaboratively and cooperatively. The trust also must go both ways; teachers must trust school leaders but that will not happen until teachers are trusted and respected as professionals. Ultimately, this cultural change is the most important factor to an effective teacher performance management system. Without a supportive culture, teacher performance management systems will not result in behavioral change for teachers.

**Tiered System.** An effective teacher performance management system would be a tiered system. Districts need to make decisions on tenure, retention and pay, but systems which result in those decisions should be different from those used to help teachers develop. Separate evaluations should be used for non-tenured teachers, tenured teachers and teachers with performance issues. As the ultimate goal in evaluating these teachers is different, the system cannot be the same. There is an inherent conflict in a system that both develops teachers and could lead to their dismissal. Some (Danielson, 2011; Goe, Biggers, & Croft, 2012) would argue that a standards based teacher evaluation can be used for both development and high stakes decisions such as tenure and termination. However, the transtheoretical model research indicates otherwise. In order to change behavior, teachers need to be able to take risks to change behavior but often fear is instilled in teacher evaluations which are used for high stakes decision. Duke and Stiggins argue that “... researchers concluded, however, that both accountability and professional development purposes could not be served easily by the same teacher evaluation system” (1990:120). The implementation of the Rhode Island model is evidence of this as teachers see that the “accountability feature trumps the continuous improvement goals and philosophy in the new evaluation system” (Slotnik et al., 2013:11). This type of system creates fear and teachers don’t want to share areas needed for development if that is going to affect their rating. Also if a teacher takes on a challenging aspect for development and then they do not not achieve it, depending on the evaluator, it might be lead to a negative evaluation. This then encourages teachers to “play it safe”, rather than attempt challenging growth.

**Teacher involvement.** In order for behavioral change to result from a teacher performance management system, teachers need to be an integral part of the system in the development, planning and implementation. This will assist in teachers trusting not only the process but also administrators and evaluators." The exclusion of teachers from the process perpetuates a them/us schism between administrators and teacher, which is fatal to teacher evaluation and reinforces a view of teacher evaluation as indifferent to teacher' professional expertise and classroom realities" (McLaughlin, 1990:406).

**Standards.** The literature is in agreement that high quality standards are the basis for an effective performance management system. Moreover, the literature has shown that standards based teacher evaluation is linked to improved student achievement. As the literature cautions against using student test scores and value added models, a standards based teacher assessment should be the primary evaluation process. "Standards based teacher evaluation systems are based on common conception of teaching, developed from empirical and theoretical literature on effective teaching behaviors, and assessed using multiple, authentic sources of teaching evidence" (Borman & Kimball, 2005:5). These standards should be clear and transparent to all stakeholders. However, they especially need to be made clear to teachers. Having clear standards will help teachers in their process of changing behavior as they outline the criteria for successful performance. Knowledge alone may move a teacher from to the pre-contemplation stage. Also, care needs to be taken that the standards are not overly burdensome and complex. The meeting of the standards should not force teachers into becoming collectors of data and paperwork. "... In some systems, the desire to capture every aspect of teaching has led to systems that require reams of paperwork to address rubrics that are dozens of pages long, and documentation that takes many hours to complete" (Darling-Hammond, 2013:137). As meaningful as artifacts and data may be, the collection and analysis should not become the central focus of a teacher's job, but rather should help in the development process. Campbell warns that in terms of teacher evaluation, ..."the form of implementation will be critical; the complexity threatens to move the mechanisms of measurement and evidence collection to the nightmarishly bureaucratic" (2013:357). This can lead teachers to see evaluation as giving the highest rating to those who can "push the most papers".

**Evidence from student learning and achievement.** Multiple measures and evidence of student learning and achievement is important to an effective teacher performance management system. Teachers may be moved through the stages of behavioral change by examining and analyzing student learning and achievement. Yet it is important how these measures are used. "...states and districts around the country are still seeking that best measures to use for teacher performance, and a consensus has yet to emerge on what constitutes fair, comparable, and equitable measures of teacher's contribution to student achievement" (Slotnik et al., 2013:11). Student measures of achievement should be used collaboratively with other educators to find areas where practice can be improved and can even provide information about what changes to make. However, using them as the main or sole criteria for teacher evaluation is detrimental. One common element in all the research on teacher evaluations is that they do include evidence of student learning and achievement. Yet, it is typically not based on standardized tests. These evaluations allow teachers to use student work samples as evidence of student achievement. These work samples would have to represent the makeup of the class and the array of achievement of students. The samples should focus on important content from student learning standards. " Though presenting certain challenges, such an approach has the advantage of avoiding most of the technical difficulties with relying on tests of student achievement, while preserving attention to the most important aspect of a teacher's work, namely whether the students are actually learning" (Danielson, 2000:43).

**Trained evaluators.** Much of the literature on teacher evaluation (Danielson, 2011; Darling Hammond, 2013; Goe et al, 2012; Milanowski, 2011) also emphasizes the importance of trained evaluators. Principals will need not only training in how to evaluate and observe but training in how to

create and lead a culture of continued learning. Principals and other school leaders need training in how to have professional learning centered conversations with teachers. It is the learning that comes out of the evidence or observations which is most important to helping teacher see the need for change in practice and supporting them in that practice. The use of any teacher performance management system relies on the training and expertise of evaluators. This will not be a onetime training and will require an investment of time and money. Danielson indicates that “our findings have been somewhat humbling; even after training; most observers require multiple opportunities to practice using the framework effectively and to calibrate their judgments” (2011).

**Alignment with HR practices.** It is also important to recognize that teacher evaluation should be part of an overall, comprehensive human resource strategy for school districts. Teacher performance management is intertwined with selection, retention, labor relations and compensation but also with the performance management systems of administrators. Yet, teacher evaluation seems to be developed independent of other human resource functions and without any overarching strategy. “To find, develop, and retain the most effective teachers, evaluations need to be complemented by other critical elements of a comprehensive, interdependent set of strategies” (Weiner& Jacobs, 2011:5). Cooperative labor management relations are important to getting teacher buy in and support from teachers’ unions. Administration or state departments of education cannot just impose evaluation systems and the expect them to be effective. “If adversarial negotiations and positional bargaining are the principal modes of joint work on new systems, then opportunities for deep collaboration will be squandered” (Weiner & Jacobs, 2011:10). In the private sector, human resource management often focuses on alignment of human resource practices and with a company’s overall business strategy as a means to improve overall performance. This piece of the conversation seems to be missing in regards to the current wave of teacher evaluation. “Teacher competency is joint product of instructional and HR practices.... HR practices must be vertically aligned to a valid teacher performance competency model, and HR practices must also be aligned horizontally to instructional practices and to each other” (Heneman & Milanowski, 2004:115). In the educational context the overall strategy is student learning. However, what constitutes student learning and moreover what is evidence of student learning is controversial as is the total impact teachers have on this learning.

### **What is Missed in the Focus on Teacher Evaluation**

Although US teacher evaluation has needed an overhaul, the hyper intensive focus on it may ultimately be displaced. Finland is seen as one of the top performing countries in education. Yet, they have no formal evaluation system for teachers. “The question of teacher effectiveness or consequences of being an ineffective teacher is not relevant in Finland... today school principals aided by their own experience as teachers, are able to help their teachers to recognize strengths and areas of work that need improvement” (Sahlberg, 2010:91). The emphasis in Finland is on the time for teachers to collaborate and work together in order to improve. Darling-Hammond argues, “There is relatively little emphasis in Finland on formal on-the-job evaluation, and much more emphasis on collaboration among professionals to promote student learning. In truth, we cannot fire our way to Finland” (2013:7). Educational policy makers and state boards of education would do well to learn from this. Moreover, US teachers spend much more time teaching and less time collaborating with colleagues than teachers in high performing countries in Europe and Asia (Darling Hammond, 2013). The collaboration and time are essential to creating a culture where teachers feel free to take risks to learning from each other and ultimately change their teaching practice.

The hyper focus on teacher evaluation may also take the focus away from areas that also highly impact student learning. By focusing on teacher evaluation, the conversation does not focus on resources, funding, class size and teaching conditions which have an equally important effect. Darling Hammond

cautions “an excellent teacher may not be able to offer high-quality instruction in a context where he or she is asked to teach a flawed curriculum or lacks appropriate materials. Similarly, a well prepared teacher may perform poorly when asked to teach outside the field of his or her preparation or under poor teaching conditions- for example without adequate teaching materials, in substandard space, with too little time, or with classes that are far too large” (2013:12). As important as teacher evaluation is to improving education in the US, these other areas cannot be overlooked.

Teachers and other stakeholders realize that there will need to be changes both from the old “drive-by method” of teacher evaluation but also changes from the “new” high stakes teacher evaluation based heavily on student test scores. “ Some of the requisite changes are obvious: professional development must respond directly to the areas identified in teachers’ evaluations rather than provided indiscriminately to large groups of teachers without regard to individual needs” (Weiner & Jacobs, 2011:12). Unfortunately, the teacher evaluation systems that have been implemented in many states fail to provide a structure for behavioral change and development. Instead, they are inordinately focused on “getting rid of bad teachers”. With the majority of focus of research has been on whether certain teacher evaluations are linked to increases in student achievement but very little has been done on whether the evaluations help teachers develop and improve. Because teacher evaluation historically was so problematic, states and districts have rushed to change teacher evaluation. “The urgency of the need has pushed ahead of the research on the subject, and states and districts are now attempting to find a balance between moving forward quickly, but also fairly” (Goe et al., 2012:22). In order to ensure that teacher evaluation is fair, its focus should be on teacher development and improvement.

#### REFERENCES

- Anderson, L. 2004. Increasing teacher effectiveness. <http://www.unesco.org/iiep>. Paris: UNESCO, International Institute for Educational Planning.
- Baker, E., Barton, P., Darling-Hammond, L., Haertel, E., Ladd, H., Linn, R., Ravitch, D., Rothstein, R., Shavelson, R., & Shepard, L. 2010. Problems with the use of student test scores to evaluate teachers. EPI briefing paper #278. <http://www.epi.org>. Washington, D.C.: Economic Policy Institute.
- Borg, L. 2013. Study: R.I. teacher evaluations seen as punitive. Providence Journal. <http://www.providencejournal.com>, September 13.
- Borg, L. 2013. High evaluation ratings for most R.I. teachers problematic. Providence Journal. <http://www.providencejournal.com>, October 11.
- Campbell, R.J., Kyriakides, L., Muijs, R.D., & Robinson, W. 2003. Differential teacher effectiveness: Towards a model for research and teacher appraisal. *Oxford Review of Education*, 29(3): 347-362.
- Coggshall, J., Rasmussen, C., Colton, A., Milton, J. & Jacques, C. 2012. Generating teacher effectiveness: the role of job-embedded professional development learning in teacher evaluation. <http://www.tqsource.org>. Washington, D.C.: The National Comprehensive Center for Teacher Quality.
- Curry, T. H. 2000. Moving from performance appraisal to performance management. In E. Kossek & R. Block (Ed.), *Managing human resources in the 21st century: from core concepts to strategic choice*:15.1-15.22. Cincinnati: South-Western College Publishing.
- Bacal, R. 2012. *Manager’s guide to performance management*. New York: McGraw Hill.
- Borman, G. & Kimball, S. 2005. Teacher quality and educational equality: do teachers with higher standards-based evaluation ratings close student achievement gaps?. *The Elementary School Journal*, 106 (1):3-20.

- Cancer Prevention Research Center. 1998. Detailed overview of the transtheoretical model. <http://www.uri.edu/research/cprc/TTM/detailedoverview.htm>. Kingston, RI: University of Rhode Island.
- Curry, T. 2000. Moving from performance appraisal to performance management. In E. Kossek & R. Block (Eds.), *Managing human resources in the 21st century*. Cincinnati: South-Western College Publishing.
- Danielson, C & McGreal, T. 2000. *Teacher evaluation to enhance professional practice*. Princeton: ETS.
- Danielson, C. 2011. Evaluations that help teachers learn. *The Effective Educator*, 68 (4):35-39.
- Darling-Hammond, L. 1990. Teacher evaluation in transition: emerging roles and evolving methods. In J. Millmand & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers*:17-32. Newbury Park, CA:Sage.
- Darling-Hammond, L. 2009. Recognizing and Enhancing Teacher Effectiveness. *The International Journal of Educational and Psychological Assessment*, 3:1-24.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E. and Rothstein, J. 2012. Evaluating teacher evaluation. *The Phi Delta Kappan*, 93 (6):8-15.
- Darling-Hammond, L. 2013. *Getting teacher evaluation right*. New York: Teachers College Press.
- Duke, D. & Stiggins, J. 1990. Beyond minimum competence: Evaluation for Professional Development. In J. Millmand & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers*:116-132. Newbury Park, CA:Sage.
- Fuhrman, S. 2010. Tying teacher evaluation to student achievement. *Education Week*. <http://edweek.org/> April 6.
- Goldstein, J. 2007. Easy to dance to: Solving the problems of teacher evaluation with peer assistance and review. *American Journal of Education*, 113 (3):479-508.
- Heinrich, C. & Marschke, G. 2010. Incentives and their dynamics in public sector performance management systems. *Policy Retrospectives*, 29 (1):183-208.
- Heneman, H., & Milanowski, A. 2004. Alignment of human resources practices and teacher performance competency. *Peabody Journal of Education*, 79 (4): 108-125.
- Heneman, H., Milanowski, A., Kimball, S. & Odden, A. 2006. Standards-based teacher evaluation as a foundation for knowledge-and skill-based pay. CPRE Policy Brief No. RB-45. Philadelphia: Consortium for Policy Research in Education.
- Jacobsen, S. & Battaglia, C. Authentic forms of teacher assessment and staff development in the US. In D. Middlewood & C. Cardno (Eds.). *Managing teacher appraisal and performance*. New York: RoutledgeFalmer.
- Johnson, S. 2009. How to best add value? EPI Briefing Paper #249. <http://www.epi.org>. <http://www.cpre.org>. Washington, D.C.: Economic Policy Institute.
- Looney, J. 2011. Developing high quality teachers: teacher evaluation for improvement. *European Journal of Education* 46 (4):440-455.
- McLaughlin, M. 1990. Embracing contraries: Implementing and sustaining teacher evaluation. In J. Millmand & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers*:403-415. Newbury Park, CA:Sage.

- The Met Project. 2013. Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study. <http://www.metproject.org/reports.php>. Bill and Melinda Gates Foundation.
- Milanowski, A. 2004. The relationship between teacher performance evaluation scores and student achievement: evidence from Cincinnati. *Peabody Journal of Education*, 79 (4):33-53.
- Milanowski, A., Kimball, S., & White, B. 2004. The relationship between standards based teacher evaluation scores and student achievement: replication and extensions at three sites. Consortium for Policy Research in Education. Madison: Wisconsin Center for Education Research.
- Milanowski, A. 2011. Strategic measures of teacher performance. *The Phi Delta Kappan*, 92(7): 19-25.
- Goe, L., Biggers, K., & Croft, A. 2012. Linking teacher evaluation and professional development: Focusing on improving teaching and learning. <http://tqsource.org>. Washington, D.C.: National Comprehensive Center for Teacher Quality.
- Natriello, G. 1990. Intended and unintended consequences: Purposed and effects of teacher evaluation. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers*:35-45. Newbury Park, CA:Sage.
- Odden, A. 2004. Lessons learned about standards-based teacher evaluation systems. *Peabody Journal of Education*, 79(4):126-137.
- Proschaska, J.O. 2008. Decision making in the transtheoretical model of behavioral change. *Medical Decision Making*, 28:845-849.
- Proschaska, J.M., Proschaska, J.O., & Levesque, D. 2001. A transtheoretical approach to changing organizations. *Administration and Policy in Mental Health*, 28 (4):247-261.
- Rockoff, J. & Speroni, C. 2011. Subjective and objective evaluations of teacher effectiveness. *The American Economic Review*, 100 (2): 261-266.
- Rockoff, J., Staiger, D., Kane, T. & Taylor, E. 2010. Information and employee evaluation: evidence from a randomized intervention in public schools. NBER Working Paper 16240. <http://www.nber.org/papers/w16240>. Cambridge: National Bureau of Economic Research.
- Sahlberg, Pasi. 2010. *Finnish lessons*. New York: Teachers College Press.
- Springer, M. & Gardner, C. 2010. Teacher pay for performance: context, status, and direction. *The Phi Delta Kappan*, 91 (8):8-15.
- Slotnik, W, Smith, M. & Guodong, L. 2013. *Focus on Rhode Island: student learning objectives and evaluation*. Boston: Community Training and Assistance Center.
- Strauss, V.2014. Statistician slam popular teacher evaluation method. <http://www.washingtonpost.com>. *The Washington Post*, April 13.
- Taylor, E., & Tyler, J. 2012. The effect of evaluation on teacher performance. *American Economic Review*, 102 (7):3628-3651.
- Tucker, P., Stronge, J., Gareis, C. & Beers, C. 2003. The efficacy of portfolios for teacher evaluation and professional development: do they make a difference?. *Educational Administration Quarterly*, 39: 572-602.
- Weiner, R. & Jacobs, A. 2011. *Designing and Implementing Teacher Performance Management*. <http://aspeninstitute.org/education>. Washington, D.C.: Aspen Institute.

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. 2009. The widget effect. <http://tntp.org>. The New Teacher Project.

Yuan, K., Le, V., McCaffrey, D., Marsh, J., Hamilton, L., Stecher, B., & Springer, M. Incentive pay programs do not affect teacher motivation or reported practices: results from three randomized studies. *Educational Evaluation and Policy Analysis*, 35 (1): 3-22.