

Teacher Evaluation: Improving the Process

Karen Phillips, Rose Balan, & Tammy Manko

Authors' Contact Information

Dr. Karen Phillips
52 Shawger Rd, Denville, NJ, USA 07834
Phone: 862.684.3169
Email: karen9679@aol.com

Dr. Rose Balan
1649 Fir Rd, Merritt, BC, Canada, V1K 0A3
Phone: 250.378.4772
Email: rozamari54@gmail.com

Dr. Tammy Manko,
186 Mesa Drive, Freeport, PA, USA 16229-2410
Phone: 412.289.2705 [mobile]
Email: tammymanko@yahoo.com

Abstract:

The association between quality of teaching and student achievement has placed an increased focus on teacher evaluation at all levels as a means of assuring educational improvement. Historically, the evaluation processes in use have been ineffective and recently many changes have been suggested in the evaluation process. Presented here is a discussion of issues impacting current evaluation practices, suggestions for improvement, and a framework that may be used to guide teacher evaluation. The instrument focuses on five elements essential to good teaching: planning and preparation, the classroom environment, instructional delivery, professionalism, and collaboration and partnerships. Our evaluation framework may be used as is or as a starting point for developing an evaluation tool specific to a particular educational jurisdiction.

Key Words:

Evaluation, drive-by evaluation method, Teacher Advancement Project (TAP), Peer Assessment and Review (PAR), Value-added Model (VAM), Plan Do Study Act (PDSA), Danielson Framework.

Introduction

Evaluation is summative in nature, based on predetermined criteria, observation-based, and involves data collection often with the use of an evaluation instrument (Nolan & Hoover, 2011). Effective teacher evaluation is essential for verifying and maintaining high quality instruction and student learning, ensuring that goals and objectives are achieved, providing a focus for instructional improvement, and holding educators accountable for their instruction. Given the scrutiny being applied to the K-12 and post-secondary educational systems, critical consideration must be paid to evaluation methods, and discussions must occur at the organizational level to ensure that the needs of educators are not subsumed by the needs of policy-makers. Current evaluation measures are fraught with problems in both design and implementation, and a more effective evaluation system is needed (Weingarten, 2010). Due to the importance placed on teacher quality, an effective evaluation model is imperative. Our goals are twofold: to examine current and past practice and to use our findings to inform an improvement strategy. A literature review provided the necessary and appropriate insight for determining effective and ineffective evaluation practices. Such insight led to the development of an evaluation model. The model can be used to advance teacher effectiveness in the classroom or as a guide to self-evaluation.

Current Evaluation Practices

Although teacher evaluation is mandated in most educational jurisdictions in K-12 and post-secondary, evaluation processes have focused more on accountability-based political agendas rather than on practices meant to improve teaching and learning (Holland, 2005). Two influential works, Toch and Rothman's (2008) *Rush to Judgment*, and Weisberg, Sexton, Mulhern, and Keeling's (2009) *The Widget Effect* laid the issue of educational problems squarely on the shoulders of poor evaluation strategies that fail to provide meaningful information, are based on binary rating scales, and do little to distinguish exemplary teaching from poor teaching.

The increased focus on teacher evaluation in K-12 has elevated to post-secondary where the process becomes even more complex. With rising tuition fees and the need for increased worker skills in a competitive global market, accountability had risen to the forefront. Compounding the issue is the rise in full-time non-tenure track positions, where competition for full-time tenure track positions has risen sharply, coupled with evaluation systems having become more complex and formal (Huber, 2002). Often the problem is not that post-secondary staff members are evaluated, but rather that a plethora of data are collected without an effective system of determining the value and application of those data (Buller, 2012).

Educational authorities have significant leeway in how evaluations are conducted and how frequently evaluations will be performed (Bouchamma, 2005; Kennedy, 2008; Santiago & Benavides, 2009). Current classroom teacher evaluations in K-12 typically involve a single, fleeting classroom observation by a supervisor based on a checklist of teacher behaviors and classroom conditions, which Toch (2008) refers to as the Traditional Drive-by Evaluation. Range, Scherz, Holt, and Young (2011) found that principals relied on the classroom walk-through as a means of gathering information to

serve as a starting point for initiating dialogue with teachers; principals were, however, most frustrated by a lack of time to do a thorough evaluation, inadequate assessment instruments, and teachers' reluctance to change. While a paucity of evaluation data may plague the K-12 system, in higher education a surfeit of information complicates the process. At the post-secondary level information may be gathered through student surveys; course quizzes, tests, and composite examinations; self-studies and growth plans; course evaluations completed by students and academic peers, administrator evaluations; and student and graduate satisfaction surveys (Buller, 2012). The lack of educational focus coupled with the inconsistencies in practice has made the evaluation process, as it is, ineffective. Compounding the issue is a lack of accountability, staffing practices devoid of incentives to make teachers care, teacher union ambivalence, and an over-reliance on credentials that may not correlate with teacher quality (Toch, 2008).

Danielson (2010) purported that many teachers view evaluation as a meaningless exercise to be endured and that to be effective, evaluation must provide accurate assessment, be meaningful, and engage teachers in dialogue about excellence in professional practice (Weingarten, 2010). Similarly, Marshall (2005) described the process by which most teachers are supervised and evaluated as "inefficient, ineffective, and a waste of time" (p. 727). Teachers' unions at all levels are reluctant to support evaluation systems that have the potential to negatively impact their members (Toch & Rothman, 2008). Union officials have not pushed for evaluation reform fearing that such measures could result in job loss or a pay scale differential based on merit (Toch & Rothman, 2008). Marshall (2005) suggested that current evaluation systems need streamlining and must be linked to strategies that will improve teaching and learning. For the process to be effective teachers and teachers' unions must see the benefit.

An additional area of concern is that evaluations are subjective in nature with no clear definition of what constitutes good teaching (Danielson, 2010; Weingarten, 2010). The issue of subjectivity is compounded when students participate in the process. While student surveys provide meaningful data about instructor efficacy, the subjective nature of students' responses means that the data only informs one facet of the evaluation process (Chisholm, Hayes, LaBrecque, & Smith, 2011; Evans & McNelis, 2000). Students may personalize their responses; when grades are high university teachers are viewed more favorably than instructors who grade using a more rigid scale (Evans & McNelis, 2000). Other factors limiting the efficacy of student surveys in the evaluation process are selection bias, timing of the surveys, student engagement or lack thereof, and the students' limited experience (Kelly, 2012).

Teacher evaluation in post-secondary education is further complicated by values that are promulgated in specific disciplines, concerns of one discipline that may be poorly understood by another, and by competing values between teaching, research, and service (Chisholm et al., 2011). While each of these values is important in academia, controversy arises when research or service takes precedence over instructional efficacy (Chisholm et al., 2011). Research and service may be easier and less time-consuming to assess than instruction, which typically occurs in isolated classrooms (Chisholm et al., 2011). Although teaching should be the main focus of post-secondary education evaluations, the emphasis on research and scholarship may

adversely influence this process by awarding tenure/promotion based on the number of publications rather than on teaching quality.

Discussions regarding teacher evaluation, instructional improvement, or rewarding teaching excellence are often contentious because various participants conceptualize the meaning of these terms differently (Kennedy, 2008). According to Kennedy (2008), educational policy makers must prioritize those teacher characteristics they consider to be essential so that appropriate assessments can be implemented. Santiago and Benavides (2009) indicated that this process has begun in several states and performance criteria have been drafted to use in performance-based evaluation rubrics. Goldrick (2002) claimed that developing an assessment framework is essential; without it, "evaluation lacks clear purpose and fails to provide policymakers and school leaders with consistently reliable and useful information about what is happening in the classroom" (p. 3). The use of such frameworks requires clearly defined standards and ongoing training on the part of evaluators (Burnett, Cushing, & Bivona, 2012).

The definition of teaching must not only focus on instructional elements, but also on behind the scenes planning, interaction with colleagues and parents, and professional growth and development (Danielson, 2010; Santiago & Benavides, 2009). All stakeholders must have a shared understanding of good teaching (Danielson, 2010). Once a clear definition of good teaching has been adopted and understood, rather than a binary rating system based on satisfactory or unsatisfactory performance, levels of performance should be established and considered when rating a teacher's performance (Danielson, 2010; Santiago & Benavides, 2009).

A standardized evaluation approach may mitigate many of the factors that result from a process that is inconsistently approached and broadly interpreted. Toch (2008) identified the Teacher Advancement Program (TAP) as one model of a standards-based approach. The TAP approach is based on a model created by Danielson and considers three aspects of teaching: designing and planning instruction, the learning environment, and instruction (Toch, 2008). This model also contains a rubric against which teacher performance may be compared (Toch, 2008). A peer evaluation system is used whereby teachers trained in evaluation participate in evaluating lessons and portfolios as part of the supervisory process. Toch (2008) described the Toledo Peer Assistance and Review Program, which uses a system of consulting teachers who have been released from the classroom specifically for this purpose. The consulting teachers are each responsible for a dozen teachers whom they evaluate based on subject knowledge, degree of professionalism, classroom management, and teaching. The consultants then meet with members of a review board who vote on each teacher's status. The principal plays an adjunct role in this process but leaves the actual evaluation to the consulting teachers (Toch, 2008).

Danielson's (1997) Framework for Teaching has become the gold standard used in many educational jurisdictions with modifications to fit the unique situation specific to that school or district. That framework formed the foundation of the Pennsylvania Evaluation System on which the evaluation tool presented here is based (see Appendix). Danielson's model initially included four domains: Planning and Preparation, The Classroom Environment, Instruction, and Professional Responsibilities. Each domain has five specific components, and each component has specific elements.

Having a framework ensures a high standard of practice, exemplifies the complexity of teaching, offers a roadmap for teachers new to the profession, and provides a common language to enable clear discourse related to the teaching profession (Danielson, 1997). Several revisions have replaced Danielson's original Framework for Teaching, the most recent of which occurred in 2013 where language was simplified to reduce ambiguity, components were eliminated in domains one and four, and language was brought in line with the requirements of the Common Core standards.

A number of researchers advocate that student achievement should be incorporated into teacher evaluation (Burnett, Cushing, & Bivona, 2012; Goldrick, 2002; Marshall, 2005; Stronge, Ward, Tucker, & Hindman, 2007); however, the importance placed on assessment data is controversial. Toch (2008) warned that the results of standardized tests should not be used in teacher evaluation processes because of the tests' narrow focuses and failure to test higher-level cognitive skills. Such tests may disadvantage more creative teachers and reduce their teaching repertoire (Toch, 2008). The validity of using student achievement has been called into question, given that many elements affect student achievement including: factors related to the student such as socio-economic status, mobility, availability of home support, peer culture, prior experiences and teachers; and those related to the school such as class size, available student support and learning resources, and specific assessment instruments used (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2011). To link achievement with teacher evaluation practices, Marshall (2005) recommended that teacher teams be created to: develop and share a common definition of good teaching; develop common lesson plans and assessments of those lessons; and conduct interim peer assessments. Marshall (2005) further recommended that time be provided each week for teacher teams to meet and that a learning culture be created in each school. By sharing evaluations with teachers, the principal's time could be used more efficiently to provide shorter more frequent evaluations with an emphasis on high-quality feedback and consideration of learning gains made by students. Final evaluations should be based on these short observations using rubrics to provide focus (Marshall, 2005).

To overcome the non-teaching aspects of student achievement, a Value-Added Model (VAM) has been suggested (Burnett, Cushing, & Bivona, 2012; Darling-Hammond et al., 2011; Glazerman, et al., 2010; Santiago & Benavides, 2010). The value-added approach, in some respects, represents a significant improvement in teacher evaluation because of its ability to separate certain factors that impact learning (Glazerman et al., 2010; Santiago & Benavides, 2010). While VAMs mitigate the effects of standardized test scores to a degree, other problems are created, and more research is needed to determine their efficacy (Darling-Hammond et al., 2011). Specifically, VAMs may be unstable; may be affected by certain students assigned to a particular class such as when student assignment is specific and non-random; and may not effectively separate the variables that affect student scores (Darling-Hammond et al., 2011; Marzano & Toth, 2013). VAMs require the addition of a great deal of data across subjects and time and may add to the expense of the evaluative process, thus making them less than ideal (Santiago & Benavides, 2010).

Evaluation that includes student achievement is supported by Marzano and Toth (2013), who recommended that teacher evaluation should be based on: (a) student

growth that is triangulated by multiple measures, over time, and in both formal and day-to-day settings; (b) evaluation data that are triangulated and collected over time with frequent observations using multiple observers; (c) planning and preparation; (d) a realistic distribution of skills among teachers; (e) strategies that improve teaching by providing support where needed; (f) a hierarchical model which evaluates principals and superintendents in their abilities to evaluate and support teachers. While progress has been made in teacher evaluation, the discussion is ongoing, and occurring beyond the school and community levels.

In response to the need for more effective means of evaluating new teachers, several states have adopted a set of standards aligned with the Common Core standards (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2014). These standards are based on classroom evidence and show the knowledge, skills, and understandings of educational theories and methods with greater stability than value-added approaches (Darling-Hammond et al., 2014). Some districts have adopted the standards-approach and use it for all teacher evaluation. Additionally, under the auspices of the Bill & Melinda Gates Foundation, new standards-based tools have been developed to assess teaching effectiveness. The IMPACT system, developed in 2009, is another comprehensive system that incorporates student achievement, classroom observation, teacher professionalism and collaboration (Rucinski & Diersling, 2014). These new evaluation methods work best when other processes are in place such as frequent evaluations by trained evaluators, ongoing professional development, and when supports are in place to assist individual teachers (Darling-Hammond et al., 2014).

Evaluation: A Model Analysis

Holland (2005) expanded on the idea of feasibility of evaluation that is anchored in meeting the needs of administrators and policy makers. Thus, the implication is that perhaps the problem with evaluation programs is not with the standards themselves, but rather the interpretation and application of the standards to the benefit of educational leaders and policy makers versus teachers (Holland, 2005). While attempting to make evaluation a meaningful process for teachers, one cannot ignore legislators' valid concerns regarding holding teachers accountable for their work (Holland, 2005). Rather than focusing on standards of performance, Holland (2005) presented six standards of the evaluation process:

1. A differentiated evaluation approach based on the teacher's level of development;
2. A collaborative approach that considers the teacher's goals, progress towards achieving those goals, and the implications of that progress;
3. Multiple observations by more than one individual;
4. Formative as well as summative evaluation, with greater emphasis on the formative approach;
5. A focus on school goals as well as individual goals; and
6. Transparent evaluation policies that are clearly articulated and defined, and understood by teachers and administrators.

Holland (2005) claimed that these standards serve the interests of teachers “as professionals and as moral agents who earn public trust by subjecting their practice to scrutiny” (p. 74).

Teacher evaluation is based on the premise that the evaluation process will improve teacher efficacy and increase student achievement (Marshall, 2005). An effective evaluation system is critical to educational improvement and reform (Stronge, 2006). To be effective, an evaluation system must (a) benefit the teacher and the school community, (b) promote effective communication, (c) create and sustain a positive evaluation climate, (d) use multiple sources for data collection, and (e) provide valid and reliable information (Stronge, 2006). Such a system uses principal and peer observations and feedback, an examination of student work, self-assessment, and student achievement on appropriate curriculum-based assessments (Goldrick, 2002). In short, an effective teacher evaluation system is one that works to promote teacher learning and growth which in turn results in improved student learning (Danielson, 2012).

Establishing a clear conceptualization of teacher quality is a starting point in this process (Kennedy, 2008). Having standards of practice provides the criteria against which teacher development can be compared (Toch, 2008) and standards for the evaluation process will ensure consistency within and between educational jurisdictions (Holland, 2005). Using peers to evaluate may be an effective alternative to principal evaluation, may increase the amount of observations that can be conducted, and may provide varying points of view (Bouchamma, 2005; Marshall, 2005; Toch, 2008). Collegiality is developed when peers work together collaboratively, and student achievement is enhanced (Marshall, 2005). While reforming teacher evaluation may be a major undertaking the results are worth the effort (Goldrick, 2002).

The authors have examined three distinct assessment models, the Pennsylvania Teacher Evaluation, the Plan Do Study Act (PDSA) model (Glickman, Gordon, & Ross-Gordon, 2007), and the Clinical Supervision model (Glickman et al., 2007; Stronge, 2006). Using information from these assessment approaches, along with the newly redesigned Danielson Framework (2013) a new evaluation form was devised. The new evaluation form is presented here along with a description and a rationale for its use detailing why this model is an effective evaluation tool.

Given our reference to the Pennsylvania Teacher Evaluation, it is important to note that Pennsylvania revamped its teacher evaluation system and piloted the new system in about 120 schools during the 2011-2012 academic year (Fioriglio, 2012). The old system used a pass-fail rating for teachers and did not require feedback to teachers, while the new system incorporates ratings of distinguished, proficient, needs improvement, or failing; and teachers receive recommendations on improvement, regardless of the ratings received (Fioriglio, 2012). The new approach not only includes classroom observation (i.e., teacher preparation, instruction, and general classroom atmosphere), but also student performance (i.e., test scores, graduation rates, and attendance) (Fioriglio, 2012). The new evaluation system was officially incorporated during the 2013-2014 school year.

Description of Evaluation Tool

The National Board for Professional Teaching Standards (1989) developed five core propositions to identify and recognize teachers who “effectively enhance student learning and demonstrate high levels of knowledge, skills, dispositions, and commitments” (p.12). With these five core propositions in mind, a framework for evaluation was designed that consists of five elements: planning and preparation, classroom environment, instruction, professionalism, and collaboration and partnership. Each element represents a specific function that teachers should know and be able to perform based on a 5-point scale indicating: exemplary, proficient, adequate, limited, or not in evidence performance, dependent upon whether the teacher is an experienced or novice teacher. Experienced teachers would be expected to obtain ratings of adequate to exemplary, while the expectations for novice teachers would consider their levels of experience. The rating indicators offer clearer articulation of instructor performance than simply checking *satisfactory* or *unsatisfactory* which may not clearly delineate a range of skills between the two endpoints. Levels of performance allow for an evidence-based judgment to be made, thus personalizing the process (Danielson & McGreal, 2000). Adding the rating indicators eliminates ambiguity while also allowing for rating teachers of all experiences levels, from novice teachers to seasoned or tenured teachers.

Effective evaluation instruments must have reliability and validity (Glickman et al., 2007). To ensure validity, the authors considered the appropriateness of their evaluation instrument in terms of whether the selected criteria contribute to effective teaching. Valid evaluation instruments include criteria that are characteristic of good teaching, and eliminate those that do not contribute; such criteria are weighted in terms of their significance to instruction (Glickman et al., 2007). Reliability is provided in the objective and evidentiary nature of the criteria. Reliability is also ensured by having trained evaluators and pre-evaluation and post-evaluation conferencing to ensure that the evaluator and teacher clearly understand the evaluation terms (Glickman et al., 2007).

Rationale

Based on the review of current evaluation models, the most significant and applicable elements of each model were identified and integrated into a new evaluation tool (see Appendix). Existing literature and research relevant to the design and implementation of an evaluation model were applied as the evaluation form was devised and refined. By taking the best components from the research, a comprehensive evaluation was created comprising five distinct elements that are interconnected in a manner that allows for thorough examination of a teacher’s performance.

Our initial evaluation tool was developed in 2009 with revisions in 2012 and 2014. The PDSA, the clinical supervision model, and Danielson’s 1997 framework formed the basis of the initial iteration of our tool. Initially the PDSA model was selected because of its familiarity to one of the authors. The Pennsylvania model is representative of the evaluation system in use in many educational areas, and includes a process of collecting data to improve the evaluation process itself, as well as for the greater goal of improving teaching. The PDSA model was based on a process of collecting data for evaluation purposes, evaluating instruction, analyzing the data to determine areas that need to be improved, and providing relevant professional development. Zatyński (2012)

supported the notion that the two main purposes of evaluation are quality assurance and professional development. The clinical supervision model is a cyclic model consisting of three phases: pre-observation conference, observation, and post-observation conference (Stronge, 2006). All three phases are necessary to the process to effectively assess teachers and facilitate self-assessment of one's teaching (Stronge, 2006). Engaging teachers in conferences to discuss instructional efforts, successes, and areas needing improvement is critical to evaluative success.

Danielson's framework focuses on four domains, which are included in this evaluation tool: (a) planning and preparation, (b) classroom environment, (c) instruction, and (d) professional responsibilities (Danielson, 1997). Aligned with Danielson's (1997) model, a fifth domain was added, collaboration and partnership. The authors felt strongly that collegiality was essential to the current professional practice with its emphasis on the formation of professional learning communities and collaborative practice. While Danielson's 1997 model formed the basis of our evaluation tool, Danielson's 2011 and 2013 revisions were consulted to ensure that our initial evaluation tool was in line with current educational practice.

Planning and Preparation

Planning and preparation are crucial to effective instruction. Orlich, Harder, Callahan, Trevisan, and Brown (2012) list planning as a hallmark of teaching. Preparation enables teachers to be more organized and coordinate their classroom activities with efficiency. Substantial time needs to be devoted to planning and preparation for instruction to be effective (Orlich et al., 2012). Planning incorporates the development of academic goals along with strategies for determining when goals have been achieved and reflect the aims of the teacher as well as those of the school. Goals must be in place for progress to occur. Goal theory explains in part why students are willing to take academic risks and what instructional characteristics factor into classroom involvement (Meyer & Turner, 2006). Balance must exist between the goals of an individual teacher and those of the school (Stronge, 2006). Effective evaluation systems promote progress and support this balance (Stronge, 2006). Marzano and Toth (2013) claimed that focusing exclusively on what happens in classrooms and ignoring what goes on before instruction does little to enhance professional growth. Thus, the element of planning and preparation is appropriate to an instructional evaluation model.

Classroom Environment

The second element, classroom environment, is relevant to the instructional evaluation process in that the creation of a positive learning climate is directly under the teacher's control (Halawah, 2005; Hallinger, 2003; Nor, Pihie, & Ali, 2008). Classrooms are unique environments that offer a myriad of possibilities to foster improved practice and facilitate student learning (Meyer & Turner, 2006). A trusting environment promotes risk taking, deep learning, and engagement where process and not product is emphasized (Robinson & Kakela, 2006). Trust is an important component in relationship building; when teachers build relationships with their students they have fewer discipline problems (Gregory & Ripski, 2008). To promote trust, teachers show respect for all students, value diversity, and encourage imagination and creativity (Robinson & Kakela,

2006). Students' emotions play a role in their learning and the environment influences emotion. Pintrich and Schunk (2002) synopsis Ford's Responsive Environment Principle with these points: (a) the alignment between a person's goals and the goals of the classroom, (b) the teacher's responsiveness to the student's competencies, (c) the provision of realistic and appropriate tasks, and (d) support for an emotional climate that fosters trust among teachers and students (p. 378).

Lazarus' Relational-Motivation-Cognitive Theory further supports the importance of creating a safe, purposeful, and equitable learning environment since Lazarus defined emotions as "emerging through person-environment interactions that change over time and situations" (Meyer & Turner, 2006, p. 378). Contemporary theorists concur with the notion that emotions, which influence learning and behavior, are contextualized (Meyer & Turner, 2006). Teachers' attitudes and behaviors definitely affect students' perceptions and performance (Meyer & Turner, 2006). Thus, the classroom environment category is an appropriate aspect for consideration when conducting teacher evaluation.

Instruction

Instruction is the third element considered in this evaluation tool. As suggested in the second element, teachers' behavior and underlying beliefs and values impact the classroom and the students. Effective instruction enables students to learn regardless of learning differences (Darling-Hammond, 2012). Instructional quality is influenced by the fit between a teacher's credentials and the subject matter he or she is required to teach (Darling-Hammond, 2012). Pedagogy, using effective teaching and assessment strategies, understanding learners and their learning needs, and motivation and supporting students in difficulty are essential to effective instruction (Darling-Hammond, 2012). Ensuring that teachers are using multiple instructional approaches is significant to making sure students are progressing. Given that the goal of evaluation is instructional improvement (Glickman et al., 2007), evaluating instruction is critical.

Professionalism

Professional development is an essential component of educational improvement in both K-12 and post-secondary educations (Balan, Manko, & Phillips, 2011). Meaningful professional development facilitates professional growth and enhances scholarly practice (Balan et al., 2011). Researchers support the need for ongoing professional development efforts (Glickman et al., 2007; Stronge, 2006). Measuring professional performance and efforts is, therefore, relevant to advancing teachers' skills and knowledge, thus improving instruction. Involving teachers in professional development decisions and goal setting is motivating to teachers and improves instructional efficacy (Costa & Garmston, 2002). Thus, professionalism is a necessary and appropriate element to include in an evaluation program.

Collaboration and Partnership

Collaboration is a method of providing support in a mentoring or collegial relationship where teachers share ideas, concerns, strategies, and gather data that will inform practice (Costa & Garmston, 2002). Collaboration offers the benefit of obtaining other

points of view; consider the axiom: *Two heads are better than one*. The literature supports the benefits of teacher collaboration on student learning and as a tool to enhance teacher quality and support improvement (Costa & Garmston, 2002; Stronge, 2006). Teaching, one of the most private professions, has primarily existed in isolation (Costa & Garmston, 2002; Stronge, 2006). Few people outside the classroom are privy to what goes on behind closed doors. Current school environments, however, are changing and an important focus of teaching is working with other teachers, support staff, and community members (Garmston & Wellman, 1999). Teachers can learn from one another no matter their length of teaching experience or level of expertise. A teacher's ability to collaborate with others transfers into classroom practice (Costa & Garmston, 2000; Stronge, 2006). More than 90% of teachers report that their teaching has been positively influenced by their teaching colleagues (Darling-Hammond, 2012). An evaluation system must recognize a teacher's contribution to the school as a whole, and may include: specific knowledge and skills, participation in shared instructional practices or specific student support programs, participation in collegial learning, and school improvement (Darling-Hammond, 2012). The value of successful collaborations to improved teaching practice makes this element an important consideration in teacher evaluation.

Although our framework was developed for use by trained evaluators, an additional recommended purpose is to guide teachers' personal self-evaluation and practice. Mielke and Frontier (2012) claimed that the most effective use of evaluation instruments is to empower teachers to self-assess and diagnose problem areas prior to formal assessments. A school culture that supports teacher self-assessment and growth promotes the belief that each individual within the school is also a learner (Mielke & Frontier, 2012). Research supports the notion that teachers are extremely proficient at identifying their own strengths and weaknesses; enabling them to use an evaluation instrument to self-reflect means that they can actively work toward achieving personal goals in a process of continual improvement and not just sporadically after formal assessment (Mielke & Frontier, 2012). Rather than strictly measuring teaching proficiency, to be effective, an evaluation system must empower teachers to be active participants in their own professional growth (Mielke & Frontier, 2012).

Implementation

No matter how effectively an evaluation system has been developed, educational improvement will not occur unless significant buy-in exists from the teachers (Fullan, 2011; White, Cowhy, Stevens, & Spote, 2012). Moreover, evaluation processes that reward or punish teachers, focus on individual teacher improvement only rather than looking at the needs of the school, and strategies that are fragmented rather than holistic will not be successful (Fullan, 2011). System-wide change will only occur if capacity building, engagement, and trust-building are part of the process (Fullan, 2011; White et al., 2012). Important to the process is the perception of a safe environment, goal-setting, collaboration, and peer support (White et al., 2012). Teachers need to be part of the decision-making process. The teachers' voice must be heard, and their feedback incorporated into the implementation process (White et al., 2012). Trust-building must be part of the process, as well as creating a culture of learners with a shared vision focused on organizational improvement (White et al., 2012). Through

collaboration and professional learning communities, ongoing discussion must occur. Teachers must feel they are the driving influence behind the evaluation process, trust the process, and take ownership for personal growth and development. Teachers must have a degree of choice in how and by whom they will be evaluated.

The current political climate is such that a push-pull dichotomy exists between education ministries, policy-makers, and school administrators on one side, and teachers and their unions on the other. On the one side a need exists for accountability; maintaining high standards of teacher quality is part of that need. On the other side, a lingering sense of mistrust and fear that evaluation processes will negatively impact job security and teacher autonomy prevents teachers from willing participation in evaluation. Given the increasing push for accountability that exists at all educational levels, it is unlikely that formal accountability-based evaluations will be discontinued within the near future.

Although improved evaluation practices are meant to improve student learning, care must be taken to ensure the new process is initiated systematically in such a way as to ensure success. White et al. (2012) recommended that the process start early but gradually to reduce anxiety and build understanding. Time must be allocated for training evaluators whether they are peer evaluators or principals. Training must occur early on in the process and continue throughout implementation (White et al., 2012). Well-trained observers can more readily discern teacher growth (White et al., 2012). Aligning teacher evaluation reform measures with other school initiatives can prevent the process from bogging down (White et al., 2012). Time must be provided for principals and peer evaluators to complete evaluations in a timely fashion. More frequent evaluations promote improvement, reduce anxiety and build trust and more accurately assess growth (White et al., 2012). The use of peer evaluators reduces the burden on the principal and allows for more frequent observations. Critical to improved evaluation practices are ongoing supervision of instruction and coaching.

Conclusion

The quality of education depends upon the quality of the teachers in the classroom (Koops & Winsor, 2005; Peterson, 2004). As such, evaluation of teachers is necessary to promote professional excellence by providing a basis to review teachers' performance (Koops & Winsor, 2005). Evaluation can be one of the most effective tools to improve the quality of instruction in schools because it promotes instructional improvement and increased student learning (Koops & Winsor, 2005); yet the evaluation system still in use in most schools falls short of this goal. Improvements in instruction will translate into improved student learning, which ultimately is the intended outcome of teacher evaluation.

An evaluation instrument or framework provides a rubric by which teaching can be assessed and is essential to the process. An effective evaluation system; however, should be much more than that. According to Marzano (2012), an effective teaching evaluation system is one that is comprehensive and specific, includes a developmental scale, and acknowledges and rewards teachers' growth. The purpose of teacher evaluation should not only be to measure teacher growth but also to promote teacher development (Marzano, 2012; Mielke & Frontier, 2012).

Our evaluation tool is one example of a comprehensive evaluation framework and may be used as is or modified to fit the particular needs of the organization adopting it. This instrument is formulated based on other evaluation instruments currently in use and on the best information available in the field of education. It is the authors' hope that the information provided will continue the teacher evaluation dialogue and aid educators in reaching their goals of improving teaching and thus, learning for every student.

References

- Balan, R.M., Manko, T.P., & Phillips, K.F. (2011). Instructional improvement through professional development. *Transformative Dialogues: Teaching & Learning Journal*, 5 (2), 1-18.
- Bouchamma, Y. (2005). Evaluating teaching personnel: Which model of supervision do Canadian teachers prefer? *Journal of Personnel Evaluation in Education*, 18(4), 289-309. doi: 10.1007/s11092-007-9025-8
- Buller, J. L. (2012). *Best practices in faculty evaluation: A practical guide for academic leaders*. San Francisco: John Wiley and Sons.
- Burnett, A., Cushing, E., & Bivona, L. (2012). *Use of multiple measures for performance based compensation*. Retrieved from <http://0-cecr.ed.gov.opac.acc.msmc.edu/>
- Chisholm, M., Hayes, E. J., LaBrecque, S., & Smith, D. (2011). The role of faculty evaluation in transformative change. *The Journal of Faculty Development*, 25(1), 36-42. Retrieved from <http://search.proquest.com/docview/1095350621?accountid=35812>
- Costa, A. L., & Garmston, R. J. (2002). *Cognitive coaching: A foundation for renaissance schools* (2nd ed.). Norwood, MA: Christopher-Gordon Publishers, Inc.
- Danielson, C. (2013). *The framework for teaching: Evaluation instrument*. (2013 ed.). Retrieved from <http://danielsongroup.org/framework/>
- Danielson, C. (2012). *Developing effective teacher evaluation systems: conversation with Charlotte Danielson*. Retrieved from <http://www.ed.gov/Teacher-Evaluation-Systems>
- Danielson, C. (2010). Evaluations that help teachers learn. *Educational Leadership*, 68(4), 35-39.
- Danielson, C. (1997). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: ASCD.
- Danielson, C., & McGreal, T. (2000). *Teacher evaluation: To enhance professional practice*. Alexandria, VA: ASCD.
- Darling-Hammond, L. (2012). *Creating a comprehensive system for evaluating and supporting effective teaching*. Stanford, CA: Stanford Center for Opportunity Policy in Education.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E. H., & Rothstein, J. (2011). *Getting teacher evaluation right: A background paper for policy-makers*. American Education Research Association. Retrieved from <http://www.njea.org/njea-media/pdf/GettingTeacherEvaluationRight.pdf?1341989791440>
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E. H., & Rothstein, J. (2014). Evaluating teacher evaluation. *Phi Delta Kappan*. Retrieved from http://www.edweek.org/ew/articles/2012/03/01/kappan_hammond.html
- Evans, M. D. D., & McNelis, P. D. (2000). *Student evaluations and the assessment of teaching: What can we learn from the data?* Retrieved from http://faculty.georgetown.edu/evansm1/wpapers_files/evalstudy.pdf
- Fioriglio, T. (2012). New Pennsylvania teacher evaluation process coming in 2013. *The Times Herald*. Retrieved from

<http://www.timesherald.com/article/20120722/NEWS01/120729899/new-pennsylvania-teacher-evaluation-process-coming-in-2013>

- Fullan, M. (2011). Introducing the drivers for whole system reform. *Centre for strategic education*. Retrieved from <http://www.michaelfullan.ca/media/13436787590.html>
- Garmston, R., & Wellman, B. (1999). *The adaptive school: A sourcebook for developing collaborative schools*. Norwood, MA: Christopher-Gordon Publishers.
- Glazerman, S., Loeb, S., Goldhaber, D., Staiger, D., Raudenbush, S., & Whitehurst, G. (2010). *Evaluating teachers: The important role of value-added*. Retrieved from http://www.brookings.edu/~media/Files/rc/reports/2010/1117_evaluating_teachers/1117_evaluating_teachers.pdf
- Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (2007). *Supervision and instructional leadership: A developmental approach* (7th ed.). Boston: Pearson.
- Goldrick, L., (2002). *Improving teaching evaluation to improve teaching quality*. Retrieved from <http://www.teacherleaders.org/node/1172>
- Gregory, A., & Ripski, M. C. (2008). Adolescent trust in teachers: Implications for behavior in the high school classroom. *School Psychology Review*, 37(3). 337-353.
- Halawah, I. (2005). The relationship between effective communication of high school principal and school climate. *Education*, 126(2), 334-345.
- Hallinger, P. (2003, November). Leading educational change: Reflections on the practice of instructional and transformational leadership. *Cambridge Journal of Education*, 33(3), 329-351. doi: 10.1080/0305764032000122005
- Holland, P. (2005). The case for expanding standards for teacher evaluation to include an instructional supervision perspective. *Journal of Personnel Evaluation in Education*, 18(1). 67-77. doi: 10.1007/s11092-006-9009-0
- Huber, M. T. (2002). Faculty evaluation and the development of academic careers. *New Directions for Institutional Research*. (114). Retrieved from <http://edweb.sdsu.edu/bober/montgomery/Article02.pdf>
- Kelly, M. (2012). *Student evaluations of teaching effectiveness: Considerations for Ontario universities*. Retrieved from <http://www.cou.on.ca/publications/academic-colleague-papers/pdfs/ac-paper---student-evaluations-may-23-12>
- Kennedy, M. M. (2008). Sorting out teacher quality. *Phi Delta Kappan*, 90(1). 59-63.
- Koops, J. B., & Winsor, K. (2005). Creating a professional learning community through faculty evaluation. *The Journal of Education*, 186(3), 61-70.
- Marshall, K. (2005). It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan*, 86(10), 727-735
- Marzano, R. (2012). The two purposes of teacher evaluation. *Educational Leadership*, 70(3), 14-19.
- Marzano, R., & Toth, M. D. (2013). *Teacher evaluation that makes a difference: A new model for teacher growth and student achievement*. Alexandria, VA: ASCD.
- Meyer, D., & Turner, J. (2006). Re-conceptualizing emotion and motivation to learn in classroom contexts. *Educational Psychology Review*, 18(4), 377-390. doi: 10.1007/s10648-006-9032-1
- Mielke, P., & Frontier, T. (2012). Keeping improvement in mind. *Educational Leadership*, 70(3), 10-13.
- National Board for Professional Teaching Standards. (1989). *Toward high and rigorous standards for the teaching profession*. Southfield, IL: National board for Professional Teaching Standards.

- Nolan, J. F., & Hoover, L. A. (2011). *Teacher supervision and evaluation: Theory into practice*. Hoboken, N.J.: John Wiley & Sons.
- Nor, S., Pihie, Z., & Ali, S. (2008). Instructional leadership practices of rural school principals. *International Journal of Learning*, 15(7), 231-238.
- Orlich, D., Harder, R., Callahan, R., Trevisan, M., & Brown, A. (2012). Teaching strategies: A guide to effective instruction. Retrieved from: http://books.google.com/books?id=DQ0KAAAAQBAJ&dq=teacher+planning+and+preparation+in+effective+instruction&lr=&source=gbs_navlinks_s
- Pennsylvania State Education Association (PSEA). (2001). *Teacher evaluation*. Retrieved from: http://www.psea.org/uploadedFiles/Publications/Professional_Publications/Advisories/TeacherEvaluationAdvisory.pdf
- Peterson, K. (2004). Research on school teacher evaluation. *NASSP Bulletin*, 88(639), 60-79.
- Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications (2nd ed.)*. Englewood Cliffs, New Jersey: Merrill.
- Range, B. G., Scherz, S., Holt, C. R., & Young, S. (2011). Supervision and evaluation: The Wyoming perspective. *Educational Assessment, Evaluation, & Accountability*. 23(3), 243-265. doi: 10.1007/s11092-011-9123-5
- Robinson, C. F., & Kakela, P. J. (2006). Creating a space to learn. *College teaching*, 54(1), 202-206. doi: 10.3200/CTCH.54.1
- Rucinski, M., & Diersling, C. (2014). America's teacher evaluation system revolution. *Harvard Political Review*. Retrieved from <http://harvardpolitics.com/united-states/americas-teacher-evaluation-system-revolution/>
- Santiago, P., & Benavides, F. (2009). *Teacher evaluation and examples of country practices*. Retrieved from <http://www.oecd.org/dataoecd/16/24/44568106.pdf>
- Stronge, J. H. (2006). Teacher evaluation and school improvement. In J. H. Stronge (Ed.), *Evaluating Teaching (2nd ed.)* (pp. 1-23). Thousand Oaks, CA: Corwin Press.
- Stronge, J. H., Ward, T. J., Tucker, P. D., & Hindman, J. L. (2007). What is the relationship between teacher quality and student achievement? An exploratory study. *Journal of Personnel Evaluation in Education*, 20(3-4), 165-184. doi: 10.1007/s11092-008-9053-z
- Toch, T. (2008). Fixing teacher evaluation. *Educational Leadership*, 66(2). 32-37. Retrieved from <http://www.ascd.org>
- Toch, T., & Rothman, R. (2008). *Rush to judgment: Teacher evaluation in public education*. Retrieved from http://www.educationsector.org/usr_doc/RushToJudgment_ES_Jan08.pdf
- Weingarten, R. (2010). *A new path forward: Four approaches to quality teaching and better schools*. Speech delivered at the National Press Club, Washington, DC. Retrieved from: http://aft.3cdn.net/227d12e668432ca48e_twm6b90k1.pdf
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. (2nd ed.). Retrieved from <http://widgeteffect.org/>
- White, B. R., Cowhy, J., Stevens, W. D., & Spote, S. E. (2012). Designing and implementing the next generation of teacher evaluation systems: lessons learned from case studies in five Illinois districts. The University of Chicago Consortium. Retrieved from <https://ccsr.uchicago.edu/publications/designing-and-implementing-next-generation-teacher-evaluation-systems-lessons-learned>
- Zatynski, M. (2012). *Revamping teacher evaluation*. Retrieved from <http://www.educationsector.org/publications/revamping-teacher-evaluation>

Appendix

Evaluation of Instruction: Observation of Classroom Performance Form

Directions:

Please complete all sections of this form, examining all sources of evidence and remaining cognizant of the aspects of instruction for each of the five elements included in this evaluation. Each element contains a rubric checklist as a guide for documenting the teacher’s performance. The box checked next to the rubric item is an indication of the teacher’s performance: exemplary, proficient, adequate or limited/not in evidence. Upon completion of the observation, assign an overall evaluation of performance, sign the form, and have the instructor sign the form. Level of Expertise (please check one):

Novice (less than 5 years of teaching experience)	Professional (more than 5 years of experience)	Master (10 or more years of experience)	Experienced teacher with a new assignment differing by grade or subject

Element 1: Planning & Preparation:

Review of the teacher’s knowledge of content and pedagogy skills and academic standards including learning benchmarks in planning and preparation; awareness of available technology and resources; and, ability to establish instructional goals, design coherent instruction, and assess student learning.

Evaluation Area	Exemplary 4	Proficient 3	Adequate 2	Limited 1	Not in evidence 0
Demonstrates knowledge of curricula.					
Exhibits knowledge of academic standards/ learning benchmarks.					
Shows knowledge of students and applies this knowledge.					
Demonstrates knowledge and application of learning theory.					
Provides instructional goals that maintain standards and reasonable expectations of students.					
Organizes the material presented logically and sequentially and based on students’ prior knowledge.					
Uses and incorporates available resources, materials and technology.					

Evaluation Area	Exemplary 4	Proficient 3	Adequate 2	Limited 1	Not in evidence 0
Designs instruction to align with teaching objectives; follow a logical sequence, and adapt to meet students' needs.					
Maintains long-term and short-term unit and lesson plans.					
Develops students' creative and critical thinking skills relative to the subject matter.					
Plans for a variety of learning situations: whole class activities, small group, partner, and individual assignments.					
Plans for appropriate assessment of students' learning aligned with instructional objectives and academic standards.					
Incorporates variety in assignments; including written work, projects, displays					
Overall rating for Element 1: Planning and Preparation					

Element 2: Classroom Environment:

Determining if the teacher has established a safe, purposeful, and equitable learning environment is critical to student achievement. Students need to feel valued, safe, respected, and motivated to learn. Element 2 seeks to review teacher-student interaction and the general learning environment.

Evaluation Area	Exemplary 4	Proficient 3	Adequate 2	Limited 1	Not in Evidence 0
Sets clear, reasonable expectations for student achievement, including high standards for student work quality.					
Provides equitable learning opportunities for all students.					
Maintains appropriate interactions between teacher and students as well as among students; demonstrative of strong working relationships and a climate of trust and respect.					

Evaluation Area	Exemplary 4	Proficient 3	Adequate 2	Limited 1	Not in evidence 0
Demonstrates effective classroom management, including clear conduct standards, and follow-through.					
Demonstrates fairness and consistency.					
Uses a variety of resources and technologies.					
Maintains a culture of respect; demonstrates positive discipline strategies.					
Supports beliefs, vision, and mission of the school.					
Encourages student inquiry, expression, and participation.					
Demonstrates flexibility and responsiveness through monitoring and modifying student assignments.					
Overall rating for Element 2: Classroom Environment					

Element 3: Instruction:

Exercising multiple instructional methods speaks to the needs of all students, thus advancing all students’ learning and skill development. Element 3 involves determination of the teacher’s ability to communicate; engage students; employ inquiry, discussion, and reflection strategies; assess students and provide feedback; and, demonstrate flexibility and responsiveness.

Evaluation Area	Exemplary 4	Proficient 3	Adequate 2	Limited 1	Not in Evidence 0
Communicates effectively; provides clear content delivery.					
Encourages student-participation; uses appropriate employment of inquiry, discussion, and reflection strategies.					
Maintains reasonable engagement of students in learning and appropriate pacing of instruction.					
Provides appropriate responses to questions, including clarification and expansion of concepts.					

Evaluation Area	Exemplary 4	Proficient 3	Adequate 2	Limited 1	Not in evidence 0
Gives productive feedback to students relative to their learning.					
Uses summative and formative assessments, both formal and informal, to monitor student progress and learning and to achieve learning objectives.					
Provides opportunities for students to develop metacognition through self-assessment.					
Demonstrates flexibility and responsiveness in fulfilling students' learning needs and goals.					
Engages students in creative and critical thinking exercises.					
Employs differentiated instruction. Modifies teaching style to ensure all students' learning styles and needs are addressed.					
Covers material in an interesting and effective manner.					
Stays on track with intended lesson plan while remaining flexible enough to address unexpected interruptions and setbacks.					
Displays evidence of instructional adaptations based on assessments.					
Provides evidence of student learning *Support with assessment data, including, but not limited to, standardized test scores, curriculum-based assessments, and daily work					
Overall rating for Element 3: Instruction					

Element 4: Professionalism:

Exhibiting professionalism and an interest in professional growth and development is critical to a teacher’s performance. Thus, Element 4 reviews the following: record keeping, communication with all stakeholders, contributions to school and district, and interest in professional development.

Evaluation Area	Exemplary 4	Proficient 3	Adequate 2	Limited 1	Not in Evidence 0
Adheres to district policies relative to attendance, punctuality, and general behavior.					
Knows of applicable policies.					
Demonstrates knowledge of and adherence to professional code of conduct.					
Complies with district policies for record keeping and stakeholder communication.					
Contributes to the work environment in a positive, productive manner.					
Is open and responsive to feedback from supervisor.					
Is open and responsive to feedback from peers/colleagues.					
Exhibits interest in and enthusiasm for work.					
Demonstrates interest in professional development opportunities.					
Attends school and district workshops and seminars.					
Participates in job-related studies, projects, and professional activities.					
Is a member of professional organizations.					
Is a reflective practitioner.					
Overall rating for Element 4: Professionalism					

Element 5: Collaboration and Partnership:

No experts exist in the complex act of teaching. All teachers can learn from one another. Collaboration offers alternative points of view, and as such, the provision of a more balanced, accurate, interpretation of professional practice. Element 5 reviews collaborative activities and professional dialogues.

Evaluation Area	Exemplary 4	Proficient 3	Adequate 2	Limited 1	Not in Evidence 0
Collaborates with other school professionals.					
Provides information, when needed, to appropriate staff and agencies.					
Works effectively with parents and community.					
Draws on school and community resources to benefit students.					
Contributes to the school, community, and/or profession.					
Respects diversity					
Engages in professional dialogues to improve instruction.					
Engages in professional conversations					
Overall rating for Element 5: Collaboration and Partnership					

Sources of Data (Check all that apply)

<input type="checkbox"/> <input type="checkbox"/> Formal Classroom Observations	<input type="checkbox"/> <input type="checkbox"/> Student Achievement Data
<input type="checkbox"/> <input type="checkbox"/> Walk-Through Observations	<input type="checkbox"/> <input type="checkbox"/> Student Work
<input type="checkbox"/> <input type="checkbox"/> Assessment Materials	<input type="checkbox"/> <input type="checkbox"/> Instructional Resources/Materials
<input type="checkbox"/> <input type="checkbox"/> Teacher Conferences/Interviews	<input type="checkbox"/> <input type="checkbox"/> Peer Feedback
<input type="checkbox"/> <input type="checkbox"/> Student Surveys/Feedback	<input type="checkbox"/> <input type="checkbox"/> Other (Please specify.) _____

Additional Notes (Optional)

Overall Justification for Evaluation:

Commendations:

Professional Development Recommendations:

I certify that employee _____ for the period of _____ has been evaluated with an overall level of proficiency that is:

Exemplary	Proficient	Adequate	Limited/ Not in Evidence
------------------	-------------------	-----------------	-------------------------------------

Signature of Principal/Assistant Principal (Evaluator)

Date

Signature of Employee (Evaluatee)

Date

Signature of Superintendent

Date

Note: This form was adapted from Pennsylvania Department of Education and William Paterson University Department of Nursing Peer Evaluation forms, and the components are based on elements of Total Quality Management (TQM) and Plan, Do, Study, Act (PDSA) and Clinical Supervision Model.