Joint White Paper from PASA and PAESSP
June 2015

Educator Effectiveness and Evaluation

Over the last five years, states across the nation have debated ways to improve educator evaluation.

Pennsylvania is no exception; in 2009 the Commonwealth developed a proposal to revamp educator evaluation in order to qualify for a Race to the Top (RTTT) grant. Although Pennsylvania was not awarded a grant in the initial RTTT competition, efforts to restructure teacher evaluation continued with the help of the Bill and Melinda Gates Foundation. More recently, Pennsylvania pursued and was awarded a Phase 3 RTTT grant of $41 million and an ESEA Accountability Waiver, both of which required substantial changes to the state’s educator evaluation system.

These converging pressures from national advocacy groups and the Federal government led to the passage of Act 82, Pennsylvania’s new educator evaluation law.

Feedback from superintendents, central office staff, principals, and teachers about Act 82 implementation inform this programmatic review.

Introduction: A Vision of High-Quality Evaluation

Excellent educators appreciate thoughtful, useful feedback on their work. But effective feedback can only be provided by an effective system of supervision and evaluation. Effective supervision and evaluation systems must be as nuanced as the jobs they are designed to examine. Current research has defined several components of effective evaluation systems.

1. **Challenging professional standards to define appropriate practice.** Professional standards tell employers what to examine and employees what to demonstrate. Employees can reflect on good practice and revise their work with a clear goal in mind.

2. **Flexibility to relate professional standards to local goals.** Within the context of professional standards, effective evaluation systems allow employers to prioritize specific behaviors, knowledge and skills.

3. **Different expectations for professionals, based upon career stage, employee and employer needs, and the purpose of the evaluation.** Effective evaluation systems help professionals grow. This means that effective evaluation systems should hold somewhat different expectations for educators over time.

4. **An understanding of organizational supports and barriers to effective job performance.** All professionals require supports; a doctor needs access to medicines and a lawyer needs access to evidence. Educators, too, need resources to be most effective. Effective evaluation systems have the capacity to link educator performance to professional supports and barriers.1

---

5. **Employee engagement, self-appraisal and feedback.** When employees participate in their evaluation the quality and quantity of information increases, and employees demonstrate higher levels of cooperation with the evaluation process.2

6. **A strong and diverse body of evidence, including opportunities for educators to produce evidence.** Education is a multifaceted profession, and the greatest amount of work is often the intellectual planning that leads to a particular behavior or instructional event. Consequently, the educator evaluation needs to capture the complexity of educators’ work.

7. **Safeguards to ensure that high-stakes decisions are based upon valid and reliable measures of an employee’s work.** A valid and reliable evaluation system needs to hold employees accountable for the work that they do and the workplace circumstances they control, not the work of others or characteristics of the workplace that are beyond the employee’s direct influence.

8. **Link to Professional Development.** The goal of any well-structured evaluation system is to improve professional practice. An evaluation system is only useful to the extent that it can produce actionable, evidence-based suggestions for professional learning.3

9. **Ongoing training for evaluators and those being evaluated.** All educators must know the standards underlying the evaluation system, the evaluation process, and the components of and criteria for evaluation. To ensure that everyone has the requisite knowledge and skills, teachers and administrators need to be thoroughly trained.4

---

**Pennsylvania’s New Educator Evaluation Models: Promising Characteristics**

The true goal [of evaluations] is to improve student achievement so that all of the children in Pennsylvania’s public schools are prepared to enter a career or post-secondary training and become productive citizens.5

Pennsylvania developed a new State approved Teacher and Principal Effectiveness System under Act 82. This system is comprised of discrete evaluation models for classroom teachers, non-teaching professionals (including educational specialists) and principals.

After two years of implementing these new evaluation models, educators have learned several lessons about effective evaluation and have grown to appreciate positive aspects of the new system. Educators also are raising questions and issues related to validity, reliability, effectiveness, and practicality of the new system.

---


5 Quote from PDE Website: [http://www.portal.state.pa.us/portal/server.pt/community/educator_effectiveness_project/20903](http://www.portal.state.pa.us/portal/server.pt/community/educator_effectiveness_project/20903)
**Promising Characteristics of the Teacher Effectiveness Model**

Pennsylvania’s Teacher Effectiveness Model contains several characteristics that are likely to improve educator effectiveness and student learning.

- The Danielson Framework provides for a thorough review of teacher effectiveness and practices in the four domains, inclusive of planning, in-class instruction, and reflection (Vision 1).
- The practice of using student data to inform and improve teacher instruction can be powerful; however how these data are used in the model warrants review.
- The use of multiple measures of student achievement encourages educators to focus on many indicators to understand the impact of their practice (Vision 6).
- Many employers are using the new evaluation system to encourage professional reflection (Vision 5).
- The State has encouraged employers to link evaluation results to individualized professional development planning, both through differentiated supervision and through state-provided online professional development related to the Danielson Framework (Vision 8).
- Education specialists such as school nurses, school guidance counselors, and home and school visitors now have specific rubrics defining effective practice within their fields (Vision 1).
- The law governing the new process includes language in it that clarifies the appropriate use of student test score data by stating, “No teacher can receive a needs improvement or failing based solely on student test scores.” This requirement sets an important expectation that test scores not be used to trump the professional judgment of educators about proficient practice (Vision 7).
- The State has provided substantial training on the new teacher evaluation system through Intermediate Units, although the extent to which this training has been shared with teachers varies (Vision 9).
- The system allows employers to determine the weight of individual components within the four domains of clinical observation (Vision 2).

**Promising Characteristics of the Principal Effectiveness Model**

The Principal Effectiveness Model also incorporates several components of effective evaluation that are likely to improve principals’ effectiveness as instructional leaders.

- For the first time, the Commonwealth has a standardized principal evaluation process (Vision 1).
- The standards for the principals’ evaluation provide for a consistent model in which to observe practice and direct professional growth opportunities (Vision 1).
- The use of multiple measures of student achievement encourages principals to adopt a holistic vision of effective leadership (Vision 6).
- Many employers are using the new evaluation system to encourage professional reflection (Vision 5).
- The State has provided substantial training on the new educator evaluation system through Intermediate Units, although the extent to which this training has been shared with building level administrators varies substantially (Vision 8, 9).
The system allows employers to determine the weight of individuals components within the four domains of clinical observation (Vision 2)

Pennsylvania’s New Educator Evaluation Models: Unintended Consequences

Although there are many positive outcomes to both the teacher and principal evaluation systems, there also have been many unintended negative consequences.

1. **Requires Extraordinary Time Commitment.** The complexity of the new system means that it requires extraordinary time and effort to administer. This complexity actually diverts attention away from professional practice and student learning and focuses attention on compliance requirements.

2. **Establishes a Disruptive Timeline for Improvement/Dismissal.** School and student data used to calculate summative ratings are released by the State each fall, pushing the release of annual summative ratings well into the following school year. This timeline is disruptive to annual professional development and professional improvement plans and/or dismissal decisions.

3. **Holds Educators Accountable for Factors Unrelated to Educator Effectiveness.** School Performance Profile (SPP) data comprise 15 to 20 percent of an educator’s summative rating. However, differences in SPP scores can be due to many factors unrelated to educator effectiveness. In fact, SPP scores correlate strongly with the percent of economically disadvantaged students in a school. In addition, there is no clear consensus among researchers about the ability of test scores to measure educator effectiveness.

4. **Misuses a Professional Learning Strategy for High-Stakes Purposes.** Similar to the SPP data, the SLO process has not been demonstrated to be a valid or reliable measure of educator effectiveness. SLOs can engage teachers in inquiry about the relationship between instructional practices and student outcomes, but using the SLO to evaluate principals or teachers assumes a causal relationship where none may exist. Depending on the work of the district, written curriculum and established learning outcomes or assessments may not be available, therefore holding teachers and principals accountable to outcomes beyond their realm of control.

5. **Fails to Emphasize Professional Growth among Satisfactory Educators.** The new system does not set different expectations for individuals at different career stages. Nor does the system capture information about specific supports for and barriers to effective practice; this information is critical to understanding the conditions under which an educator works and to building supports to improve future practice.

6. **Reduces Focus on Targeted Professional Learning Specifically Related to Practice.** In many cases teachers are finding that they are spending substantial time learning how to comply with evaluation requirements at the expense of time to collaborate with colleagues around issues of professional practice or engage in content-based professional learning.

7. **Provides a Disincentive to Work with Struggling Students or to Collaborate.** Because student and teacher test scores comprise a significant portion of the summative rating, the new system encourages educators to “shop” for jobs in
schools with higher scores. In a related vein, some educators may avoid tested
grades and subjects because of concerns about the validity and reliability of test-
based educator evaluations. Finally, the new evaluation system may inadvertently
decrease professional collaboration because the high-stakes use of student
achievement data places teachers in competition for students and classes that are
more likely to perform well on tests.

8. **Calculates a Penalty for Attaining Expected Growth.** The School Performance
Profile uses PVAAS scores to calculate a significant portion of a school’s overall
score. However, if an elementary or middle school attains one year’s worth of
growth according to PVAAS, the school only attains 7 of the possible 10 points
possible on the SPP. If one year’s growth is attained in each content area, the school
can lose as much as 9 points of 30 available. This means that schools that attain
expected one year’s growth are penalized unnecessarily in their overall score.

9. **Makes Dismissal More Difficult and Less Fair.** The new system makes it difficult
and time-consuming for principals to issue unsatisfactory ratings to teachers,
particularly in schools where students perform well on state tests. Conversely,
teachers in struggling schools may find their scores suppressed because of where
they teach rather than because of deficiencies in practice.

10. **Fails to Discriminate between Supervision and Evaluation.** Effective supervision
and evaluation each play an important role in fostering effective professional
practice, but their roles are different. Traditionally, supervision is designed to foster
professional growth, whereas evaluation is designed to measure educators against a
standard. Supervision is ongoing and encourages risk-taking, constructive self-
analysis, and collegial learning, whereas evaluation is standardized, hierarchical,
and encourages educators to demonstrate their best possible work. The new
evaluation system muddies the line between supervision and evaluation by defining
“differentiated supervision” as part of a high-stakes evaluation process and
neglecting the role of supervision in professional learning.

11. **Overlooks the Unique Job Responsibilities of Assistant Principals.** Assistant
Principals’ roles in districts vary according to their job description. Their work is
unique to their position and should have their own rubric specific to their job
responsibilities.

12. **Provides a Disincentive for Accepting Student Teacher Placements.** Because of
the high-stakes consequences attached to student test performance and the inability
to determine what the impact of a different teacher may be on student learning,
classroom teachers are more reluctant to accept student teacher placements in their
classrooms.

**Critical Lessons from Implementation of the Educator Effectiveness Models**

While the new Educator Effectiveness System is an important first step in
developing a robust and fair evaluation process, educational professionals across the
Commonwealth have discovered important lessons through the first years of
implementation:

1. **Importance of Simplicity.** A system that is too complex will not be administered
with fidelity. The system needs to be simplified for teachers and administrators.
2. **Importance of Brevity.** A system that is too time-consuming will not be administered with fidelity. The system needs to be implementable for both teachers and administrators within the time available.

3. **Need for Across-the-Board Training.** When those on the receiving end of the new system do not receive ample training they are unlikely to have any faith in or understanding of the results. The system needs to be supported with comprehensive, ongoing training for all participants.

4. **Focus on Professionalism and Professional Expertise.** Professional decision-making is a critical aspect of educator effectiveness. The system needs to be designed to focus all educators on high-quality decision-making rather than on student test results and to allow administrators to exercise professional expertise in determining proficiency.

5. **Emphasis on Improvement.** Staff development needs to be the focus of any system to improve professional skills. The system needs to be designed so that all educators have access to excellent staff development and a few educators are subject to intensive clinical evaluation when warranted.

6. **Importance of Timing.** Issuing summative ratings in the fall based upon data from the prior year makes it hard to address professional learning needs or take personnel action in a timely manner. The system should provide immediate, actionable feedback to help educators improve their practice and to ensure appropriate personnel decisions.

7. **Perception of Fairness.** Using measures of student achievement and growth impacts different groups of educators differently and in ways that may have nothing to do with educator effectiveness. This impacts the “face validity” of the system among participants and reduces the utility of results because they are not perceived as accurate and meaningful. The system needs to be designed to ensure that no educator is penalized for working in a less effective school, in a less effective district, or with struggling students.

8. **Need to Support Collaboration among Educators.** Evaluating educators on student test scores means they may compete for “better students,” “better classes,” or “better schools.” Resentments can arise because of perceived differences in the students teachers are assigned to teach. The system should encourage professionals to collaborate, not compete.

9. **Appropriate Use of Student Test Score Data.** A test score cannot determine which teaching practices or other circumstances led to specific levels of student achievement or which practices an individual teacher should change to achieve different outcomes. The system should encourage the use of student test score data as a signaling device that may lead to further examination of practice, not as a high-stakes measure of educator effectiveness.

**Recommended Strategies to Address Unintended Consequences and Build an Effective Educator Evaluation System**

Effective evaluation is a critical component of professional growth. However, in order for evaluation results to be useful they need to be perceived as valid and reliable, be efficient and simple to implement, support professional collaboration, allow for timely
administrative decisions about necessary professional support or personnel actions, and be supported by high-quality training.

Current national initiatives also require that measures of student achievement be incorporated into any educator evaluation system. PAESSP, PASA, and PSEA agree on the building blocks that support an effective educator evaluation system that meets these criteria.

**Effective Components of the Educator Effectiveness Project**

The new educator effectiveness system established as a result of Act 82 includes several promising components that foster a high-quality vision of evaluation. These components should be retained

- Continue to evaluate instructionally certified educators with a rubric based upon the Danielson Framework for Teaching (Vision 1).
- Continue to evaluate school principals and assistant principals with a rubric based upon Pennsylvania’s Framework for Leadership (Vision 1).
- Continue to evaluate education specialists and licensed professionals with Danielson-based rubrics that reflect the unique nature of their work (Vision 1).
- Continue to gather annual student achievement data at both the state and local level and issue teacher specific data as applicable and appropriate (Vision 6).
- Continue to encourage differentiated data collection for the evaluation of experienced, satisfactory educators (Vision 3).
- Continue to allow an educator to be dismissed if the educator receives two consecutive unsatisfactory scores issued at least four months apart.
- Continue to base dismissal decisions upon “deficiencies in practice” as required by Act 82 and defined by Pennsylvania’s professional standards rather than upon student test scores (Vision 7).
- Continue to develop training for educators on the applicable standards of practice and on components of effective evaluation and supervision (Vision 9; Unintended Consequence 10; Lesson 3).
- Continue to implement the Student Learning Objective process for the collection and analysis of local student achievement data (Vision 6).
- Continue to allow employers to determine how to weight each individual component with a Domain (Vision 2).

**Modifications to Improve the Educator Effectiveness Project**

Based upon lessons learned from the initial implementation of the educator effectiveness system, several components of the new system should be modified to improve the validity, reliability, utility, simplicity, and efficiency of the new system.

- **Use Student Achievement Data to Inform Individualized Professional Learning, Focus Supervision Activities, and Direct Supervisor Decisions about Clinical Observation.** Use measures of student achievement as signaling devices for supervisors to determine which educators may need additional observation and support. Allow supervisors, in consultation with educators, to use their judgment in determining what additional supports an educator may need and if an educator’s student achievement data raise questions about professional performance that
warrant greater supervision and clinical observation. Focus all educators’ professional development on individualized data analysis and targeted instructional planning (Vision 2, 3, 4, 5, 6, and 8; Unintended Consequences 3, 4, 6, 7, 8, 9, 12; Lesson 4, 7, 9).

- **Allow an educator to be dismissed if the educator receives two consecutive unsatisfactory scores, issued at least four months apart, on the “Observation and Practice” portion of the evaluation** (Vision 7; Unintended Consequences 9; Lesson 1, 4, 7, 9)

- **Focus on Professional Growth.** Require that an improvement plan be developed for any educator who receives a needs improvement or failing on the “Observation and Practice” portion of the rating form. Ensure that all improvement plans are targeted to address clearly defined issues of professional practice, not simply test score growth. For “satisfactory” teachers, use student achievement data to direct critical reflection, collaboration, and ongoing professional learning (Vision 3; Unintended Consequences 5; Lesson 1, 5, 8).

- **Maintain a Statewide Standard Minimum Score.** Continue to define the four rating categories—distinguished, proficient, needs improvement, or failing—according to the point scale defined in current regulations:
  
  - 0.00 – 0.49  Failing
  - 0.50 – 1.49  Needs Improvement
  - 1.50 – 2.49  Proficient
  - 2.50 – 3.00  Distinguished (Vision 1)

  Apply this rating scale to the “Observation and Practice” portion of the evaluation system to determine the summative rating of an educator for employment decisions.

- **Issue Annual Summative Ratings by the End of the Academic Year.** Because student assessment data issued by the state is used to direct planning and supervision activities, all summative ratings can be completed in May with the most recent student achievement data available (Unintended Consequences 2; Lesson 1, 2, 6).

- **Evaluate All Instructionally Certified Staff with the Same High Standards.** Rather than dividing instructionally certified staff into “classroom teachers” and “nonteaching professionals,” adopt the same clinical observation process for all instructionally certified staff, and use student achievement data to inform professional development, supervision, and additional clinical observation needs of individual educators with instructional certificates (Vision 1).

- **Discriminate between Supervision and Evaluation to Encourage Educator Growth.** Develop a universal evaluation system that is administered annually for all educators. In addition, the state should develop guidelines for effective supervision and provide professional development for all educators on the different roles. (Unintended Consequences 10).

- **Streamline the process of gathering and reporting student data.** Develop one section of reporting local student data for Teacher Specific Data and Student Learning Outcomes reducing the complexity and redundancy between these two components. (Unintended Consequences 1, 2, 10)
• **Recalculate the scoring for Indicators of Calculated Growth on SPP.** Develop a higher calculation for attaining goal of one year’s growth. Example: one year’s growth - 9 of 10 points; 1 standard deviation above -9.5 points, and 2 standard deviations above 10 points.) (Unintended Consequences 8)

• **Evaluate Assistant Principals under the Non-Teaching Professionals/Specialists Evaluation tool.** Develop a specific rubric for assistant principals’ role and evaluate using the Specialists Evaluation tool. (Unintended Consequences 11).
CLASSROOM TEACHER RATING FORM

Last Name | First | Middle
---|---|---
District/LEA | School | 
Rating Date: | Evaluation: (Check one) | □ Semi-annual | □ Annual

(A) Teacher Observation and Practice

<table>
<thead>
<tr>
<th>Domain</th>
<th>Title</th>
<th>Rating* (A)</th>
<th>Factor (B)</th>
<th>Earned Points (A x B)</th>
<th>Max Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>Planning &amp; Preparation</td>
<td>20%</td>
<td></td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>II.</td>
<td>Classroom Environment</td>
<td>30%</td>
<td></td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>III.</td>
<td>Instruction</td>
<td>30%</td>
<td></td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>IV.</td>
<td>Professional Responsibilities</td>
<td>20%</td>
<td></td>
<td>0.60</td>
<td></td>
</tr>
</tbody>
</table>

(1) Teacher Observation Practice Rating: 3.00

*Domain Rating Assignment* 0 to 3 Point Scale (A)

<table>
<thead>
<tr>
<th>Rating</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Failing</td>
<td>0</td>
</tr>
<tr>
<td>Needs Improvement</td>
<td>1</td>
</tr>
<tr>
<td>Proficient</td>
<td>2</td>
</tr>
<tr>
<td>Distinguished</td>
<td>3</td>
</tr>
</tbody>
</table>

(B) Student Performance—Building Level Data, Teacher Specific Data, and Elective Data

Building Level Score (0—107)

(2) Building Level Score Converted to 3 Point Rating

(3) Teacher Specific Rating

(4) Elective Rating

(C) Final Teacher Effectiveness Rating—All Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Rating</th>
<th>Factor (D)</th>
<th>Earned Points (C x D)</th>
<th>Max Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Teacher Observation Practice Rating</td>
<td>50%</td>
<td></td>
<td>1.50</td>
<td></td>
</tr>
<tr>
<td>(2) Building Level Rating</td>
<td>15%</td>
<td></td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>(3) Teacher Specific Rating</td>
<td>15%</td>
<td></td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>(4) Elective Rating</td>
<td>20%</td>
<td></td>
<td>0.60</td>
<td></td>
</tr>
</tbody>
</table>

Total Earned Points: 3.00

Conversion to Performance Rating

<table>
<thead>
<tr>
<th>Total Earned Points</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00-0.49</td>
<td>Failing</td>
</tr>
<tr>
<td>0.50-1.49</td>
<td>Needs Improvement</td>
</tr>
<tr>
<td>1.50-2.49</td>
<td>Proficient</td>
</tr>
<tr>
<td>2.50-3.00</td>
<td>Distinguished</td>
</tr>
</tbody>
</table>

□ Rating: Professional Employee, OR □ Rating: Temporary Professional Employee

I certify that the above-named employee for the period beginning _______ and ending _______ has received a performance rating of:

□ DISTINGUISHED □ PROFICIENT □ NEEDS IMPROVEMENT □ FAILING

resulting in a FINAL rating of:

□ SATISFACTORY □ UNSATISFACTORY

A performance rating of Distinguished, Proficient or Needs Improvement shall be considered satisfactory, except that the second Needs Improvement rating issued by the same employer within 10 years of the first final rating of Needs Improvement where the employee is in the same certification shall be considered unsatisfactory. A rating of Failing shall be considered unsatisfactory.

Date Designated Rater/Position: Date Chief School Administrator

I acknowledge that I have read the report and that I have been given an opportunity to discuss it with the rater. My signature does not necessarily mean that I agree with the performance evaluation.

Date Signature of Employee