

# Update

## Regents Task Force on Teacher and Principal Effectiveness

**Co-Chairs of the Task Force**  
Vice Chancellor Milton Cofield  
Regent Lester Young



Joseph P. Frey, Deputy Commissioner  
Amy McIntosh, Senior Fellow  
January 10, 2011

# Chapter 103 of the Laws of 2010

## What Does the New Law Require?

- New performance evaluation system for teachers and principals
  - 20% - State student growth data or comparable measure of student growth (increased to 25% upon implementation of a value-added growth model)
  - 20% - Locally selected measures of student achievement that are determined to be rigorous and comparable across classrooms in accordance with regulations of Commissioner (decreased to 15% upon implementation of value-added model)
  - 60% - Multiple measures of teacher/principal effectiveness based on standards prescribed in the Regulations of the Commissioner
- Four rating categories: highly effective; effective; developing; ineffective
- Resulting in a single composite score of teacher or principal effectiveness

# Chapter 103 of the Laws of 2010

## What Does the New Law Require, continued

- Appropriate training for all evaluators
- Teacher/Principal Improvement Plans (developing/ineffective)
- Utilize evaluation results as a factor in career ladder decisions and other recognition
- Locally-developed appeal process
- Expedited 3020a process (single hearing officer/after two consecutive ineffective ratings)

## Timelines in Chapter 103 of the Laws of 2010

- **July 2010** – New collective bargaining agreements must be consistent with the requirements of Chapter 103 of the Laws of 2010
- **July 2011** – New performance evaluation system takes effect for classroom teachers of common branch subjects, ELA or math in grades 4-8 along with their respective building principals
- **July 2012** – New Performance evaluation system goes into effect for remaining teachers and building principals
- **2012-2013 school year and thereafter** – Implementation of teacher and principal improvement plans and implementation of a value-added growth model to be used within the teacher and principal performance evaluation system

# Regents Task Force on Teacher and Principal Effectiveness

- Education Law 3012c requires Commissioner to establish an advisory committee consisting of representatives of:
  - Teachers
  - Principals
  - Superintendents of Schools
  - School boards
  - School district and BOCES officials
  - Other interested parties
  
- The advisory committee makes recommendations to the Commissioner

# Regents Task Force: Committee

## Members of the workgroups:

- Teachers: \* 13\*\*
- Principals: \* 11\*\*
- Superintendents: 4\*\*
- Union: 5
- Administrators: 6
- Higher Ed: 5
- School Boards: 2
- BOCES: 4

\*Selected by many different constituency groups

\*\*Includes school districts and BOCES personnel

## Organizations:

- NYSSBA
- NYSCOSS
- NYSUT
- District Superintendents
- SAANYS
- CSA
- Big 5 School District
- SCDN Network
- PSPB
- Representatives of the Arts
- ITI

# Regents Task Force: Committee

- Currently there are four work groups:
    - Non-Tested Subjects
    - Locally Selected Assessments
    - 60% Non-Growth Measures for Teachers
    - 60% Non-Growth Measures for Principals
  - Each work group has SED senior staff and Regents Research Fellow support
  - Website created for research and best practice examples\* provided by Fellows, SED staff, and TF members
    - Enables sharing of discussion documents, comments
- \* See Appendix for highlights

# Regents Task Force: Meetings

- Work Groups have met monthly in-person since Sept. on the second day of the Regents meetings with conference calls between sessions
  - Morning and afternoon WG sessions:
  - SED Resource person, Fellow, Senior Management
  - Mid-day full Task Force session to discuss broad-based policy issues
  
- Examples of presenters on interim calls:
  - Laura Goe, ETS (supporting AFT innovation districts) to Teacher WG
  - TNTP to teacher WG (supporting NYC pilot)
  - Rob Meyer, Uwisc. to Non-Tested WG
  - Matt Keleman, New Leaders for New Schools to Principal WG



# Full Task Force: Presentations

- Introduction to student growth – how it can be measured using student growth percentiles
- Introduction to value-added models – how do you control for variables outside of teacher/principal control
- With the assistance of three researchers and the Center for Assessment:
  - Hamilton Lankford, SUNY Albany
  - Jonah Rockoff, Columbia University
  - James Wyckoff, UVA
  - Center for Assessment

# Regents Task Force: Work Groups

- Each work group will summarize their initial conclusions to the Full Task Force this month
- The Commissioner will receive an initial report in February Task force meeting and another in March
- Work groups will continue to meet on additional topics in February and March
- Task Force recommendations to Regents in April

# Regents Task Force: Progress

- The Work Group on 60% Non-Growth Measures for Teachers has:
  - Agreed the NYS Teaching Standards should be the foundation of the 60% component of teacher evaluation
  - Reviewed specific criteria to define proposed teacher practices rubrics
  - Discussed how the approval of 3<sup>rd</sup> party and district-developed rubrics should occur
  - Considered a range of tools to measure teacher effectiveness

# Regents Task Force: Work Groups

- The Work Group on Non-Tested Subjects has:
  - Identified key teacher groups to prioritize given their size and similar assessment issues
    - E.g. K-2 educators, High School, Performance courses (Arts, CTE, etc.)
  - Reviewed assessment options evaluated on four criteria:
    - E.g. Comparability; rigor/validity; cost and feasibility of implementation; and effect on instruction
  - Began the process of identifying assessment options

# Regents Task Force: Work Groups

- The Work Group on Locally Selected Assessments has:
  - Focused first on grades 4-8 ELA/Math since this requires first implementation
  - Identified primary options for local assessments
    - E.g. existing local assessments, vendor-built custom tests, etc.
  - Discussed criteria for selecting local assessments

# Regents Task Force: Work Groups

- The Work Group on Principal 60% other measures has:
  - Discussed the State's previous work with Wallace Foundation on a coherent leadership system that includes principal evaluation
  - Considered options for leadership standards
  - Discussed appropriate degree of local flexibility in choosing approaches to principal evaluation
  - Discussed how districts might establish multiple measures of principal effectiveness within the "60%" other metrics

# Regents Timetable

	Board of Regents Agenda
January	Policy Options Discussion: “60%” teacher and principal measures
February	Policy Options Discussion: Local assessments and Non-tested subjects
March	Policy Options Discussion: <ul style="list-style-type: none"><li>• VA/growth modeling for “20%” state assessment</li><li>• Determining scores and ratings (Highly Effective, Effective, Developing, Ineffective)</li></ul>
April	Regents Task Force Recommendations
May	Draft regulations for 2011-12 implementation
June	Emergency adoption of regulations

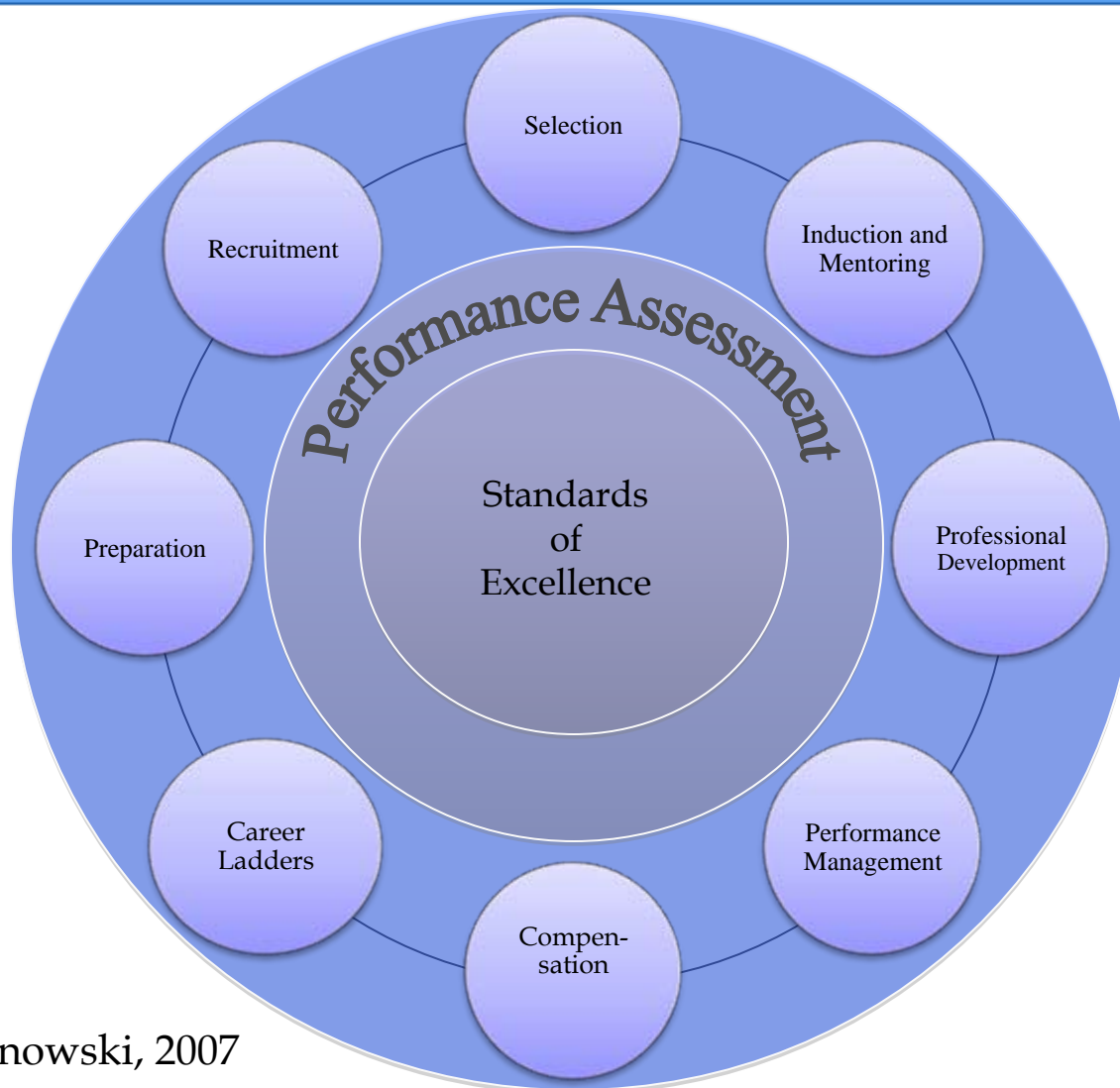
# Overview of Policy Options

## Introductory discussion

- Teacher and Principal evaluation
- Options for “60% Other” measures of teacher or principal effectiveness



# Ensuring Teacher and Principal Excellence



# Teacher 60%: Policy Options – Design Considerations

## **Design considerations from recent research:**

1. Annual evaluations for all
2. Clear, rigorous expectations for instructional excellence, prioritizing student learning
3. Multiple measures of performance
4. Multiple ratings: at least 4 performance levels to describe differences in teacher effectiveness
5. System should encourage regular constructive feedback and ongoing development
6. Significance: results are a major factor in employment decisions

Source: The New Teacher Project (2010). Teacher Evaluation 2.0. Available: [www.TNTP.org](http://www.TNTP.org)

# Teacher “60%”: Setting Clear, High Expectations

- Starting point: New NYS Teaching Standards, on Regents Agenda for adoption in January
- Teacher practice rubrics required to expand on NYS Teaching Standards
  - Describe differences in performance levels (Highly Effective, Effective, Developing, Ineffective)
  - Articulate specific, observable behaviors of students and teachers as evidence
- Options:
  - One state-wide rubric
  - State provides menu of rubric options that meet state-determined criteria with district variance procedure for other choices
  - Districts choose or develop own rubrics

# Teacher “60%”: Teacher Practice Rubric Examples

- **General rubrics**
  - Broadly cover classroom and non-classroom practices
  - E.g. Charlotte Danielson Framework for Teaching
    - Widely used or adapted in NY State, nationally
    - NYSUT innovation districts adapted this to directly align with NYS standards and added more granular indicators
- **Classroom Observation Only Rubrics**
  - Focus deeply on evidence available from classroom observation
  - E.g. CLASS from University of Virginia
- **Subject-specific**
  - Tailored to math, ELA, or science instruction in secondary school
  - E.g. Mathematical Quality of Instruction (Hill and Ball); PLATO for ELA (Grossman)
- **Designed by district, charter or non-profit organization**
  - D.C. and Denver
  - Teach for America: Teaching as Leadership framework
- Note: Gates Measurement for Effective Teaching Research will assess many of the examples above for their alignment to student learning outcomes and other measures of teaching effectiveness
- More on 75 teacher evaluation tools at <http://learningpt.org/tqsource/GEP>

Teacher “60%”:

## Multiple Measures – Classroom Observations

- **Principal and other administrator observation** of classroom practice (using a rubric) is usually a major part of teacher evaluation
- **Independent observers**, usually expert teachers, are becoming more common
  - Examples: Denver, Hillsborough, New Haven, D.C., Cincinnati, Toledo
- **Training and ongoing monitoring** of all evaluators' accuracy and reliability is very important

Teacher 60%:

## Multiple Measures – Beyond Classroom Observation

NYS Standards cover practices that are not observable in the classroom visits. Examples of tools\*:

- **Structured review of student work and/or teacher documents**
  - E.g. NYSUT districts piloting “evidence binders” with a variety of performance tasks that are individually assigned based on need, and scored by rubric
- **Mid-year and year-end conferences** address other domains of standards
  - E.g., D.C. and NYC pilot
- **Student surveys** like the Tripod surveys from Ron Ferguson at Harvard and Cambridge Education ([www.tripodproject.org](http://www.tripodproject.org)), which generated initial strong correlation to student outcomes in initial Gates METS research results ([www.metproject.org](http://www.metproject.org))
- **Teacher attendance**

\* Goe, Bell, and Little (2008) provide a synthesis of other options ([www.tqssource.org](http://www.tqssource.org))

# Teacher “60%”: Other States

- Most RTTT states have adopted **state teaching standards**
- **Rubrics:**
  - 6 states (DE, GA, NC, OH, RI, TN) have adopted a single state wide rubric (2 based on Danielson, 1 based on CLASS)
  - 2 states will choose a default option and provide criteria for local selection of rubrics (MA, MD)
  - 1 state will allow for local choice including Danielson or other best-practices (HI)
- **Observations:**
  - Principal, or other supervisor, conducts teacher assessments in all 11 states where we gathered information on evaluation models
  - Independent and/or peer observers in addition to supervisors is described by 7 states as either an option or required (CO, FL, GA, HI, LA, MD, TN)

# Teacher 60%: Specific District Examples

	<b>Hillsborough County</b>	<b>D.C.</b>	<b>NYSUT Innovation Districts Pilot 2011-11</b>	<b>NYC Pilot 2010-11</b>	<b>Denver</b>
<b>Rubric</b>	Danielson	Developed own framework	Danielson adapted to NYS standards	Danielson	Developed own district-wide - rubric
<b>Multiple Measures</b>	<ul style="list-style-type: none"> <li>• Administrator Observations</li> <li>• Roving full-time expert Observers</li> </ul>	<ul style="list-style-type: none"> <li>• Administrator Observations</li> <li>• Independent Observers</li> <li>• Roving Full-time Expert Observers</li> <li>• Commitment to School Community (5%)</li> </ul>	<ul style="list-style-type: none"> <li>• Administrator Observations</li> <li>• Teacher Portfolio/Evidence Binders</li> </ul>	<ul style="list-style-type: none"> <li>• Administrator Observations</li> <li>• School-selected Measures (10%)</li> </ul>	<ul style="list-style-type: none"> <li>• Administrator Observations (2 unannounced)</li> <li>• Independent Observers (2 unannounced focused on learning environment and instruction)</li> <li>• Roving Full-time Expert Observers</li> </ul>
<b>Total Weight From These Measures</b>	60%	55%	60%	60%	50%
<b>Notes</b>	<ul style="list-style-type: none"> <li>• 5 Rating Categories</li> <li>• # of observations vary by teacher experience and prior ratings</li> </ul>	<ul style="list-style-type: none"> <li>• 4 Levels</li> </ul>	<ul style="list-style-type: none"> <li>• Peer Assistance</li> </ul>	<ul style="list-style-type: none"> <li>• 4 Rating Categories</li> <li>• District-Wide Peer Assistance</li> </ul>	<ul style="list-style-type: none"> <li>• 4 Major Rating Categories (7 counting Sub-Ratings)</li> </ul>



# Principal “60%”: Design Considerations

Key principles from SED’s recent “Cohesive Leadership System” project through Wallace Foundation

- Use clearly articulated leadership standards, specifically the Interstate School Leaders Licensure Consortium (ISLLC) Standards for School Leaders
- Collect feedback and evidence from multiple sources including supervisor; parents, students, teachers; self-assessment and other sources
- Be based on research, best practice, and experiential learning
- Promote learning for all students
- Be sensitive to the diversity and the context of the school and district
- Should promote principal-supervisor collaboration and trust and be a shared responsibility
- Lead to professional growth and development of the leader that is being evaluated

# Principal “60%”: Standards, Rubrics and Assessment Tools

- ISLLC or local adaptations are the most common **leadership standards** used
  - The ISLLC Standards have helped guide leadership policy and practice in more than 40 states since they were released in 1996
  - Recently updated version used to inform leadership standards in DE, FL, RI, OH, MD, NC
- **Rubrics** that differentiate performance levels include:
  - 360 degree survey tools like Val-Ed from Vanderbilt University ([www.valed.com](http://www.valed.com))
  - New Leaders for New Schools recently-released rubric ([www.nlms.org](http://www.nlms.org))
  - School-wide practices rubric for inspection visits like NYC Quality Review modeled on Ofsted (UK) (<http://schools.nyc.gov/Accountability/tools/review>)
  - State and/or local developed (FL, RI, MA, OH, HI, MD, NC)
- Parent, teacher and student **surveys** (independent of a 360 degree feedback instrument)
  - E.g. New Teacher Center Working Conditions survey of teachers, in Gates METS study; NYC teacher, student, parents surveys
  - Used by DE, GA, HI, MD, OH

# Principal “60%”: Individual Goal-Setting

- Annual goal setting between superintendent and principal and progress monitoring throughout the year\*
  - Debate in the literature around whether or not goals should be part of a principal’s evaluation
  - Also should goals be related to school and district education plans only or should they also include individual professional growth goals?
- Examples where goals are part of evaluation: DE, TN, OH, D.C., Hillsborough County, New Haven, NYC, and Rochester

\* E.g. New Leaders for New Schools, National Comprehensive Center for Teacher Quality

# Principal “60%”: Other Options

- Measures of principal impact on **teacher effectiveness**:
  - Inputs like timely and thorough teacher evaluations, development plans
  - Outcomes like teacher attendance, effectiveness of teachers offered tenure; alignment of ratings with student learning results; retention of more effective teachers
  - Examples: FL, GA, LA, MA, HI, MD, D.C., Hillsborough County
- **Operations**:
  - Budget, compliance, safety metrics
  - Examples: HI, NC, RI, D.C., Hillsborough County, NYC, Rochester

# Principal “60%”: Differing Examples from States

	<b>Delaware</b>	<b>Georgia</b>	<b>Ohio</b>	<b>Massachusetts</b>
<b>Standards</b>	ISLLC	Georgia’s Leadership Performance Standards: Leader Keys	Ohio Standards for Principals	Will develop framework and rubrics
<b>Multiple Measures</b>	<ul style="list-style-type: none"> <li>• 360 Survey of principal, teachers, evaluator</li> <li>• Individual goals</li> <li>• School or district improvement plan goals</li> </ul>	<ul style="list-style-type: none"> <li>• Rubric-based review of practice</li> <li>• Retention of effective teachers</li> <li>• Student, staff, parental feedback</li> <li>• Student attendance</li> <li>• Developing additional tools to assess student engagement</li> </ul>	<ul style="list-style-type: none"> <li>• 360 degree survey</li> <li>• Choice of McRel, Val-Ed or “other educational impact tools”</li> <li>• Goals (2-3) set with supervisor</li> <li>• Student attendance</li> <li>• Graduation rates</li> <li>• Suspensions and expulsions</li> <li>• Percent of students in AP classes</li> </ul>	<ul style="list-style-type: none"> <li>• Effectiveness measures of leadership skills (exemplars and models will be provided by MA)</li> <li>• Peer evaluation</li> <li>• Teacher effectiveness</li> <li>• Self-assessments of professional skills</li> </ul>
<b>Total Weight From These Measures</b>	75%	30%	60%	60%
<b>Notes</b>	<ul style="list-style-type: none"> <li>• 4 Rating Categories</li> </ul>		<ul style="list-style-type: none"> <li>• 5 Rating Categories</li> <li>• Districts determine weights of components and how to arrive at final rating</li> </ul>	<ul style="list-style-type: none"> <li>• At Least 3 Rating Categories</li> </ul>

# Principal “60%”: Differing Examples from Districts

	Hillsborough County	D.C.	Rochester (today)	NYC (today)
<b>Rubric</b>	Val-Ed (360 degree assessment)	Developed own framework	Developed own leadership standards rubric	Quality Review
<b>Multiple Measures</b>	<ul style="list-style-type: none"> <li>• Assessment of leadership, school climate, academic expectations (30%)</li> <li>• Management of school operations (10%)</li> <li>• Retention of teachers and ability to effectively evaluate/differentiate teachers (10%)</li> <li>• Student attendance and discipline (10%)</li> </ul>	<ul style="list-style-type: none"> <li>• Instructional leadership (10%)</li> <li>• Organizational leadership (5%)</li> <li>• Creating a safe and effective learning environment (5%)</li> <li>• Family and community engagement (10%)</li> <li>• Annual stakeholder survey (5%)</li> <li>• Instructional superintendent’s assessment (5%)</li> <li>• Retention of high-performing teachers (10%)</li> <li>• Special education goal (10%)</li> </ul>	<ul style="list-style-type: none"> <li>• Goal setting with superintendent</li> <li>• Principal self-assessment against goals and leadership standards</li> <li>• Superintendent assessment on leadership standards</li> <li>• Evidence of principal’s work</li> </ul>	<ul style="list-style-type: none"> <li>• Quality Review</li> <li>• Goals and objectives (individually set between principal and supervisor)</li> <li>• Compliance</li> <li>• Attention to special education and ELL populations</li> <li>• Surveys: students, parents, teachers</li> <li>• Student attendance</li> </ul>
<b>Total Weight From These Measures</b>	60%	60%		
<b>Notes</b>		•4 Rating Categories	•4 Rating Categories •Includes different process for tenured versus non-tenured	•5 Rating Categories (uses 0-4 scale)

# Further Considerations

- Implementation fidelity at least as important as design decisions
- Monitoring needed for constituent feedback and medium and long-term impacts
- Continuous improvement necessary and desirable

# Appendix

## Overview of ISLLC 2008

- **Standard 1:** An education leader promotes the success of every student by facilitating the development, articulation, implementation, and stewardship of a vision of learning that is shared and supported by all stakeholders.
- **Standard 2:** An education leader promotes the success of every student by advocating, nurturing, and sustaining a school culture and instructional program conducive to student learning and staff professional growth.
- **Standard 3:** An education leader promotes the success of every student by ensuring management of the organization, operations, and resources for a safe, efficient, and effective learning environment.
- **Standard 4:** An education leader promotes the success of every student by collaborating with faculty and community members, responding to diverse community interests and needs, and mobilizing community resources.
- **Standard 5:** An education leader promotes the success of every student by acting with integrity, fairness, and in an ethical manner.
- **Standard 6:** An education leader promotes the success of every student by understanding, responding to, and influencing the political, social, economic, legal, and cultural context.



# Appendix

## Highlights of Research in Support of Task Force

- Over 90 documents provided to Task Force on website
  - Research reports about TLE topics from National Comprehensive Center on Teacher Quality; Dr. Laura Goe at ETS; National Center for Improvement of Assessment; The New Teacher Project; Vanderbilt University; McKinsey; Brookings Institution; New Teacher Center; CPRE; among others
  - Samples of teacher and principal assessment tools and rubrics
  - Summary of evaluation approaches in other states, districts
  - All discussion documents and presentations for work group meetings

Note: state and district examples included within this presentation come from research conducted in October and November using public sources, or materials from TLE-related convenings of states and districts.

# Appendix

## Technical Support Convenings Attended By Staff and Fellows Related to Task Force

- Ed Counsel/Gates network of states: TLE focus
  - Two in person meetings
  - Several webinars
  
- USDOE convening of RTTT winners
  - Two in person meetings
  - Series of webinars
  
- Gates Measures of Effective Teaching Partnership meetings

# Appendix

## Highlighted Bibliography

- Laura Goe, Courtney Bell, Olivia Little (2008). Approaches to Evaluating Teacher Effectiveness: A Research Synthesis. Available: [www.tqsource.org/publications/EvaluatingTeachEffectiveness.pdf](http://www.tqsource.org/publications/EvaluatingTeachEffectiveness.pdf)
- Herbert G. Heneman , Anthony T. Milanowski (2007). Assessing Human Resource Alignment: The Foundation for Building Total Teacher Quality Improvement. Available: [www.smhc-cpre.org/resources](http://www.smhc-cpre.org/resources).
- Teacher Evaluation Tools: Available: [www3.learningpt.org/tqsource/GEP/Default.aspx](http://www3.learningpt.org/tqsource/GEP/Default.aspx)
- Issue Analysis: Building a Thriving Teacher WorkForce. Teacher Evaluation 2.0. The New Teacher Project. Available: <http://www.tntp.org/files/Teacher-Evaluation-Oct10F.pdf>
- Melissa Brown-Sims (2010). Evaluating School Principals: Tips & Tools. Available: [www.tqsource.org/publications/KeyIssue\\_PrincipalAssessments.pdf](http://www.tqsource.org/publications/KeyIssue_PrincipalAssessments.pdf)