



MORE THAN A SCORE

Ideas for a better teacher evaluation system

March 2015

**NEW YORK STATE COUNCIL OF
SCHOOL SUPERINTENDENTS**
7 Elk Street ♦ Albany, NY 12207
518-449-1063
www.nyscoss.org

How Annual Professional Performance Review (APPR) works now

Each teacher or principal gets a “composite effectiveness score” based on 3 parts:

- 20% based on “state-provided growth score” based on performance of educator’s students on state assessments, or “comparable measures” developed locally for educators not covered by state assessments.
- 20% based on locally selected measures of student achievement.
- 60% based on other measures –
 - at least 31 points from multiple observations by principal or other administrator;
 - remaining points may come from observations by other professionals, survey feedback from students or parents, or review of lesson plans, student work, or other items.

Results from each component are added together to produce a “quality rating”

- Quality ratings: Ineffective, Developing, Effective, Highly Effective
- Scoring bands are used to convert points from each APPR component into quality ratings:

	State Growth Score/ Comparable Measures*	Locally Assessed 20%*	Other 60%**	Composite Effectiveness Score*
Ineffective	0-2	0-2	Locally negotiated	0-64
Developing	3-8	3-8	Locally negotiated	65-74
Effective	9-17	9-17	Locally negotiated	75-90
Highly Effective	18-20	18-20	Locally negotiated	91-100

* prescribed by state law/regulation

** determined through local negotiations

How a district APPR plan is developed and approved

Most of a district APPR plan is determined through negotiation with local unions, including:

- Measures used for the locally assessed 20% and “other 60%” in the composite effectiveness score and choice of “rubric” used in conducting observations.
- Scoring bands for other 60% measures
- Other elements such as procedures for resolving appeals of evaluations.

The state prescribes scoring bands for the 2 student performance-based measures in the composite effectiveness score.

The entire plan is subject to approval by the State Education Department.

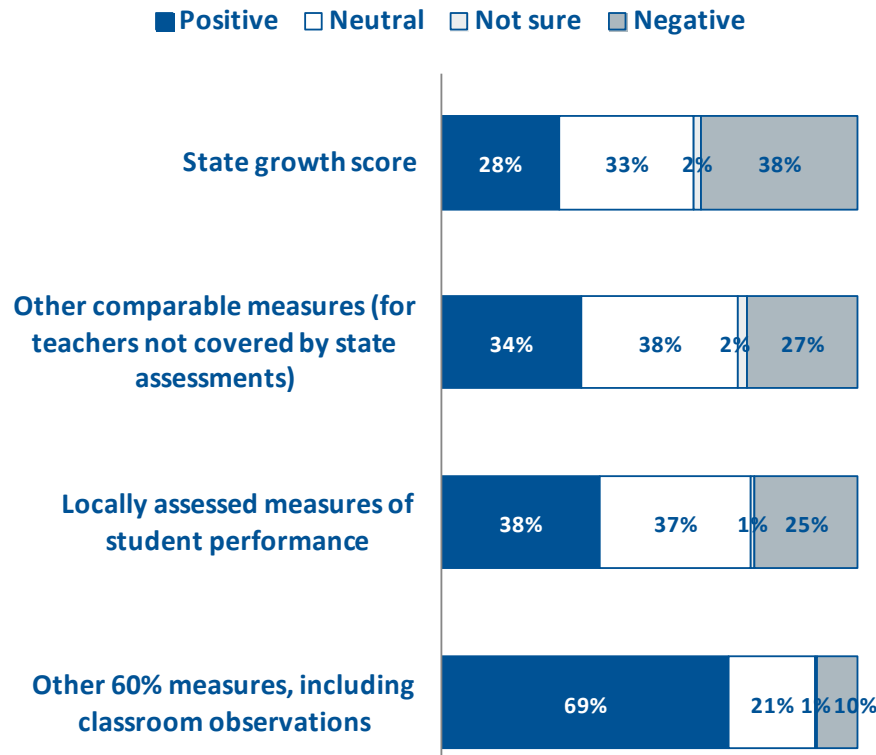
Districts were required to negotiate and gain SED approval for original plans by January 17, 2013 or face loss of state aid.

Evaluations should be about more than a score

- The goal of an evaluation system is **not** to rank and sort teachers and principals into categories.
- The goal is to improve teaching and school leadership – to help raise student achievement.
- An evaluation system can help improve teaching and school leadership by:
 - Giving educators information to help improve their daily work; and
 - Giving districts information to make sound personnel decisions, including to identify educators who need to improve or be removed.
 - The law requires evaluation results to be a “significant factor” in employment decisions, including granting tenure, making promotions, and seeking dismissal.

What do superintendents think about APPR now?

How would you assess the *impact of the various components* of the evaluations in improving teaching in your district's schools?



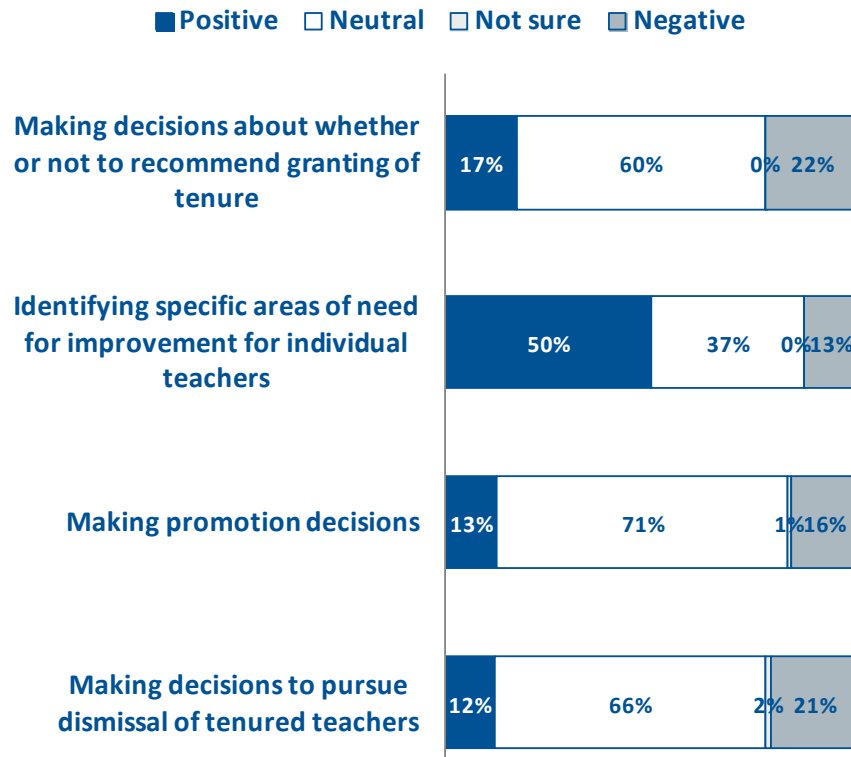
Superintendents say that as a result of APPR, their districts are doing classroom observations of their teachers more carefully and the observations are helping to improve classroom instruction.

No more than 38% of superintendents see any of the test-based measures as having a positive impact.

SOURCE: Council survey of superintendents, summer 2014.

What do superintendents think about APPR, continued

How would you assess the *impact of APPR in making personnel decisions* about teachers in your schools?



Superintendents see some value from APPR in providing feedback to help teachers improve their work.

At best, superintendents are withholding judgment about the value of APPR for making formal personnel decisions.

Anecdotally, superintendents report the system missing in both directions – good teachers rated poor, weak teachers rated strong.

SOURCE: Council survey of superintendents, summer 2014.

Superintendents on APPR in their own words...

I have very mixed responses to the APPR requirements. In some respects, I believe it is a waste of valuable time, effort and other resources... However, in other respects the process has led to some excellent and powerful professional development on shaping the classroom environment and practicing effective learning strategies. Principals were compelled to become instructional leaders, spending time in class-rooms and analyzing the teaching learning process. The teachers are much more engaged in professional development and students benefit from the new learning. So the bureaucratic aspect is ridiculously burdensome, but the impact on professional growth is significant. -- Lower Hudson Valley Suburb

The "other 60%" has been moderately helpful in systematizing good practices that were already more or less in place. They have also generated good conversations about teaching and learning. Test derived data has been far less useful and overly convoluted. -- Lower Hudson Valley Suburb

The APPR process is overly prescriptive and too time consuming. Time is taken from principals that could be used to assist struggling teachers. The most exemplary teacher has to go through the same process as all others. Wrong! -- Mid-Hudson Valley Rural

The mistake which was made, in my opinion, was tying teacher performance to these assessments before we really had time to work with the CCLS aligned assessments, understand their format and what they really tell us... In my opinion, this is the single biggest factor in the common core backlash we experienced this past year. The observation piece of the APPR has been very beneficial - we are using the Danielson Rubric - conversations between Principals and teachers have been very productive and are resulting in improved instructional practices. -- Western New York Rural

The APPR process has essentially made it more difficult to get rid of bad teachers. The observation rubrics have been helpful in getting focus on instruction. Student data from state tests are somewhat useful. The local assessments are all over the place - very time consuming and expensive - whether you buy them or develop them. The APPR process so far - is a boon to bad teachers - because it is so full of procedural issues and hoops to jump through - that school districts have great difficulty navigating given everything else going on. Those who negotiated this system have absolutely no concept about how systems work in public schools. The process needs to be simplified and made completely transparent - or it will fall of its own weight and complexity. -- Capital Region Suburb

APPR has prompted rich conversations about instruction that previously did not occur with regularity or structure. This is the main and maybe only benefit. -- Southern Tier City

What does Governor Cuomo propose?

1. Eliminate the locally selected 20% measures of student performance.
2. Increase the share based on the state provided growth score or comparable measures from 20% to 50%.
3. Base the remaining 50% of the composite effectiveness score on observations:
 - a. 35% on observations done by independent evaluators – a principal from another school or district, a college faculty member or another outside professional.
 - b. 15% on observations done by others
4. Have the State Education Department prescribe scoring bands for the observation-based components

(currently the state prescribes scoring bands for student performance based components; bands for the “other 60%” measures are locally negotiated).

Our reactions to the Governor's proposals

- 👎 **Oppose** basing 50% of score on state growth score or comparable measures – *this would amplify stress over standardized testing and put too much reliance on a measure superintendents question.*
- 👎 **Oppose** requiring use of independent evaluators – *this would undermine one part of APPR that is working – interactions between teachers and administrators focused on improving daily practice.*
- 👍 **Support** having the state prescribe scoring bands for all APPR components, including the current other 60% measures (which include classroom observations) – *this would address concerns about overly positive evaluation results, without the damage the first two proposals would cause.*

Our goals for a better evaluation system:

1. Maximize value of the evaluations as a tool for improving instruction and school leadership.
2. Improve the soundness of individual results so that they are seen as a fair and reliable tool in making formal employment decisions.
3. Streamline administrative demands of the system to enable leaders and teachers to devote their time to other priority responsibilities.
4. Reduce the impact of system in creating needs for student testing and in contributing to an over-emphasis on standardized test results as a “gold standard” for gauging the performance of students, educators, schools, and districts.
5. Structure any changes in state requirements to avoid or minimize the need for additional local collective bargaining and staff training to implement the changes.

The system should create a fork in the road: help successful teachers become even more effective and put inferior teachers on a structured path toward improvement or removal.

One option: a matrix

1. Each teacher would receive two ratings, one based on observations and practice measures, one based on impact on student learning.
2. The observation and practice measure could be the current “other 60%” measures.
3. The student impact measure could be the state provided growth score for teachers covered by state assessments, or a school- or group-wide measure of student performance on state assessments for other teachers. This could be a transitional measure pending recommendations from the commission suggested in #6 below. Multi-year averages of student data could be used to address concerns about inaccuracies with a single year result.
4. The State Education Department should prescribe scoring bands for each component; the bands should have similar or identical distributions of points across the ratings categories.
5. Consequences would vary based on the combination of results on the two measures (next slide).
6. A commission of practitioners could be created to recommend alternatives for the measure of impact on student learning and to recommend what consequences would apply for each combination of ratings. An option could be to maintain the current APPR system, pending the commission’s recommendations.

An example...

Hypothetical Matrix Structure for Teacher Evaluations					
Observation & Practice Measures ("Other 60%")	Highly Effective	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Effective	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Developing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Ineffective	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		Ineffective	Developing	Effective	Highly Effective
		Student Performance Measures			

- Green zone – teacher develops personal growth plan
- Yellow zone – district creates and assigns teacher improvement plan
- Red zone – year may not be counted toward satisfying probation requirement toward tenure

Another option: “75/25”

1. Retain the single composite score, but base 75% on observation and practice measures, 25% on a measure of impact on student learning.
2. Could use same measures as in matrix option:
 - The observation and practice measure could be current “other 60%” measures.
 - The student impact measure could be the state provided growth score for teachers covered by state assessments, or a school- or group-wide measure of student performance on state assessments for other teachers. This could be a transitional measure pending recommendations from the commission proposed in #3 below. Multi-year averages of student data could be used to address concerns about inaccuracies with a single year result.
 - The State Education Department should prescribe scoring bands for each component; the bands should have similar or identical distributions of points across the ratings categories.
3. As with matrix model, a commission of practitioners could be created to recommend alternatives for the measure of impact on student learning. An option could be to maintain the current APPR system, pending the commission’s recommendations.

Other proposals by the Governor

- 👍 The Council **supports** the Governor’s proposal to extend the probationary period for earning tenure from 3 to 5 years and to clarify the authority of districts to terminate teachers during the probationary period.
- 👎 The Council **opposes** the Governor’s proposal to require teachers to earn Effective or Highly Effective APPR ratings 5 years in a row in order to receive tenure.
- 👍 The Council **supports** the Governor’s proposed changes to disciplinary procedures for tenured educators (section 3020-a of the Education Law):
 1. Revoke certification immediately of any person convicted of a violent felony against a child
 2. Authorize suspension without pay for any teacher charged with child abuse
 3. Require all 3020-a hearings to be conducted by a single hearing officer
 4. Provide two consecutive ineffective ratings shall constitute prima facie evidence of incompetence rebuttable only by clear and convincing evidence of fraud in calculation of composite effectiveness ratings
 5. Provide for reciprocal discovery – “full and fair disclosure of witnesses and evidence by both parties...”
 6. Failure of district to remediate or correct the behavior of the employee shall not be a defense and shall not be considered by the hearing officer



OUR PUBLIC SCHOOLS

**OF THE PEOPLE...
BY THE PEOPLE...
FOR THE PEOPLE...**