

APPR Discussion  
Board of Regents  
May 18, 2015

# Timeline Related to New York State's Evaluation System

## **2007:**

- Chapter 57 of the Laws of 2007 added Education Law §3012-b which, for purposes of evaluating a candidate for tenure, required an evaluation of the extent to which the teacher successfully utilized analysis of available student performance data and other relevant information when providing instruction, but provided that tenure couldn't be granted or denied based on student performance data.

## **2008:**

- Chapter 57 of the Laws of 2008 created a sunset for Education Law §3012-b, which expired on July 1, 2010. Chapter 57 of the Laws of 2008 also created a legislative commission on value added assessments, but this commission was never empanelled.

## **2010:**

- Governor Paterson signed Chapter 103 of the Laws of 2010, which added a new section 3012-c to the Education Law, establishing a comprehensive evaluation system for teachers and principals, effective July 1, 2010.
- USDE announced that New York is selected for a RTTT award of approximately \$700M.

# Timeline Related to New York State's Evaluation System

## **2011-12:**

- First year of State-provided growth score results for all 4-8 ELA and math teachers and their building principals.
- Evaluations for teachers and principals are done in some NYS districts (e.g., School Improvement Grant and Teacher Incentive Fund).
- Evaluation Law is revised. Governor Cuomo signed the bill into law on March 27, 2012 (Chapter 21 of the Laws of 2012). The Board of Regents adopted emergency regulations to conform to the major 2012 legislative changes.

## **2012-13:**

- All NYS districts must have an approved APPR plan by January 17, 2013 or risk state aid increases.
- Evaluations for teachers and principals are done in all districts except for NYC. NYC is required by law to have a State-imposed evaluation plan.
- The Legislature further amends the Evaluation Law (Part A of Chapter 57 of the Laws of 2013).

## **2013-14:**

- Second year of evaluations for all districts in NYS, except NYC. First year for NYC.
- The Legislature further amends the Evaluation Law (Chapter 56 of the Laws of 2014).
- NYC's state-imposed plan yields greater differentiation than systems in place in other states.

## **2014-15:**

- Third year of evaluations for all districts in NYS, except NYC. Second year for NYC.
- The Legislature further amends the Evaluation Law (Chapter 56 of the Laws of 2015).

# Milestones for Implementing Education Law §3012-d

Date	Milestone
<b>May 7, 2015</b>	APPR Learning Summit. Statewide day of discussion to solicit State and national expert feedback
<b>May 18, 2015</b>	Present recommendations to the Board of Regents and seek feedback for June regulations
<b>June 15, 2015</b>	Board of Regents meeting to approve Commissioner's regulations

# Statewide Stakeholder Engagement

The Department received more than 3,000 emails to [eval2015@nysed.gov](mailto:eval2015@nysed.gov). In addition, staff has consulted the following State stakeholder groups at the Learning Summit and in individual meetings:

## Statewide Stakeholder Engagement\*

District Superintendents	NYSSBA
Big 5 Districts/NYC DOE	NYSCOSS
Small City School Districts	NYSPTA and related parent groups
NYSUT/UFT	Members of the NYSED Assessment Technical Advisory Committee
SAANY/ESSAA/NYSFSA	Professional Standards and Practice Board (PSPB)
Commissioner's Advisory Council for NYS Teachers	NYSCOSS Commissioner's Advisory Council

\*Stakeholders listed above that did not participate directly in the Learning Summit were provided with time for individual conversations with Department staff.

# National Expert Engagement

The Department has consulted with the following national experts via the Learning Summit\*:

- Catherine Brown, Center for American Progress
- Stephen Caldas, Manhattanville College
- Lesley Guggenheim, The New Teacher Project
- Sandi Jacobs, National Council on Teacher Quality
- Tom Kane, Harvard Graduate School of Education
- Aaron Pallas, Teachers College, Columbia University
- Jesse Rothstein, University of California, Berkeley

\*National experts who were not able to participate were invited to submit comments and materials to be posted on the Learning Summit web page.

# Guiding Principles

**These recommendations are guided by the following principles:**

- **We remain mindful that the evaluation law was adopted in 2010 and re-affirmed by state elected officials four additional times (in 2012, 2013, 2014, and 2015).**
- **The purpose of New York State's evaluation system is and should be to support teaching, learning, and talent management decisions.**
- **Technical parameters alone will not ensure that teachers receive meaningful feedback. This will require extensive professional development and a comprehensive approach to talent management by school districts.**

# Summary of Recommendations: Observations

- **Technical parameters can support but cannot ensure meaningful feedback for teachers and principals. Proper professional development is critical.**
- **Existing observation rubrics should remain in place.**
- **Observations should focus on specific observable professional behaviors, while ensuring that all observable teaching standards are assessed each year. Artifacts should be allowed to the extent they constitute evidence of an otherwise observable rubric subcomponent.**
- **Observation parameters (number, duration, etc.) should be established as differentiated minimum standards that allow for local best practices. “School Building” generally should be defined by BEDS code for the purpose of independent observers.**
- **Multiple observations (principal/supervisor, independent, peer) should be combined through a weighted average. Weights should reflect the role of the principal as the instructional leader of a school.**
- **To ensure that the process for scoring rubrics is comparable across districts, scores on observations could be expressed as a percentage of possible points for observed subcomponents (in addition to the familiar 1-4 ratings).**
- **Scoring ranges to create Observation HEDI scores should be set based on a common-sense approach to percentage of points earned (65%, 75%, 90%).**



# Summary of Recommendations: Student Performance

- **Growth can be represented through the use of the required student growth subcomponent and the optional student growth subcomponent.**
- **The Department should convene advisory groups to recommend enhancements for the next generation of assessments and growth models (performance-based assessment tasks, additional growth covariates, new high school growth metrics, multi-year growth models, possible adjustments to the normative method to determine HEDI ratings and/or development of criterion-referenced measures of growth).**
- **Until next generation growth models are available and adopted, existing methods to establish growth scores should be continued. SLO targets should reflect a year of expected student growth, which will vary by a student's academic preparedness and learning needs.**
- **Multiple growth measures (i.e., required and optional student growth subcomponents) should be combined through a weighted average. Weights should not incentivize additional tests for students.**
- **Once multiple measures are combined through a weighted average, it is recommended that scoring ranges to create a Student Performance HEDI score should be set based on a common-sense approach to percentage of points earned (e.g., 65%, 75%, 90%).**

# Summary of Recommendations: Other Areas

- **Some aspects of the principal evaluation should be different than the teacher evaluation. Similar to teachers, independent observers could include anyone outside of the principal's building, defined by BEDS code. Although professional goal setting is now a prohibited element of principal evaluations, organizational goal setting could be used to the extent that it is evidence of an observable component of the practice rubric.**
- **With few exceptions, the provisions of 3012-c should be carried forward to 3012-d to limit the burden of new negotiations where local practices are successful.**
- **Waivers from the general prohibition against assigning a student to an Ineffective teacher for two consecutive years should be granted only if the district has an improvement and/or removal plan in place for the teacher in question, consistent with law and regulation.**
- **Short-term hardship waivers from the November 15 deadline should be accompanied by good faith attempts to collectively bargain and train for the new system.**
- **In response to field request, the Department should provide a model plan for optional field consideration once regulations have been adopted.**

# Observation Category

**Current:** Various scoring ranges exist for observation rubrics that are selected and implemented locally under Education Law §3012-c.

	NYSED-Recommended; determined locally		NYSUT – Recommended; determined locally		NYC – Commissioner Imposed	
	<i>Min</i> (% of points)	<i>Max</i> (% of points)	<i>Min</i> (% of points)	<i>Max</i> (% of points)	<i>Min</i> (% of points)	<i>Max</i> (% of points)
H	3.7 <b>(93%)</b>	4.0 (100%)	3.5 <b>(88%)</b>	4.0 (100%)	3.26 <b>(82%)</b>	4.0 (100%)
E	2.9 <b>(73%)</b>	3.6 (92%)	2.5 <b>(63%)</b>	3.4 (87%)	2.51 <b>(63%)</b>	3.25 (81%)
D	1.9 <b>(48%)</b>	2.8 (72%)	1.5 <b>(38%)</b>	2.4 (62%)	1.76 <b>(44%)</b>	2.50 (62%)
I	1.0 <b>(0%)</b>	1.8 (47%)	1.0 <b>(0%)</b>	1.4 (37%)	1.0 <b>(0%)</b>	1.75 (43%)

**Current:** In order to combine into the 20/20/60 = 100-point scale, various conversions of the 60-point scoring range to HEDI categories exist. These conversions can result in very little differentiation in the Highly Effective and Effective ranges.

	NYSUT – Recommended; determined locally		NYC – Commissioner Imposed	
	<i>Min</i> (% of points)	<i>Max</i> (% of points)	<i>Min</i> (% of points)	<i>Max</i> (% of points)
H	59 <b>(98%)</b>	60 (100%)	55 <b>(92%)</b>	60 (100%)
E	57 <b>(95%)</b>	58 (97%)	45 <b>(75%)</b>	54 (91%)
D	50 <b>(83%)</b>	56 (94%)	39 <b>(65%)</b>	44 (74%)
I	0 <b>(0%)</b>	49 (82%)	0 <b>(0%)</b>	38 (64%)

The 60-point conversion is no longer necessary because of the matrix approach to combining Student Performance and Observation categories.

**Recommended:** Under Education Law §3012-c, districts have negotiated a variety of ways for converting rubric scores to 0-60 points and corresponding HEDI rating categories. SED’s recommendation is to treat the overall observation score as a percentage of possible points and apply common-sense percentage-of-points cut scores that will be comparable across locally-selected and locally-weighted rubrics.

	NYSUT and UFT Recommended		NYSED Recommended		NYCDOE Recommended	
	<i>Min</i> (% of points)	<i>Max</i> (% of points)	<i>Min</i> (% of points)	<i>Max</i> (% of points)	<i>Min</i> (% of points)	<i>Max</i> (% of points)
<b>H</b>	3.5 <b>(88%)</b>	4.0 (100%)	3.6 <b>(90%)</b>	4 (100%)	3.76 <b>(94%)</b>	4.0 (100%)
<b>E</b>	2.5 <b>(63%)</b>	3.49 (87%)	3 <b>(75%)</b>	3.59 (90%)	2.76 <b>(69%)</b>	3.75 (93%)
<b>D</b>	1.5 <b>(38%)</b>	2.49 (62%)	2.6 <b>(65%)</b>	2.99 (74%)	1.76 <b>(44%)</b>	2.75 (68%)
<b>I</b>	1 <b>(0%)</b>	1.49 (37%)	1 <b>(0%)</b>	2.59 (64%)	1 <b>(0%)</b>	1.75 (43%)

**Note that no technical parameter will ensure that teachers receive meaningful feedback about their relative strengths and weaknesses. Meaningful feedback will occur only if quality training is provided that incorporates an understanding of the technical parameters.**

# **Recommended: Scoring Ranges for Observation Category**

(necessary to be entered into the Evaluation Matrix)

	Statewide	
	<i>Min</i> (% of points)	<i>Max</i> (% of points)
H	3.6 <b>(90%)</b>	4 (100%)
E	3 <b>(75%)</b>	3.59 (90%)
D	2.6 <b>(65%)</b>	2.99 (74%)
I	1 <b>(0%)</b>	2.59 (64%)

## How it would work:

- Each observation (principal/supervisor, independent, peer) would be completed using a 1-4 rubric, producing an overall score between 1-4.
- Multiple observations would be combined using a weighted average (depending on the weights adopted in regulation), producing an overall Observation category score between 1-4.
- Because evaluation categories will be combined according to the matrix, there is no need to convert into a 60-point scale.
- This overall observation score would be converted into an HEDI rating and entered into the Evaluation Matrix to determine the overall evaluation rating.

# **Recommended: Subcomponent Weights for Observation Category**

## **How it would work\*:**

For the independent evaluator, recommended percentages include:

- NYCDOE: 5 to 20%, determined by the district.
- NYSUT: 1 to 5%, subject to local bargaining.
- UFT: 1 to 25%, subject to local bargaining

- If there are no optional peer observations, the role of the principal as instructional leader would be reflected as –
  - Principal/supervisor observations weighted 80%
  - Independent observations weighted 20%
- If there are optional peer observations, the role of the principal as instructional leader would be reflected as –
  - Principal/supervisor observations weighted 80%
  - Peer observations weighted 10%
  - Independent observations weighted 10%

\*The alignment of these subcomponents among each other and with Student Performance category will be subject to audit and corrective action as permitted under Sections 9(a) and 9(b) of Education Law §3012-c.



# Evaluation Matrix

The statute mandates the use of the “matrix” below to determine a teacher’s composite score based on the two categories of the evaluation (see §3012-d (5)):

		Observation			
		Highly Effective (H)	Effective (E)	Developing (D)	Ineffective (I)
Student Performance	Highly Effective (H)	H	H	E	D
	Effective (E)	H	E	E	D
	Developing (D)	E	E	D	I
	Ineffective (I)	D*	D*	I	I

\* If a teacher is rated Ineffective on the Student Performance category, and a State-designed supplemental assessment was included as an optional subcomponent of the Student Performance category, the teacher can be rated no higher than Ineffective overall (see §3012-d (5)(a) and (7)).

# **Recommended: Number, Frequency, and Duration of Observations**

How it would work:

- **At least two observations (one principal/supervisor; one independent), each at least 20-minutes in duration for non-tenured teachers and tenured teachers who were not rated Highly Effective or Effective overall the prior year.**
  - **Independent observers must be trained and selected by the district. This may include observations by other administrators, department chairs, or peers (e.g., teacher leaders on career ladders) so long as they are not from the same building (defined as same BEDS code) as the teacher being evaluated.**
- **The required minimum duration for the two required observations is reduced to 10 minutes each for a tenured teacher rated Highly Effective or Effective overall in the prior year.**
- **At least one observation must be unannounced.**
- **In addition to the above, short walkthrough observations (5-10 minutes) are permissible.**
- **Observations may occur live or by live or recorded video.**

# **Recommended: Observation Rubrics**

## **How it would work:**

- **Districts can choose a teacher practice rubric from a menu of state-approved practice rubrics. The currently approved list, under Education Law §3012-c, will remain in effect.**
- **The evaluator may select a limited number of observable rubric subcomponents for focus within a particular observation (which determines the total number of possible points), so long as all observable Teaching Standards/Domains are addressed across the total number of annual observations.**
- **Under Education Law §3012-d(6), artifacts are a prohibited element of teacher evaluations. However, evidence documented during an observation cycle may be considered to the extent that it constitutes evidence of an otherwise observable rubric subcomponent (e.g., a lesson plan viewed during the course of the observation cycle may constitute evidence of professional planning).**
- **Teaching Standards/Domains that are part of the rubric, but not observable during the classroom observation, may be observed during a pre-observation conference or post-observation review or other natural conversations between the teacher and the principal/supervisor and incorporated into the observation score. Regardless, points shall not be allocated based on artifacts submitted to or reviewed by the evaluator outside of the observation cycle.**

# Student Performance Category

**Current: Scoring Ranges for Student Learning Objective 20 Points**  
 (necessary for combination into a 20/20/60 100-point scale when a state-provided growth score is not available)

Rating	Rest of State (Recommended in guidance)		Percent of Points (20 possible points)	NYC (Imposed)		
	Scoring Range	Percent of Students Meeting Target		Scoring Range	Percent of Students Meeting Target	
I	0	0-65%	0%	0	0-4%	I
I	1		5%	1	5-8%	I
I	2		10%	2	9-12%	I
D	3	66-77%	15%	3	13-16%	I
D	4		20%	4	17-20%	I
D	5		25%	5	21-24%	I
D	6		30%	6	25-28%	I
D	7		35%	7	29-33%	I
D	8		40%	8	34-38%	I
E	9	78-85%	45%	9	39-43%	I
E	10		50%	10	44-48%	I
E	11		55%	11	49-54%	I
E	12		60%	12	55-59%	I
E	13		65%	13	60-66%	D
E	14		70%	14	67-74%	D
E	15		75%	15	75-79%	E
E	16		80%	16	80-84%	E
E	17	85%	17	85-89%	E	
H	18	86-100%	90%	18	90-92%	H
H	19		95%	19	93-96%	H
H	20		100%	20	97-100%	H

See Appendix for flexibility to address sensitivity in small group situations.

## Current: Scoring Ranges for State-Provided Growth Score 20 Points\*

(necessary for combination into a 20/20/60 100-point scale)

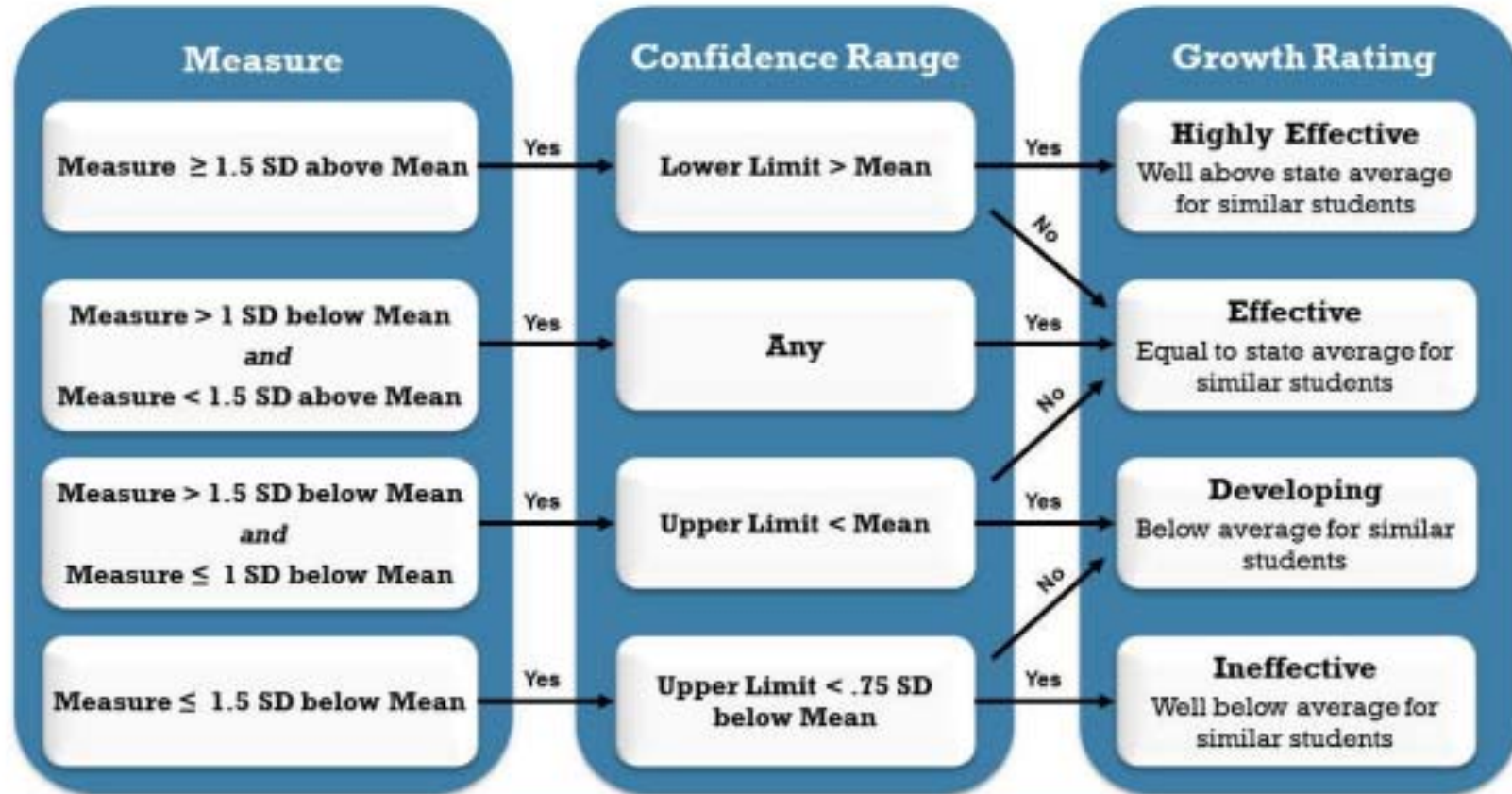
Rating	Rest of State		Percent of Points	NYC		Rating
	Scoring Range	MGP Range		Scoring Range	MGP Range	
I	0	3-28	0%	0	3-23	I
I	1	29-32	5%	1	24	I
I	2	33-35	10%	2	25	I
D	3	29-35	15%	3	26	I
D	4	36	20%	4	27	I
D	5	37	25%	5	28	I
D	6	38	30%	6	29	I
D	7	39	35%	7	30	I
D	8	40	40%	8	31	I
E	9	36-43	45%	9	32	I
E	10	44-45	50%	10	33	I
E	11	46-48	55%	11	34	I
E	12	49-50	60%	12	35	I
E	13	51-52	65%	13	29-37	D
E	14	53-55	70%	14	38-40	D
E	15	56-57	75%	15	36-48	E
E	16	58-61	80%	16	49-55	E
E	17	62-68	85%	17	56-68	E
H	18	67-68	90%	18	67-68	H
H	19	69-72	95%	19	69-72	H
H	20	73-94	100%	20	73-94	H

Each score between 0 and 20 is associated with a “mean growth percentile” (MGP) value, or the mean of each teacher’s individual student growth percentiles (SGP). SGPs are a measure of growth for each student compared to similar students.

\*MGP ranges are based on 13-14 school year results and may differ slightly in future years based on the distribution of teachers’ MGPs.

## Current: Method to determine points with the Scoring Ranges for State-Provided Growth Score 20 Points

- For each teacher, an MGP and a confidence range are reported, representing the upper and lower limits on the MGP within a 95% statistical confidence.
- To determine a teacher's HEDI rating, the teacher's MGP value is compared to the mean and standard deviation of MGPs for all teachers, and the teacher's confidence range is used to confirm the rating category in which he or she should be placed.
- After all teachers are assigned to a rating category, growth score points (0-20) are distributed across teachers in each category so that higher MGPs earn higher points, and the number of teachers receiving each score is approximately proportional to the number of score point values in the category.



NYCDOE suggested cut scores for MGPs that are more directly related to an above/below the mean determination.

Measure	Confidence Range	Growth Rating
Measure > 1.5 SD above Mean	Lower Limit > Mean	Highly Effective
Measure above Mean	Any	Effective
Measure below Mean	Upper Limit > Mean	Effective
Measure below Mean	Upper Limit < Mean	Developing
Measure > 1.5 SD below Mean	Upper Limit < below Mean	Ineffective



NYSUT suggested cut scores for MGPs that would define Ineffective as two or more standard deviations below the mean (vs. 1.5 currently), reducing the percent of Ineffective ratings.

Measure	Confidence Range	Growth Rating
Measure $\geq$ 1.5 SD above Mean	Lower Limit $>$ Mean	Highly Effective
Measure $<$ 1.5 above Mean	Any	Effective
Measure $>$ 1 SD below Mean	Any	Effective
Measure $>$ 2 SD below mean	Upper Limit $<$ Mean	Developing
Measure $\leq$ 1 SD below Mean	Upper Limit $<$ Mean	Developing
Measure $\leq$ 2.0 SD below Mean	Upper Limit $<$ .75 below Mean	Ineffective

# **Recommended: Parameters for growth scores**

## **How it would work:**

- **Maintain existing normative method to establish growth scores for the required and optional student growth subcomponents on existing State assessments and new State-designed supplemental assessments.**
- **Maintain in the growth model the full list of characteristics described in Section D1 of the APPR Guidance (e.g., prior academic history, English language learner status, disability status, poverty).**
- **Explore with stakeholders and technical experts future assessment and metrics options, new covariates for the growth model, new high school growth metrics, multi-year growth models, possible adjustments to normative method to determine HEDI ratings, and/or criterion-referenced measures of growth.**
- **Superintendents continue to have sole discretion to determine SLO targets. These targets must reflect a year of expected student growth, which will vary by a student's academic preparedness and learning needs (see Appendix).**
- **SLOs may incorporate group measures, including school-wide measures. Linked group measures (group measures based only on a teacher's roster) are encouraged.**

## **Recommended:** Optional Student Growth Subcomponent with No Additional Testing

### **Examples of how it could work:**

- **Computed by the State based on the percentage of students who achieve a State-determined level of growth on a State assessment (e.g., at least average for similar students). Such measures could incorporate multiple years of data.**
- **State calculated school-wide results based on the State-provided growth scores of all students in the school taking the grades 4-8 State ELA or math assessment.**
- **Locally-computed school-wide results based on all or a subset of State-provided growth scores.**

**Recommended:** Optional Student Growth Subcomponent and/or Required State Growth SLOs with Locally-Selected Additional Testing

**How it would work:**

- **In order to accommodate locally-selected additional testing, the Department will issue, with advice from stakeholders and experts in assessment and growth metrics, a Request for Qualifications (RFQ) for assessments certified to provide acceptable instructional and psychometric qualities and the ability to generate acceptable measures of growth consistent with the requirements of §3012-d:**
  - 1) **As an optional locally-selected State-designed supplemental assessment to be used with a State-provided or approved growth model; and/or**
  - 2) **As a measure of expected student growth to be used in the required student growth subcomponent for SLOs that do not use an existing State assessment.**

**Recommended:** Scoring Ranges for Student Performance Category  
 (necessary to be entered into the Evaluation Matrix)

	Statewide	
	<i>Min</i> (% of points)	<i>Max</i> (% of points)
H	18 <b>(90%)</b>	20 (100%)
E	15 <b>(75%)</b>	17 (85%)
D	13 <b>(65%)</b>	14 (70%)
I	0 <b>(0%)</b>	12 (60%)

**How it would work:**

- Each performance measure (Required Student Growth subcomponent and Optional Student Growth subcomponent) would result in a growth score between 0 and 20 points.
- Multiple measures would be combined using a weighted average (depending on the weights adopted in regulation), producing an overall Student Performance category score between 0 and 20 points.
- This overall student performance score would be converted into a HEDI rating and entered into the Evaluation Matrix to determine the overall evaluation rating.

# Recommended: Detailed Scoring Ranges for Growth Scores

Note that, for SLOs, NYSUT and UFT recommended:  
 I = 0-29% meeting target  
 D = 30-54% meeting target  
 E = 55-84% meeting target  
 H = 85-100% meeting target

See Appendix for flexibility to address sensitivity in small group situations.

Rating	SLOs		Percent of Points	State-Provided Growth Scores		Rating
	Percent of Students Meeting Target	Scoring Range		Scoring Range	MGP Range	
I	0-4%	0	0%	0	3-23	I
I	5-9%	1	5%	1	24	I
I	10-14%	2	10%	2	25	I
I	15-19%	3	15%	3	26	I
I	20-24%	4	20%	4	27	I
I	25-29%	5	25%	5	28	I
I	30-34%	6	30%	6	29	I
I	35-39%	7	35%	7	30	I
I	40-44%	8	40%	8	31	I
I	45-49%	9	45%	9	32	I
I	50-54%	10	50%	10	33	I
I	55-59%	11	55%	11	34	I
I	60-64%	12	60%	12	35	I
D	65-69%	13	65%	13	29-37	D
D	70-74%	14	70%	14	38-40	D
E	75-79%	15	75%	15	36-48	E
E	80-84%	16	80%	16	49-55	E
E	85-89%	17	85%	17	56-68	E
H	90-92%	18	90%	18	67-68	H
H	93-96%	19	95%	19	69-72	H
H	97-100%	20	100%	20	73-94	H

\*MGP ranges are based on 13-14 school year results and may differ slightly in future years based on the distribution of teachers' MGPs.

# **Recommended: Subcomponent Weights for Student Performance Category**

(necessary to be entered into the Evaluation Matrix)

## **How it would work\*:**

- **If there is no optional student growth subcomponent:**
  - **Required student growth subcomponent (State-provided growth scores or SLOs) would be weighted 100%.**
- **If there is an optional student growth subcomponent, the following weights are recommended to avoid the creation of an incentive for additional testing:**
  - **Required student growth subcomponent (State-provided growth scores or SLOs) would be weighted 80%.**
  - **The optional student growth subcomponent would be weighted 20%.**

For the required student growth subcomponent, recommended percentages include:

- NYSUT: no more than 20%
- UFT: no more than 40%

\*The alignment of these subcomponents among each other and with Observation category will be subject to audit and corrective action as permitted under Sections 9(a) and 9 (b) of Education Law §3012-c.

# Evaluation Matrix

The statute mandates the use of the “matrix” below to determine a teacher’s composite score based on the two categories of the evaluation (see §3012-d (5)(b)):

		Observation			
		Highly Effective (H)	Effective (E)	Developing (D)	Ineffective (I)
Student Performance	Highly Effective (H)	H	H	E	D
	Effective (E)	H	E	E	D
	Developing (D)	E	E	D	I
	Ineffective (I)	D*	D*	I	I

\* If a teacher is rated Ineffective on the Student Performance category, and a State-designed supplemental assessment was included as an optional subcomponent of the Student Performance category, the teacher can be rated no higher than Ineffective overall (see §3012-d (5)(a) and (7)).



# **Recommended: Principal Evaluation**

## **How it would work:**

**Parameters for teacher evaluations would largely apply to principals as well:**

- 1) For the required student growth subcomponent, principals would continue to use State-provided growth scores where available. All other principals would use SLOs with the same options as for teachers.**
- 2) For the optional student growth subcomponent, principals would have the same options as teachers, except that school-wide ELA and math measures would not be allowed.**
- 3) Video will not be allowed for the principal observation measures.**
- 4) Similar to teachers, independent observers may include anyone outside of the principal's building, defined by BEDS code (superintendent, other principals, department chairs/directors).**
- 5) Under Education Law §3012-d(6), professional goal-setting is now a prohibited element of principal evaluations. However, organizational goal-setting may be used to the extent that it is evidence of an observable component of the practice rubric.**
- 6) Similar to teachers, districts can choose a principal practice rubric from a menu of state-approved practice rubrics. The currently approved list, under Education Law §3012-c, will remain in effect.**

## Recommended: Continuation of 3012-c

Provision of §3012-c	Area of the Law	Recommended for carryover to §3012-d?	Revisions
§3012-c(2)(d)	Evaluator training	Yes	<ol style="list-style-type: none"> <li>1) Add language on training of independent and peer observers</li> <li>2) Eliminate requirements to the extent they do not comply with 3012-d</li> </ol>
§3012-c(2)(k)	Submission of plans	Yes	<ol style="list-style-type: none"> <li>1) Eliminate 9/1 deadline for approval of plans</li> <li>2) Add 3/1 deadline for submission of material changes</li> <li>3) Eliminate references to 2012 school year</li> <li>4) Eliminate references to annual or multi-year plans</li> <li>5) Eliminate language requiring written list of deficiencies</li> </ol>
§3012-c(2)(k-1)	Material changes to reduce assessments	Yes	No changes needed
§3012-c(2)(k-2)	Reduction of time spent on field tests	Yes	No changes needed
§3012-c(2)(l)	Triborough amendment	Yes	No changes needed

## Recommended: Continuation of 3012-c

Provision of §3012-c	Area of the Law	Recommended for carryover to §3012-d?	Revisions
3012-c(4)	TIPs/PIPs	Yes	<ol style="list-style-type: none"> <li>1) Provide for management discretion on developing and implementing TIPs/PIPs</li> <li>2) Require plans to be implemented by October 1</li> </ol>
3012-c(5)	Appeals	Yes	<ol style="list-style-type: none"> <li>1) Clean up to be consistent with §3012-d</li> </ol>
3012-c(5-a)	NYC appeals	Yes	<ol style="list-style-type: none"> <li>1) Clean up to be consistent with §3012-d</li> </ol>
3012-c(9)	Monitoring	Yes	<ol style="list-style-type: none"> <li>1) Preserve the Department's authority to monitor as intended by the statute</li> <li>2) Clean up to be consistent with §3012-d</li> </ol>
3012-c(10)	FOIL/personal privacy of APPR data	Yes	<ol style="list-style-type: none"> <li>1) Clean up to be consistent with current practice and §3012-d</li> </ol>

## **Recommended: Waiver to Assign Students to an Ineffective Teacher for Two Consecutive Years**

### **How it would work:**

- **If a district wishes to assign a student to an Ineffective teacher in the same subject for two consecutive years, the district must request a waiver.**
- **Waivers may be assigned if the district cannot make alternate arrangements (e.g., too few teachers qualified to teach the subject).**
- **Since consecutive assignment to an Ineffective teacher has a demonstrated negative impact on a student, waivers will be granted only if a true hardship is demonstrated and the district has an improvement and/or removal plan in place for the teacher in question, consistent with law and regulation.**

## **Recommended: Hardship Waiver for November 15 Approval Deadline**

### **How it would work:**

- **To be considered for a waiver, as part of its submission, the district would need to submit evidence of its good-faith attempts to negotiate a new APPR plan consistent with 3012-d and train staff in the new required procedures prior to November 15 .**
- **If a waiver is not granted or district does not meet the deadline to apply, the district forfeits the increase in state aid for 2015-16 and must secure approval for an evaluation system aligned to 3012-d when a successor agreement is reached.**
- **The previously approved (2014-15) APPR plan remains in effect during any approved waiver period pursuant to 3012-d(12).**
- **If granted, waivers will be in effect for renewable two-month periods.**
- **If a district wishes to request an additional two-month waiver, evidence of additional good faith collective bargaining and appropriate training since the last waiver must be submitted two weeks prior to the expiration of the current waiver.**
- **APPR plans approved prior to March 1, 2016 will apply to the 2015-16 school year. Plans approved after March 1, 2016 will apply to the 2016-17 school year.**
- **The final deadline for plan approval to secure 2015-16 state aid increases is September 1, 2016.**

# Appendix

# Student Learning Objectives (SLOs)

# Current: Scoring Ranges for Small n Student Learning Objectives 20 points

Comes from NYSED Guidance on Setting SLOs with small “n” sizes:  
<https://www.engageny.org/resource/alternative-target-setting-models-within-student-learning-objectives-slos>

- Points, from 0-3, are assigned based on each student’s movement from a baseline performance level from 1-4 to a summative performance level from 1-4 aligned with the qualitative descriptions.
- Points are then averaged for all students on a teacher’s course roster.

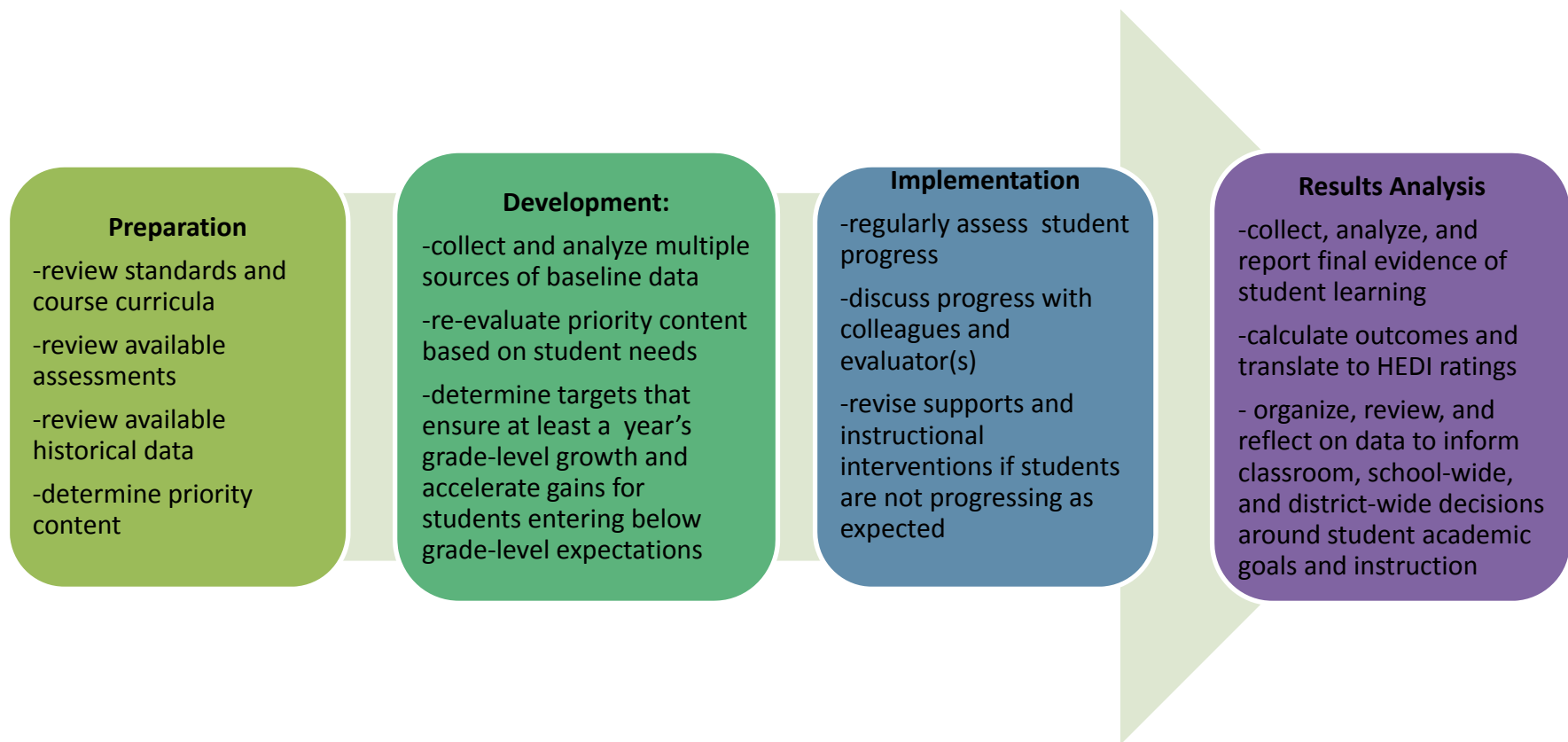
**Level 1= performance is well-below average/expectations**  
**Level 2= performance is below average/approaching expectations**  
**Level 3= performance is average/meeting expectations (also aligned with concept of proficiency)**  
**Level 4= performance is well-above average/ exceeding expectations (also aligned with concept of mastery)**

Rating	Highly Effective	Effective	Developing	Ineffective
Average Points	2.7 – 3.0	2.3 – 2.6	1.9 – 2.2	0-1.8

HIGHLY EFFECTIVE			EFFECTIVE			DEVELOPING		INEFFECTIVE												
20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
3.0	2.9	2.7-2.8	2.6	2.5	2.3-2.4	2.1-2.2	1.9-2.0	1.68-1.8	1.54-1.67	1.40-1.53	1.26-1.39	1.12-1.25	.98-1.11	.84-.97	.70-.83	.56-.69	.42-.55	.28-.41	.14-.27	0-.13



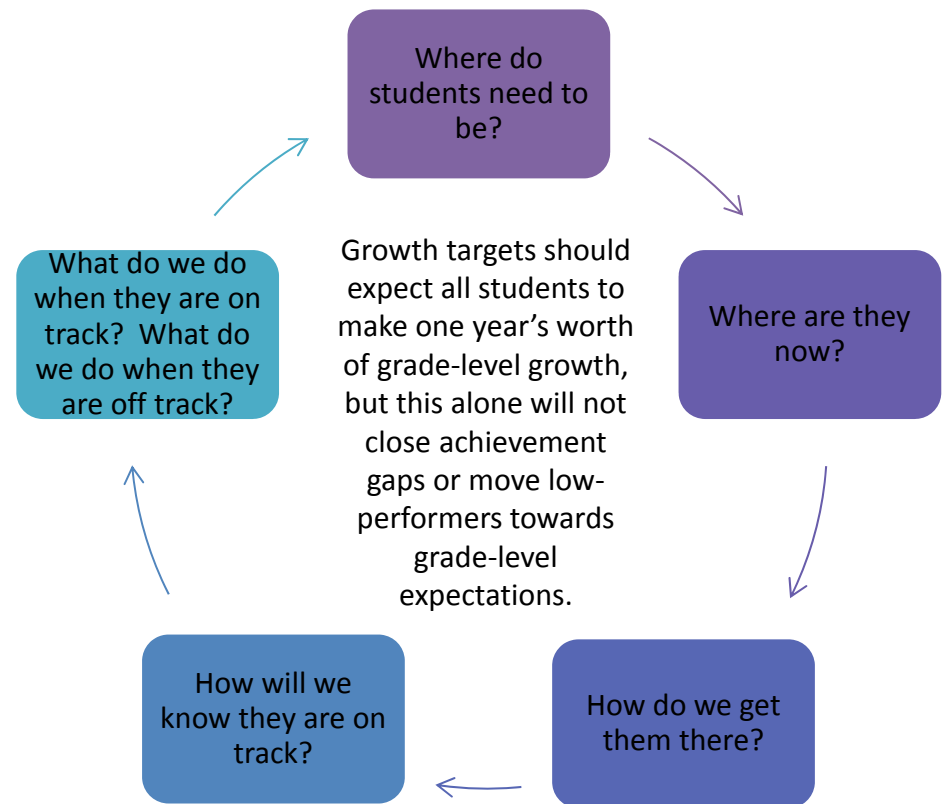
**Improving student learning is at the center of all our work. Setting long-term goals, as seen in Student Learning Objectives, allows educators to be strategic as they plan backwards from a vision of student success.**



When done thoughtfully, the SLO process can lead to higher quality discussions focused on student growth and learning, clearer indications of when and how to adjust instruction to meet student needs, and more strategic planning of professional development offerings.

## Educators must set a minimum growth target of one year's grade-level growth, except for those students with IEPs, wherein the district may need to specify an alternative growth target.

- Educators should look to standards and course curricula to determine what knowledge and skills students are expected to gain over the interval of instruction
- The minimum rigor target for all students should reflect proficiency of the relevant course or grade-level standards, or in other words, “one year's grade-level growth”
- Students begin a course with varying levels of preparedness so educators must determine what a year's worth of grade-level growth will look like for students who enter significantly below or significantly above grade-level expectations



**It is expected that all students should be making at least one year's expected grade-level growth, however, targets may be differentiated based on students' level of preparedness.**

Determine what the most important learning is for the specified course/grade level; decide what students need to know and be able to do in order to be successful in the subsequent course/grade level.

Use multiple sources of baseline data to identify how prepared each student is to meet these expectations.

Some students will enter the course lacking prerequisite knowledge or skills.

Some students will enter the course with the necessary prerequisite knowledge or skills.

Some students will enter the course with prerequisite knowledge or skills that exceed the expectation.

Determine targets that will accelerate student gains and close achievement gaps.

Determine targets that will ensure students master the relevant course content and prepare them for the next level of instruction.

Determine targets that will continuously challenge students to grow and deepen their understanding.

*Students who begin the course significantly below grade level expectations will need to make more than a year's worth of grade level growth, in order to "catch up" to their peers. Targets should be set that encourage accelerated gains and close achievement gaps.*

# Statutory Language

## Statutory Language: Observation Weighting and Scoring Ranges

### Statutory Language

3012-d(4)(b) states: “The commissioner shall determine the weights, and/or weighting options and scoring ranges for the subcomponents of the observation category that result in a combined category rating” and 3012-d(7) further states that the process by which weights and scoring ranges are assigned to subcomponents and categories is transparent...and must ensure that it is possible...to obtain any number of points in the applicable scoring ranges, including zero, in each subcomponent”

## Statutory Language: Observation Rubrics

### Statutory Language

3012-d(4)(b) states: “The observations category for teachers shall be based on a state-approved rubric and shall include up to three subcomponents”

## Statutory Language: Number, Frequency and Duration of Observations and Observation Parameters

### Statutory Language

3012-d(4)(b) states, “The commissioner shall also determine the minimum number of observations to be conducted annually, including frequency and duration, and any parameters therefor.”

## Statutory Language: Student Performance Weights and Scoring Ranges

### Statutory Language

3012-d(4)(a)(2) states: “The commissioner shall determine the weights and scoring ranges for the subcomponent or subcomponents of the student performance category that shall result in a combined category rating” and 3012-d(7) further states that the process by which weights and scoring ranges are assigned to subcomponents and categories is transparent...and must ensure that it is possible...to obtain any number of points in the applicable scoring ranges, including zero, in each subcomponent”



# Statutory Language: Parameters for State-provided Growth Model

## Statutory Language

3012-d(4)(a)(1)(A) states: “for a teacher whose course ends in a state-created or administered test for which there is a state-provided growth model, such teacher shall have a state-provided growth score based on such model...”

3012-d(4)(a)(2) states: “The commissioner shall also set parameters for appropriate targets for student growth for both subcomponents, and the department must affirmatively approve and shall have the authority to disapprove or require modifications of district plans that do not set appropriate growth targets, including after initial approval.”

## Statutory Language: Parameters for Growth Targets (SLOs)

### Statutory Language

3012-d(4)(a)(1)(B) states: “for a teacher whose course does not end in a state-created or administered test such teacher shall have a student learning objective (SLO) consistent with a goal-setting process determined or developed by the commissioner, that results in a student growth score; provided that, for any teacher whose course ends in a state-created or administered assessment for which there is no state-provided growth model, such assessment must be used as the underlying assessment for such SLO.”

3012-d(4)(a)(2) states: “The commissioner shall also set parameters for appropriate targets for student growth for both subcomponents, and the department must affirmatively approve and shall have the authority to disapprove or require modifications of district plans that do not set appropriate growth targets, including after initial approval.”

## Statutory Language: State-Approved Assessments (Optional and Required Component)

### Statutory Language

3012-d(4)(a)(2)(A) and (B) state that the optional student performance category may be either “a second state-provided growth score on a state-created or administered test...”, or “a growth score based on a state-designed supplemental assessment, calculated using a state provided or approved growth model. The optional second subcomponent shall provide options for multiple assessment measures that are aligned to existing classroom and school best practices and take into consideration the recommendations in the testing reduction report...”

3012-d(2)(d) defines state-designed supplemental assessments as “a selection of state tests or assessments developed or designed by the state education department, or that the state education department purchased or acquired from (i) another state; (ii) an institution of higher education; or (iii) a commercial or not-for-profit entity, provided that such entity must be objective and may not have a conflict of interest or appearance of a conflict of interest; such definition may include tests or assessments that have been previously designed or acquired by local districts, but only if the state education department significantly modifies growth targets or scoring bands for such tests or assessments or otherwise adapts the test or assessment to the state education department's requirements.”

# Statutory Language: Overall Rating

## Statutory Language

Section 3012-d(5)(a) states:

The following rules shall apply: a teacher or principal who is

- “(1)rated using two subcomponents in the student performance category and receives a rating of ineffective in such category shall be rated ineffective overall; provided, however, that if the measure used in the second subcomponent is a state-provided growth score on a state-created or administered test pursuant to clause (A) of subparagraph one of paragraph a of subdivision four of this section, a teacher or principal who receives a rating of ineffective in such category shall not be eligible to receive a rating of effective or highly effective overall;
- (2) rated using only the state measure subcomponent in the student performance category and receives a rating of ineffective in such category shall not be eligible to receive a rating of effective or highly effective over- all; and
- (3) rated ineffective in the teacher observations category shall not be eligible to receive a rating of effective or highly effective overall.”

Section 3012-d(5)(b) lists the overall rating combinations based on all available student performance and teacher observation category ratings (i.e., the matrix presented on slide 21)

## Statutory Language: Principal Regulations

### Statutory Language

3012-d(14) states, “The commissioner shall adopt regulations to align the principal evaluation system as set forth in section three thousand twelve-c of this article with the new teacher evaluation system set forth herein.”

## Statutory Language: Continuation of §3012-c

### Statutory Language

Section 3012-d(15) states:

“The provisions of paragraphs d, k, k-1, k-2 and l of subdivision two and subdivisions four, five, five-a, nine, and ten of section three thousand twelve-c of this article, as amended, shall apply to this section to the extent determined by the commissioner.”

## Statutory Language: Waiver Process for Assignment of Students to Ineffective Teachers

### Statutory Language

3012-d(8) states, “A student may not be instructed, for two consecutive school years, by any two teachers in the same district, each of whom received a rating of ineffective under an evaluation conducted pursuant to this section in the school year immediately prior to the school year in which the student is placed in the teacher's classroom; provided, that if a district deems it impracticable to comply with this subdivision, the district shall seek a waiver from the department from such requirement.”

## Statutory Language: Prohibited Elements

### Statutory Language

- 3012-d(6) states, “The following elements shall no longer be eligible to be used in any evaluation subcomponent pursuant to this section:
- a. Evidence of student development and performance derived from lesson plans, other artifacts of teacher practice, and student portfolios, except for student portfolios measured by a state-approved rubric where permitted by the department;
  - b. Use of an instrument for parent or student feedback;
  - c. Use of professional goal-setting as evidence of teacher or principal effectiveness;
  - d. Any district or regionally-developed assessment that has not been approved by the department; and
  - e. Any growth or achievement target that does not meet minimum standards set forth in regulations of the commissioner adopted hereunder.”