New York State
Teacher and Principal Effectiveness
Evaluation Policy Options
March 8, 2011

John King, Senior Deputy Commissioner
Amy McIntosh, Senior Fellow

Regents Timetable

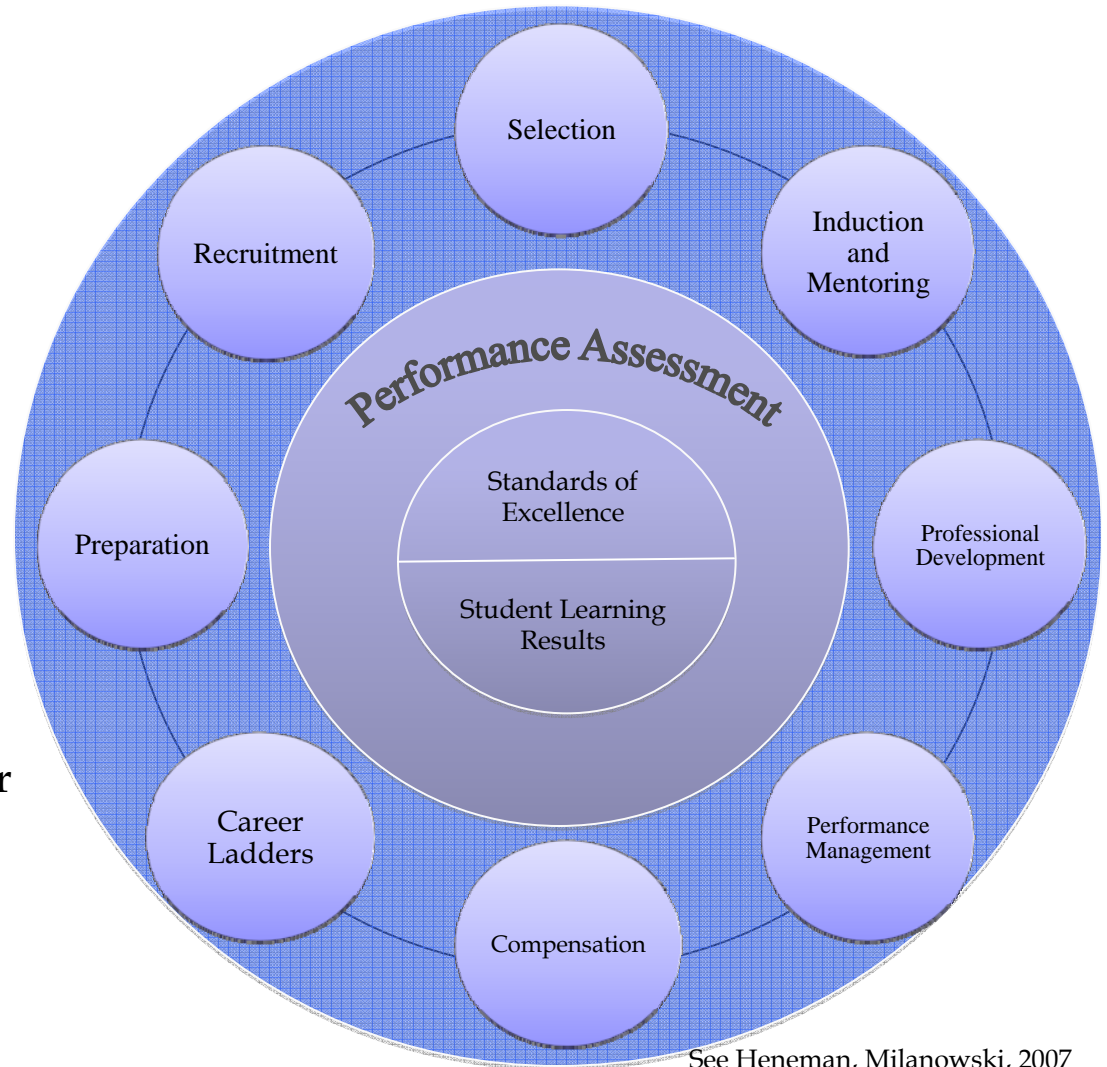|	Board of Regents Agenda
January	Policy Options Discussion: "60%" teacher and principal measures
February	Policy Options Discussion: Local assessments and Non-tested subjects
March	Policy Options Discussion: Review and Discussion • Student Achievement Measures •"Other" Measures of Effectiveness • Determining scores and ratings (Highly Effective, Effective, Developing, Ineffective)
April	Regents Task Force Recommendations
May	Draft Regulations for 2011-12 Implementation
June	"Emergency" Adoption of Regulations

Overview of Evaluation Policy Decisions and Options

- ➢ Ensuring Teacher and Principal Excellence
- ➢ Student Achievement Measures
  - Growth and Value-added Model Basics
  - State growth measures : teachers and principals
  - Local assessment measures: teachers and principals
- ➢ Other Measures of Effectiveness
  - Teacher s
  - Principals
- ➢ Scoring and Ratings
- ➢ Questions/Comments

Ensuring Teacher and Principal Excellence

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

Selection

Induction and Mentoring

Recruitment

Performance Assessment

Professional Development

Preparation

Standards of Excellence

Student Learning Results

Performance Management

Career Ladders

Compensation

Student Achievement Measures

|	Teachers		Principals	
	Grades 4-8 ELA/Math	All Other (2012 +)	Elementary Middle	High School (2012 +)
State Growth				
Local Assessments				

Regent Decisions:
- Approval of state growth/value-added models
- Teacher of Record and student linkage policies
- Which assessments?

Next topics
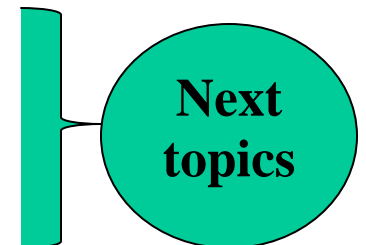
State Student Growth Measures

Most Race to the Top states have plans for growth measures based on state tests.

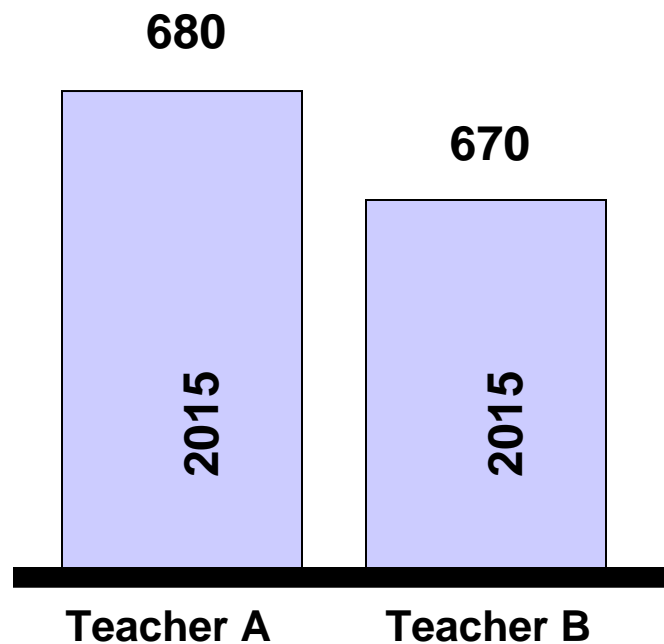Percent (of 100%) for state growth measures in evaluation

20-25%	New York State*
26% - 40%	FL, MD*, TN*
41% - 60%	CO, D.C., GA, HI, LA, RI
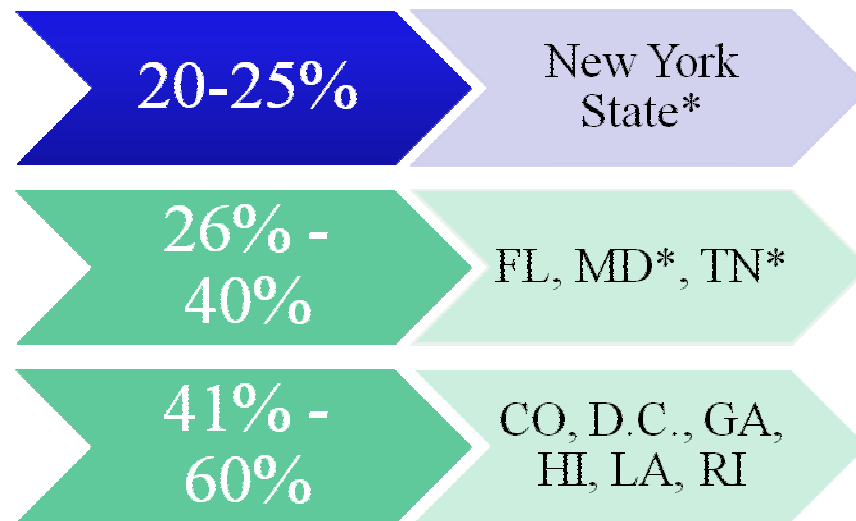
* Incorporating local assessment measures brings NYS to 40% and MD and TN to 50% weight on student achievement.

VA/Growth Basics: Achievement

Avg. Student Achievement (2015)

5th grade math
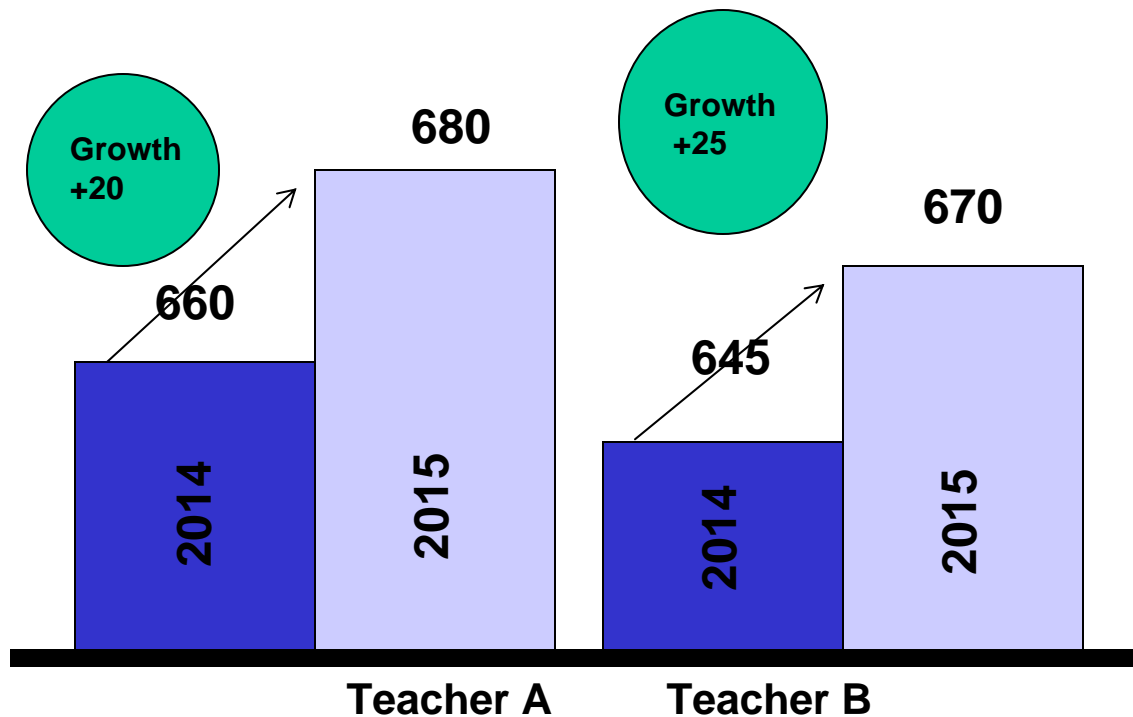
Illustrative Scale Scores

Annual achievement scores say more about students than teachers.

VA/Growth Basics: Growth

Avg Student Growth (2014-2015)

Illustrative Scale Scores

Adding average prior achievement for the same students shows Teacher B's students had higher growth.

VA/Growth Basics

Value-added: Growth in achievement compared to what?

- For educator evaluation: what growth did my students achieve compared to the average of "similar" students?
    - i.e. Expected or predicted growth
- Difference between this "expected" or "predicted" growth and actual is Value-added

VA/Growth Basics: Value-Add

Avg Student Growth vs. Similar Students (2014-2015)

Illustrative Scale Scores

Comparing growth to the average growth of "similar" students gives teacher A the higher "value-added" result.
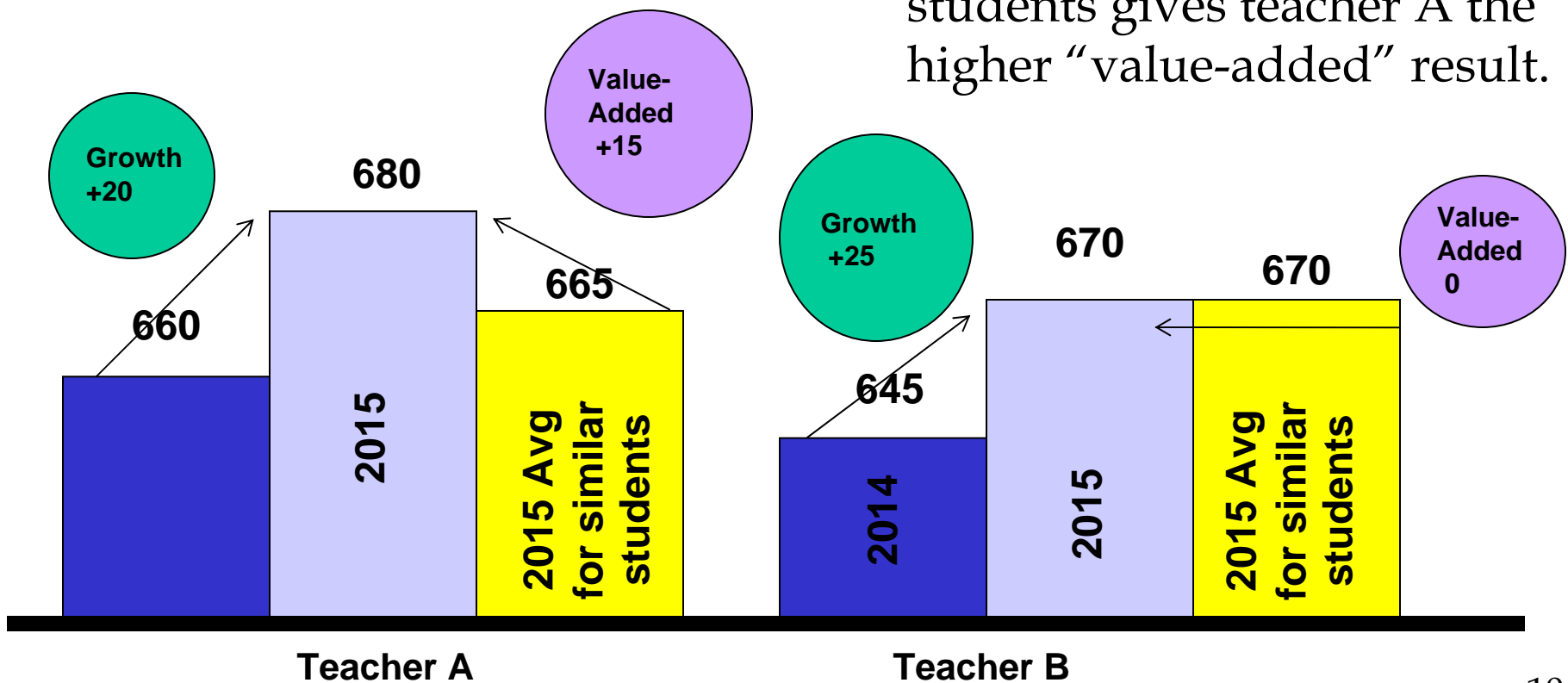
Teacher A

Teacher B

VA/Growth Basics: Calculating "similar student growth"

- Apply statistical analysis to (lots of) data to isolate impact of factors outside of teacher's control
- Result puts teachers on a more level playing field.

Measurable Student Characteristics	Classroom or School Characteristics
Academic history: previous year or multiple years	% in class/school at previous level
Poverty	% in class/school
Special Ed/ELL	% in class/school
Repeat a grade	% in class/school
Attendance history	Class Size/average class size in school
Etc.	Etc

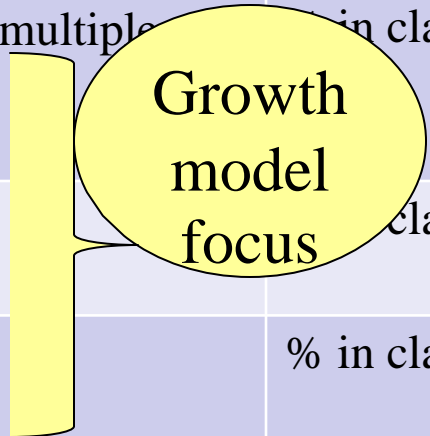# VA/Growth Basics: Calculating "similar student growth"

- Apply statistical analysis to (lots of) data to isolate impact of factors outside of teacher's control
- Result puts teachers on a more level playing field.

Measurable Student Characteristics	Classroom or School Characteristics
Academic history: previous year or multiple years	in class/school at previous level
Poverty	class/school
Special Ed/ELL	% in class/school
Repeat a grade	% in class/school
Attendance history	Class Size/average class size in school
Etc.	Etc

Growth model focus

Basic Findings from VA Research

- Substantial variation in VA across teachers
  - Difference between teachers at 75th and 25th percentiles is ~1/5th of the racial test score gap
  - A bit more variation in math than ELA
  - Much of the variation is <u>within</u> schools
- VA estimates appear to contain real power to predict teacher effectiveness as measured by student achievement
  - Enough stability across years to be useful
  - Year to year reliability ranges from 0.3 to 0.5

Slide from November 2010 Presentation to NY Regents Taskforce by Jonah Rockoff,
Columbia Business School, Hamp Lankford, SUNY Albany, and Jim Wyckoff, UVA

13

Incomplete and Useful Performance Data

Table 1. How the Ten Players with the Highest BAs in 1998 Did in 1997 and 1999

	1997	1998	1999
Larry Walker	.366	.363	.379
John Olerud	.294	.354	.298
Bernie Williams	.328	.339	.342
Mo Vaughn	.315	.337	.281
Eddie Perez	.215	.336	.249
Dante Bichette	.308	.331	.298
Albert Belle	.274	.328	.297
Mike Piazza	.362	.328	.303
Eric Davis	.304	.327	.257
Jason Kendall	.294	.327	.332

- Stability in batting average ~ 0.4
- Stability in earned run avg. (pitchers) ~ 0.35
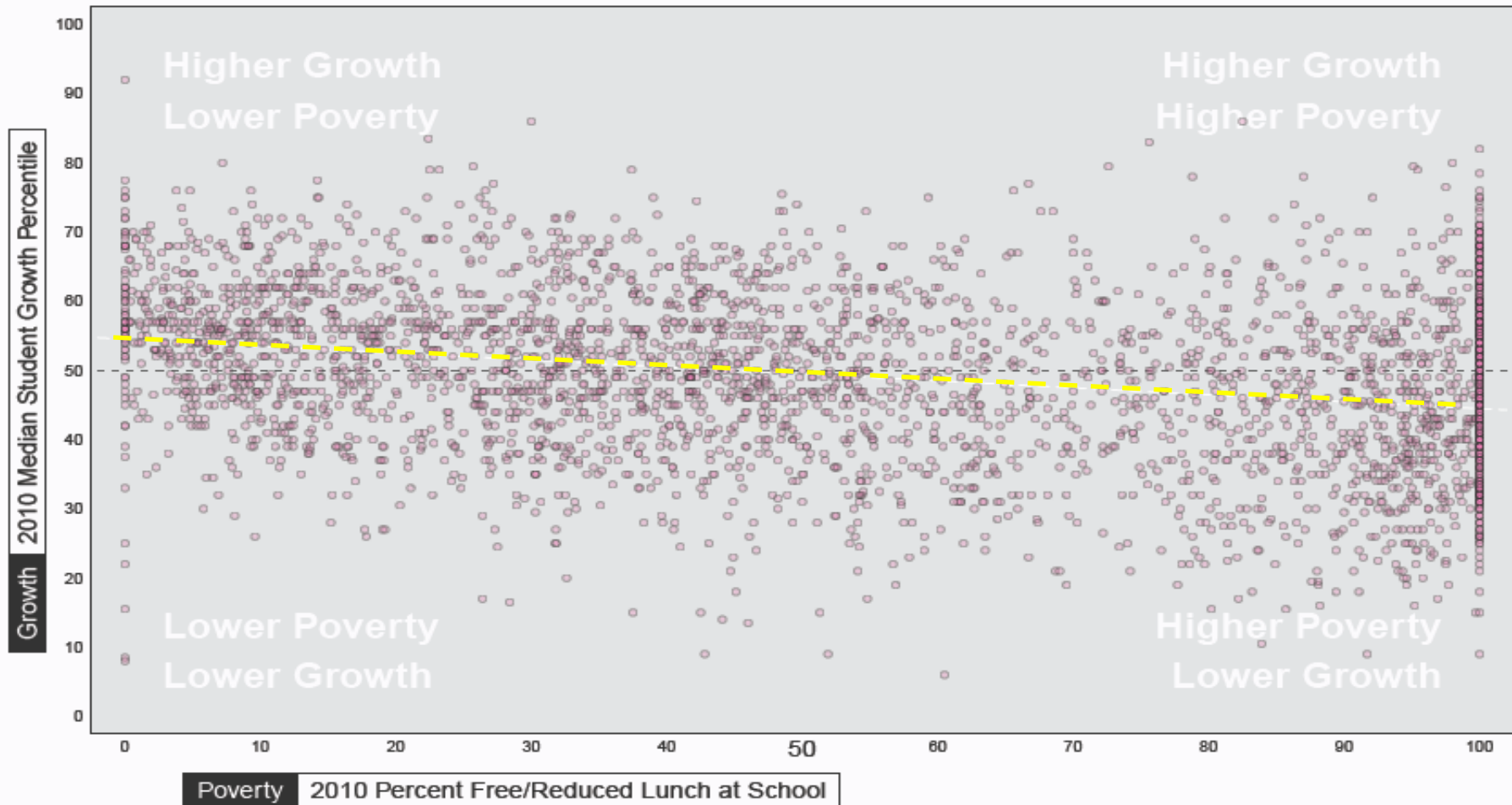
Status of Growth/VA Modeling

- Center for Assessment under contract to explore whether student growth percentiles could be a basis for a revised institutional accountability model.
- Have completed student growth percentiles (SGPs) calculations for all students with 2009-10 test results and at least 1 year prior history. (also 08-09, 07-08)
  - Quality checks and statistical analysis still in process so all results preliminary
- Initial student level results show model works generally as expected:
  - No correlation at student level between prior student achievement and student growth percentile.
  - Model more predictive the more years of student data (i.e. older students) but quite predictive even at 4th.
  - More variation in SGP in math, than ELA
- School level results show other expected trends:
  - Wide range of performance across state and within districts
  - Wide range of median SGP by prior achievement

Incorporating student growth percentile data into principal or teacher evaluations

- At school level, median SGP shows some correlation with other student demographics which may need to be considered in assigning evaluation scores
  - Poverty
  - Students with disability
  - Possibly others (ELL)
- Also exploring ways to correct for false negatives
  - Identify statistical outliers where small changes in test scores of generally very high or low achieving students could lead to poor results on growth percentile comparisons
  - How to identify and use statistical confidence intervals

Growth and School Poverty

New York State School Performance
2010 Mathematics Growth & School Poverty

Slide from NYC Model Summary: February 24th, 2011 (p. 39)

Data elements to collect/consider for VA modeling

Which of the following items should SED provide and instruct its chosen vendor to test and recommend in a growth/Value-added model?

State Data Elements
(list below includes items for which SED believes it has a valid data source now or in the near future)

Student Characteristics	Other Characteristics
• Student State assessment history • Free/reduced price lunch or other poverty indicators • Disability indicators • English language learner indicators • Ethnicity/race • Gender • % daily student attendance (period attendance is not feasible in the near future) • Student suspension data • Retained in grade • Summer school attendance • Student new to school in a non-articulation year •Student age (.i.e. overage for grade?)	• Classroom characteristics o Class size o % of students with each demographic characteristics in a class • School characteristics o % of students with each demographic characteristic o Average class size o Grade configuration (K-5, K-12 etc) • Educator experience level in role

Next Steps: Growth/VA Modeling

- Issue RFP for vendor(s) to develop full principal and teacher value-added models and report results to educators and district/SED leadership.
  - Develop value-added model for ELA/Math grades 4-8 and recommend which student, classroom, school and educator characteristics should be incorporated based on empirical evidence and policy considerations
  - Advise on converting student growth and VA model results to evaluation scores in 2011-12 and beyond
  - Expand beyond 4-8 ELA and Math as models can be developed for existing and new state assessments.
- In advance of vendor selection, work with Task Force researchers and student growth percentile data to prepare illustrative scenarios showing how school-level median student growth percentiles could be incorporated into principal evaluations
- Teacher linkage data for 10-11 school year will provide data set for modeling at teacher level

Teacher of Record and Student Linkage Policy

➤ State must collect data linking teachers, courses and students

  – Where teachers verify data, credibility of result is higher

➤ 2010-11 Policy Decisions:

  – Teacher of Record: Districts must identify one teacher with primary responsibility for instruction

    • Team Teachers: if some districts are able to identify, should they?

  – Best way with current data to determine which students count for each teacher?

➤ As data collection improves, policy options increase

  – Teacher assignment changes during year

  – Student mobility and attendance

  – Role of push-in/pullout teachers

➤ Challenge: operational complexity vs. precision

Student Achievement Measures

	Teachers		Principals	
	Grades 4-8 ELA/Math *2011-12 and beyond*	All Other *2012-13 and beyond*	Elementary Middle *(with grades 4-8 in 2011-12) All thereafter*	High School *2012-13 and beyond*
State Growth	Results of Model	NEXT TOPIC	Results of model	NEXT TOPIC
Local Assessments				

State Growth Beyond Grades 4-8 ELA/Math

- Regents may choose to expand state assessments in core academic subjects:
  - Planned: ELA 9-11 (2011-12)
  - Possibilities subject to funding availability/approval:
    - Science 6-8
    - Social Studies 6-8
    - PARCC ELA/Math 3-11 (2014-15)
- Vendor would then be tasked to develop Growth/Value-added models for approval using existing (Regents, science) and new state assessments.
  - Same-subject growth/VA models would be prioritized (ELA sequence; Math from 8th through Regents)

Student Achievement Measures

	Teachers		Principals	
	Grades 4-8 ELA/Math	All Other	Elementary Middle	High School
State Growth	Results of Model	Results of model in other tested subjects	Results of model in tested grades/subjects	Results of model in other tested grades/subjects
Local Assessments	**NEXT TOPIC**			

Next topic: Which assessments?
Teacher: local: 4-8 ELA/Math and All Other

Local Assessment Criteria

- ➤ May be an achievement or a growth measure
- ➤ Rigorous
  - – Aligned with NY State Common Core Standards
  - – Other industry standards for rigor
- ➤ Comparable across classrooms
  - – Does this mean 1 assessment district-wide per grade and subject?
  - – If district wants more than 1, should there be a process for approval?
- ➤ Districts may have other considerations

  Effect on Instruction

  Considerations of cost, timing

Local Assessment Options: Should all/any of these be allowable for Districts to choose from?

1. Select from pre-approved list of commercially available assessments by grade/subject group
2. District develops own assessments with vendors and/or teacher teams
3. Group or team measure from state or local assessments
4. Standard, structured district student growth goal-setting process between principals and teachers, using any state, district or school assessment
5. Utilize state tests for local measures. (achievement instead of growth, or full 40 points, or other)
6. Other?

Local Assessments Used Today

Districts select an assessment to use across all classrooms in a specific grade / subject; these are just examples:

- NWEA/MAP
- Scantron Performance Series, Achievement Series
- CTB Acuity
- ACT Explore, PLAN
- AP or IB
- PCTS
- ITBS
- InView
- Meta Metrics

- GMADE
- DIBELS
- DRA
- OWLS
- Scholastic SPI / SRI
- iStation
- College Board ReadiStep, PSAT, and SAT

Student Achievement Measures

	Teachers		Principals	
	Grades 4-8 ELA/Math	All Other	Elementary Middle	High School
State Growth	Results of Model	Results of model as approved to other tested subjects	Results of model in tested grades/subjects	Results of model in ELA/Math
Local Assessments	List of allowable options for District choice			

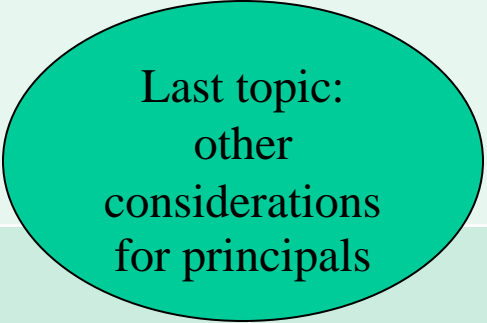NEXT TOPIC: **Remaining Non-tested subjects**

State test coverage

As NY State adds state tests, and even after addressing highest priority needs, there will be teachers without a state measure of student growth.

Non-Tested Growth Metric: Options in lieu of State assessments

1. State requires school-wide or grade/subject-wide team measure of growth. For example:
   - Art teachers earn points for school-wide ELA growth
2. State requires districts to choose a growth measure from the local assessment options and assign the related 20 points, as well as choosing their approach to the "local assessment" 20 points.
3. The State pre-approves specific assessment options as suitable for growth measures in each grade/subject category and requires districts to use one of these for each grade/subject.
   - Also provides approach to assigning the "state growth" evaluation points in each case.
4. State distributes the growth points in some proportion to the other two evaluation categories for these teachers.
5. Other

Student Achievement Measures

	Teachers		Principals	
	Grades 4-8 ELA/Math	All Other	Elementary Middle	High School
State Growth	Results of Model	Results of model as approved to other tested subjects Regent decision on remaining grades/subjects	Results of model in tested grades/subjects	Results of model in ELA/Math
Local Assessments	List of allowable options for District choice		Last topic: other considerations for principals	

Principal Student Achievement Measures: Other State and District Examples

Principal Student Achievement Measures			
(Examples from States and Districts)			
State Growth/VA Measures	**% Proficient**	**Focus on Sub-Groups**	**High-School Measures**
- Many states and districts aggregate state growth/VA measures for principals across tested grades in K-8	• GA • HI • MA • MD • OH • RI • TN • DC • Hillsborough County • NYC	- Growth of low achievers and/or special needs: • Hillsborough County • NYC - Growth of high achievers: • Ohio (AP take rates)	- Value-added for end of course tests with predictors: • GA • RI • Hillsborough County • NYC - Grad rates and credit accumulation: • NYC

Principal Student Achievement: Other options

PRINCIPALS	Elementary/Middle	High School
State Growth	• Results of model in tested grades/subjects •Growth of subgroups: e.g. SWD, ELLs, high or low achievers; Required for all or where achievement gaps exist	• Results of model in ELA/Math • Progress toward graduation: credit accumulation or other predictive measure •Growth of subgroups: e.g. SWD, ELLs, high or low achievers; Required for all or where achievement gaps exist
Local Assessment: should all below be allowable? Others?		
	• Achievement on state tests (% proficient) • Performance of student subgroups (SWD, ELL, highest or lowest achievers) • Selected metrics from local assessments used for teachers	Additional HS options: •Graduation Rates (or college-ready grad rates) • Regent participation rates • % Regent pass rate (or pass with college-ready score) • Drop out rates • PSAT/SAT take rates and scores • AP, IB other Regent equivalent take and pass rates

32

Student Achievement Measures: Implementation Steps

SED plans to proceed with:

- RFQ for local assessment options and subset of those for "in lieu of" state growth measures
- RFP for VA model and reporting provider(s)

Overview of Evaluation Policy Decisions and Options

➢ Ensuring Teacher and Principal Excellence

➢ Student Achievement Measures

- Growth and Va Model Basics
- State growth measures : teachers and principals
- Local assessment measures: teachers and principals

➢ Other Measures of Effectiveness

- Teachers
- Principals

➢ Scoring and Ratings

➢ Questions/Comments

Teacher 60%: Differing Examples from States and Districts

	Florida	Georgia	DC	NYSUT Innovation Districts (Pilot 2010-11)	NYC (Pilot 2010-11)
Rubrics	Developing own	Developing own based on CLASS	Developed own framework	Danielson adapted to NYS standards	Danielson
Multiple Measures	• Administrator Observations • Roving Full-Time Expert Observers	• Administrator Observations • Roving Full-Time Expert Observers • Student/parent Surveys • Teacher Portfolio/Evidence Binders	• Administrator Observations • Roving Full-Time Expert Observers • Student/parent Surveys • Teacher Portfolio/Evidence Binders	• Administrator Observations • Teacher Portfolio/Evidence Binders	• Administrator Observations (2 unannounced) • Roving Full-Time Expert Observers (2 unannounced focused on learning environment and instruction)
Notes	• Comprehensive range that includes highly effective and effective categories	• Teachers select individual growth plans based on self-reflection and self-evaluation	• 4 Levels	• Peer Assistance	• 4 Rating Categories • District-Wide Peer Assistance

Principal 60%: Differing Examples from States and Districts

	Ohio	Massachusetts	DC	Rochester	NYC
Rubrics	Ohio Standards for Principals	Will develop frameworks and rubrics	Developed own framework	Developed own leadership standards rubric	Quality Review
Multiple Measures	• 360 Degree Survey • Choice of McRel, Val-Ed, or "other education impact tools" • Goals (2-3) set with supervisor • Student Attendance • Graduation rates • Suspensions and expulsions • % of students in AP classes	• Effectiveness measures of leadership skills (exemplars and models will be provided by MA) • Teacher effectiveness • Self-assessment of professional skills	• Superintendent on leadership standards (30%) • School specific goals (10%) • Family and community engagement (5%) • Retention of high performing teachers (5%) • Special education compliance (10%)	• Goals sett with superintendent • Principal self-assessment against goals and leadership standards • Superintendent assessment on leadership standards • Evidence of principal's work	• Quality Review • Goals and objectives (individually set between principal and supervisor) • Compliance • Attention to SPED/ELL populations • Surveys: students, parents, teachers • Student attendance
Notes	• 5 Rating Categories • Districts determine weights of components and how to arrive at final rating	• At least 3 Rating Categories	• 4 Rating Categories	• 4 Rating Categories • Includes different process for tenured versus non-tenured	• 5 Rating Categories (uses 0 -4 scale)

Other Effectiveness Measures: Regents Decisions

Key Decisions:	Teachers	Principals
What Standards?		
What rubrics to assess performance vs. standards?		
What options may districts consider?		
Are there requirements or restrictions on District options?		

Other Effectiveness Measures: Regents Decisions

Key Decisions:	Teachers	Principals
What Standards?	NY State Teaching Standards	• ISLLC • Develop own NYS standards
What rubrics to assess performance vs. standards?		
What options may districts consider?		
Are there requirements or restrictions on District options?		

Other Effectiveness Measures: Regents Decisions

Key Decisions:	Teachers	Principals
What Standards?	NY State Teaching Standards	•ISLLC •Develop own NYS standards
What rubrics to assess performance vs. standards?	• 1 state-wide rubric • Menu of pre-approved rubrics from state RFQ with district variance option • Free district choice	
What options may districts consider?		
Are there requirements or restrictions on District options?		

Other Effectiveness Measures: Regents Decisions

Key Decisions:	Teachers	Principals
What options may districts consider?	**Should any be removed/added?** •Classroom Observation: •Observation by principal or other administrator •Observation by trained evaluators independent of the school •Observation by trained in-school peer teachers •Observation using video of classroom practice and any approved evaluator •Structured review of student work and/or teacher artifacts using "portfolio" or "evidence binder" processes •Feedback from students, parents, and/or other teachers using structured survey tools •Teacher attendance •Individual professional growth goals •Teacher self-reflection •Other?	**Should any be removed/added?** •Superintendent assessment of principal leadership practice •School visits by other trained evaluators independent of the school •Teacher effectiveness: - Principal actions to implement and conduct teacher evaluation effectively (for example, timely observation and feedback) - Evidence of improved effectiveness of teaching staff (for example, improved retention of higher performers, student growth scores of teachers granted vs. denied tenure) •Operational (for example, budget performance, compliance with regulations, safety, management) •Feedback from students, parents, and/or other teachers •Teacher and/or student attendance •School academic or learning environment goals •Individual professional growth goals •Principal self-reflection •Other?

Other Effectiveness Measures: Regents Decisions

Key Decisions:	Teachers	Principals
Should there be requirements or restrictions on District options? Examples:	• Prioritize classroom practice or principals' instructional leadership standards by requiring a majority of points on these criteria? • Capping points for standards that have less rigorous measurement tools? (professional growth, contribution to community) • Requiring criteria like educator attendance?	
	• Require multiple classroom observations for all teachers? • Require at least one source of evidence besides classroom observation?	• Require metrics of impact on teacher effectiveness • Require input from other sources besides superintendent, e.g. teachers, families, students?

Other Effectiveness Measures: Implementation Issues

- Timing for proceeding with related RFQs for rubrics.

- Requirements districts must address in their district performance review plans. What is missing?
  - Criteria and assessment approaches for teachers and principals
  - How the rating categories (HEDI) will be used to differentiate professional development, compensation, and career decisions for teachers and principals
  - How rating points will be determined for local assessment and "other metrics" subcomponents of evaluation. Process must be transparent to all educators and must be designed in a way that it is possible to earn any amount of rating points within the subcomponent
  - How districts/BOCES will ensure all evaluators are trained and "certified" to conduct evaluations?
  - How district or BOCES will ensure that evaluators will have the time required to complete requirements of the evaluation system. For example, ensuring a reasonable ratio of teachers to trained evaluator to accomplish required observations and conferences?
  - How the district/BOCES will provide timely and constructive feedback to teachers and principals
  - How the district or BOCES will address the performance of teachers or principals whose performance is evaluated as needing an individual improvement plan
  - Other?

Overview of Evaluation Policy Decisions and Options

- ➢ Ensuring Teacher and Principal Excellence
- ➢ Student Achievement Measures
  - Growth and Value-added Model Basics
  - State growth measures : teachers and principals
  - Local assessment measures: teachers and principals
- ➢ Other Measures of Effectiveness
  - Teachers
  - Principals
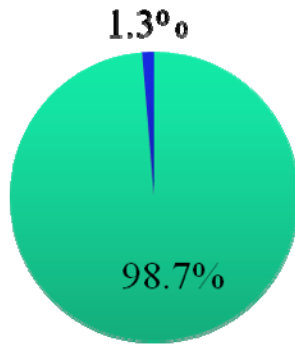- ➢ Scoring and Ratings
- ➢ Questions/Comments
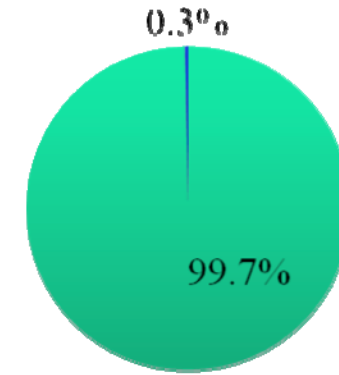
Scoring and Rating

Decisions:

- How will 20 points (25 with a Value-added model) be assigned to educators based on the "state growth" portion of evaluations?

- How will regulations address district assignment of points for "local" and "other" subcomponents?

- What scoring ranges from 1-100 points will be required for each of the summative ratings:
  - Highly Effective
  - Effective
  - Developing
  - Ineffective

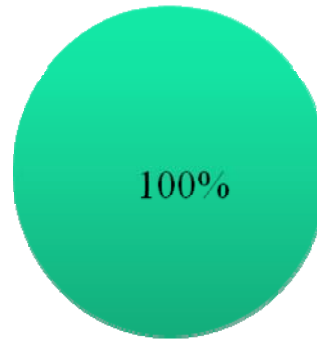# Research Shows Many Districts Do Not Differentiate Teachers
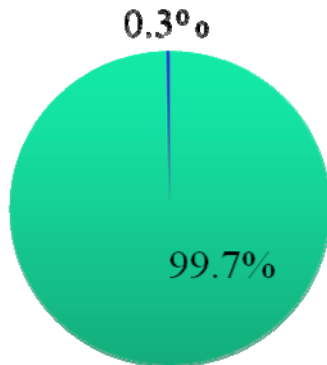
Denver Public Schools SY 05-06 to 07-08

Jonesboro Public SY 03-04 to 07-08

1.3%

0.3%

98.7%

99.7%

Springdale Public Schools SY 05-06 to 07-08

Pueblo City Schools SY 05-06 to 07-08

100%

Toledo Public Schools SY 03-04 to 07-08

0.3%

0.3%

99.7%

99.7%
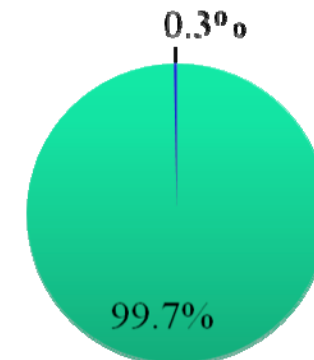
■ Satisfactory (or equivalent) ■ Unsatisfactory (or equivalent)

Weisberg, D., Sexton, S., Mulhern, J. & Keeling, D. (2009) *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness*. New York: The New Teacher Project.

44

Scoring and Rating Key Decisions

- 3 Subcomponent Ratings:
  - How will 20 points (25 with a Value-added model) be assigned to educators based on the "state growth" portion of evaluations?
  - How will regulations address district assignment of points for "local" and "other" subcomponents?

- Overall Composite: What scoring ranges from 1-100 points will be required for each of the summative ratings:
  - Highly Effective
  - Effective
  - Developing
  - Ineffective

State Growth Subcomponent:

- Points assigned according to Commissioner formula using approved assessments and models
- Normal distribution of educators across state
  - Alternative?

- Highly illustrative example:

Performance	Educator Percentile in State (illustrative)	Point Range (of 20) (illustrative)
Well below average	Bottom 15%	0 – 4
Below Average	16 – 35%	5 – 9
Average	36 – 70%	10 – 15
Above Average	71%+	16 – 20

Local and Other subcomponents

What degree of regulatory guidance to districts around assigning these points? (examples)

- None
- Minimal:
  - publicly transparent process for assigning points.
  - no all or nothing scoring i.e. use full range of points available
  - Not limited to "considerations" like student and school factors outside educator control; performance vs. district goals
- Require 4 bands within subcomponents (Highly effective, E, D, I)
  - Local HEDI bands linked to specific growth and/or proficiency results on local assessment results
  - Other 60 points linked to rating levels on rubrics
- Require normal distribution of local assessment points across district except where too few teachers or mostly high achieving students

Overall Composite Score and Ratings:

If districts generally control scoring of 80 points, Regents will set the state growth and the overall composite score bands only. One example (assuming previous state growth

State growth performance	Educator Percentile in State (illustrative)	Point Range (of 20) (illustrative)
Well below average	Bottom 15%	0 – 4
Below Average	16 – 35%	5 – 9
Average	36 – 70%	10 – 15
Above Average	71%+	16 – 20

Overall	HEDI Rating
0 – 64	Ineffective
65 – 79	Developing
80 – 90	Effective
91+	Highly Effective

Overall Composite Score and Ratings:

Alternative could incorporate performance level descriptions and scoring ranges for each subcomponent (state, local, 60) and composite HEDI bands.

- STATE GROWTH 20: normal distribution for similar students. (same as above)

- LOCAL ASSESSMENT 20: normal distribution across district modeled after state growth

- Other 60: Absolute standards anchored in rubrics and tools that have 4 scoring levels.

• Illustratively for 2011-12:

Rating Level	Other 60 Points	Local Assessments	State Growth	Overall	HEDI Rating
Lowest	0-30	0-4	0-4	0-50	Ineffective
	30-45	5-9	5-9	51-75	Developing
	45-54	10-15	10-15	76-90	Effective
Highest	55-60	16-20	16-20	91+	Highly effective

Related Implementation Issues:

- Will SED collect subcomponent ratings for all 3 subcomponents in addition to overall composite score?
- Timing of data availability for state measures vs. current evaluation cycles.
- How to monitor for intended results
- Should SED be able to change composite score bands every year?

Appendix

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.

(From: Educational Leadership Policy Standards: ISLLC 2008) 4

Appendix

- 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

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

Highlighted Bibliography

➢ Braun, H. (2005). Using student progress to evaluate teachers: A primer to value-added models. Princeton, NJ: Educational Testing Service.

➢ Evaluating Teachers: The Important Role of Value-Added. November 17, 2010: The Brookings Brown Center Task Group on Teacher Quality. Available: www.brookings.edu/reports/2010/1117_evaluating_teachers.aspx

➢ Gordon, R., Kane, T., Staiger, D. (2006). Identifying effective teachers using performance on the job. *Brookings Institution Hamilton Paper.*

➢ Measuring Student Growth for Teachers in Non-Tested Grades and Subjects: A Primer. Document created for the Race to the Top Technical Assistance Network held by ICF International (2010).

➢McKinsey & Company. (2010). *Glossary of teacher effectiveness and evaluation terms.* New York: Author.

Appendix

Highlighted Bibliography

➢NYC Teacher Data Initiative: Technical Report in the NYC Value-Added Model 2010. Report by the Value-Added Research Center. Available: schools.nyc.gov/Teachers/Teacher Development/TeacherDataToolkit/default.htm

➢ Rockoff, J., Lankford, H., Wyckoff, J. (November 2010). Presentation to NY Regents Taskforce. Value-added Analysis and Measurement of Effective Teaching.

➢Laura Goe, Courtney Bell, Olivia Little (2008). Approaches to Evaluating Teacher Effectiveness: A Research Synthesis. Available:

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

Appendix

Highlighted Bibliography

➢ Teacher Evaluation Tools: Available:
www.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

➢ Slotnik, W.J., & Smith, M.D. (2004). Catalyst for change: Pay for Performance in Denver (Final Report). Available: cecr.ed.gov/pdfs/guide/other69Percent.pdf

➢ Melissa Brown-Sims (2010). Evaluating School Principals: Tips & Tools. Available:
www.tqsource.org/publications/KeyIssue_PrincipalAssessments.pdf