engage ny
Our Students. Their Moment.

New York State Growth Model for Educator Evaluation 2011-12

June 18, 2012

Draft Discussion
Board of Regents Meeting

AIR®

Today's Agenda

- Context
- New York State Growth Model Description
- Using Growth Measures for Educator Evaluation
- What Data Will Be Available and When?

Our Challenge
Graduating *All* Students College & Career Ready

New York's 4-year high school graduation rate is 74% for All Students
However, the gaps are disturbing.

June 2011 Graduation Rate

Graduation under Current Requirements

	% Graduating
All Students	74.0
American Indian	59.6
Asian/Pacific Islander	82.4
Black	58.4
Hispanic	58.0
White	85.1
English Language Learners	38.2
Students with Disabilities	44.6

Calculated College and Career Ready*

	% Graduating
All Students	34.7
American Indian	16.8
Asian/Pacific Islander	55.9
Black	11.5
Hispanic	14.5
White	48.1
English Language Learners	6.5
Students with Disabilities	4.4

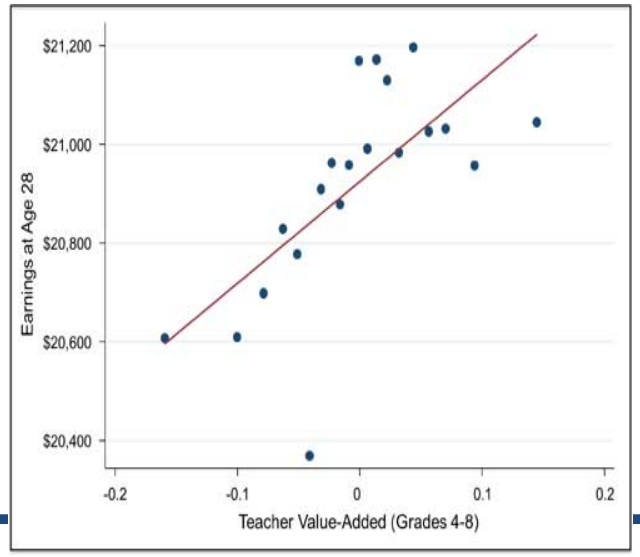
*Students graduating with at least a score of 75 on Regents English and 80 on a Math Regents, which correlates with success in first-year college courses.

Source: NYSED Office of Information and Reporting Services

Teacher Effectiveness Matters

Three new studies show that having an effective teacher in the classroom makes a difference in student outcomes in the classroom and beyond

- The Long-Term Impacts of Teachers: Teacher Value-added and Student Outcomes in Adulthood (Chetty, Friedman & Rockoff). http://obs.rc.fas.harvard.edu/chetty/value_added.html
- The MET Project: Multiple Measures of teaching
- Learning Denied: The Case for Equitable Access to Effective Teaching in California's Largest District. (The Education Trust-West)



Measure	Predictive power	Reliability	Potential for Diagnostic Insight
Value-added	H	M	L
Student survey	M	H	M
Observation	L	M/H	H

Evaluating Educator Effectiveness

Growth 20%	• Student growth on state assessments (state-provided) • Student learning objectives
Locally Selected Measures 20%	• Student growth or achievement • Options selected through collective bargaining
Other Measures 60%	• Rubrics • Sources of evidence: observations, visits, surveys, etc.

American Institutes for Research (AIR)—Our Growth & Value-Added Vendor

Large national non-profit research and service provider in educational assessment and policy areas, as well as health and other areas

- AIR has contributed over many years to the research and early implementation work of growth and value-added models
- Current or previous customers: Florida, Baltimore, MD; Wisconsin; Washington; Oregon; Washington, D.C.; and USDE

Technical Advisory Committee—Growth & Value-Added Technical Experts

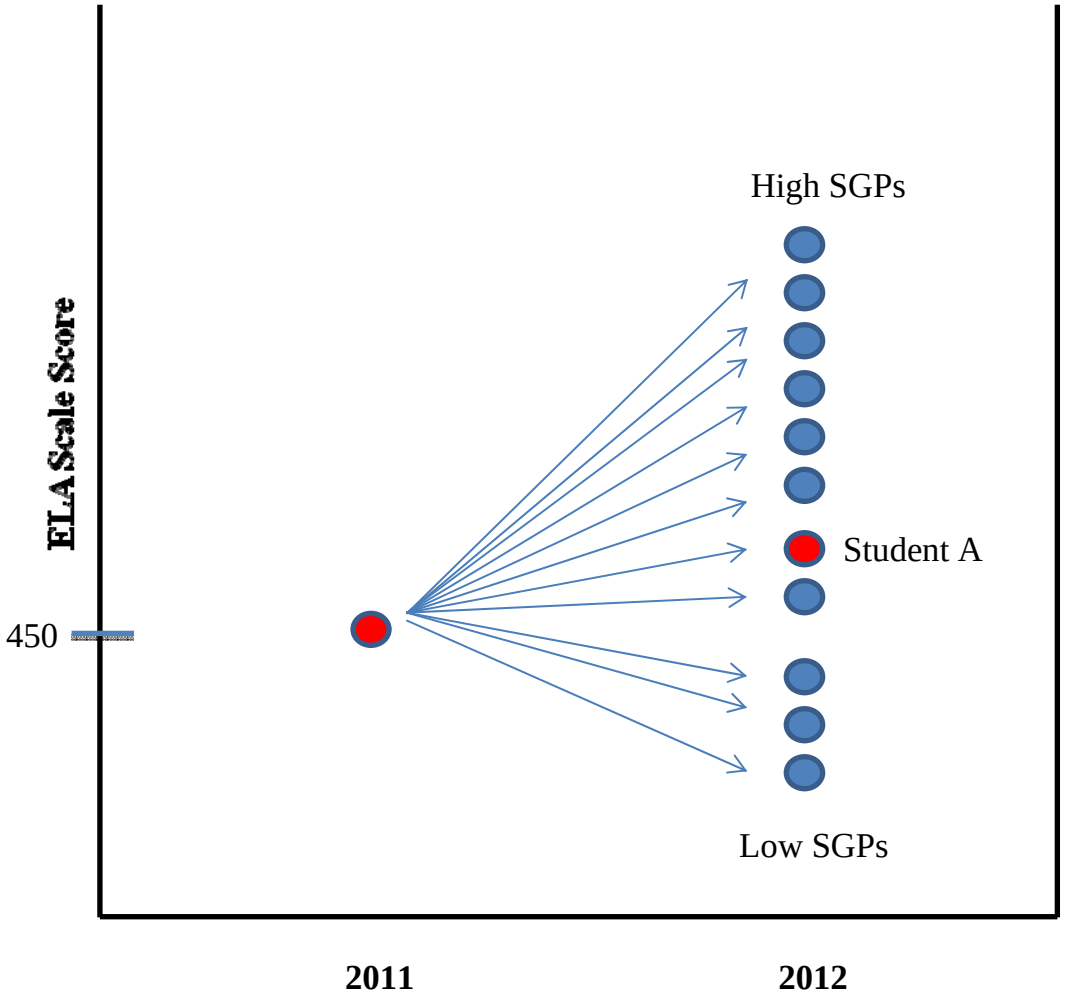
- **Dr. Dan Goldhaber,** University of Washington
  - Professor and Director of the Center for Education Data & Research
- **Dr. Hamilton Lankford,** SUNY Albany
  - Professor and Research Fellow with the American Statistical Association
- **Dr. Dan McCaffrey,** RAND Corporation
  - PNC Chair in Policy Analysis
- **Dr. Jonah Rockoff,** Columbia Graduate School of Business
  - Associate Professor and Faculty Research Fellow at the National Bureau of Economic Research (NBER)
- **Dr. Timothy Sass,** Georgia State University, Distinguished Professor
- **Dr. Douglas Staiger,** Dartmouth College
  - Professor & Research Associate at NBER
- **Dr. Martin West,** Harvard Graduate School of Education & Research
  - Assistant Professor and affiliate at Harvard's Kennedy School and of the CESifo Research Network
- **Dr. James Wyckoff,** University of Virginia
  - Professor and Director of the Center on Educational Policy and Workforce Competitiveness at University of Virginia

Key Points about NYS Growth Measures

- We are measuring student growth and not achievement
  - Allow teachers to achieve high ratings regardless of incoming levels of achievement of their students
- We are measuring growth compared to similar students
  - Similar students: Up to three years of the same prior achievement, three student-level characteristics (economic disadvantage, SWD, and ELL status)

Every educator has a fair chance to do well on these measures regardless of the composition of his/her class or school.

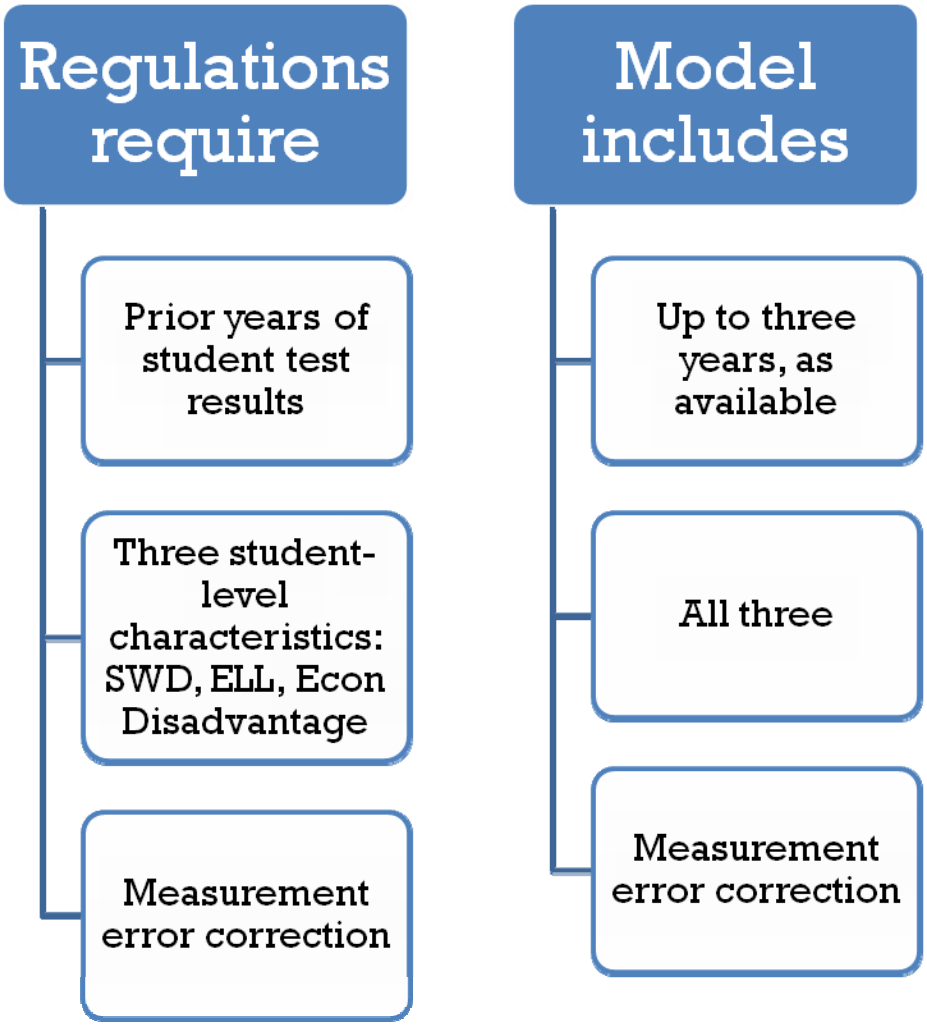
Student A's Current Year Performance Compared to "Similar" Students



Student A's current score is compared to other students who had the same prior score (450). The result can be described as a **"student growth percentile" (SGP).** Student A's SGP is the result of a statistical model and in this example is 45, meaning she performed better in the current year than ***45 percent of similar students.***

State Growth Model Summary

Growth model for 2011-12 only for grades 4-8 ELA/Math for teachers and principals



From Student Growth to Teachers and Principals

Ms. Smith's Class	
	SGP
Student A	45
Student B	40
Student C	70
Student D	60
Student E	40

To measure teacher performance, we find the mean growth percentile (MGP) for her students. To find an educator's mean growth percentile, take the average of SGPs in the classroom. In this case:

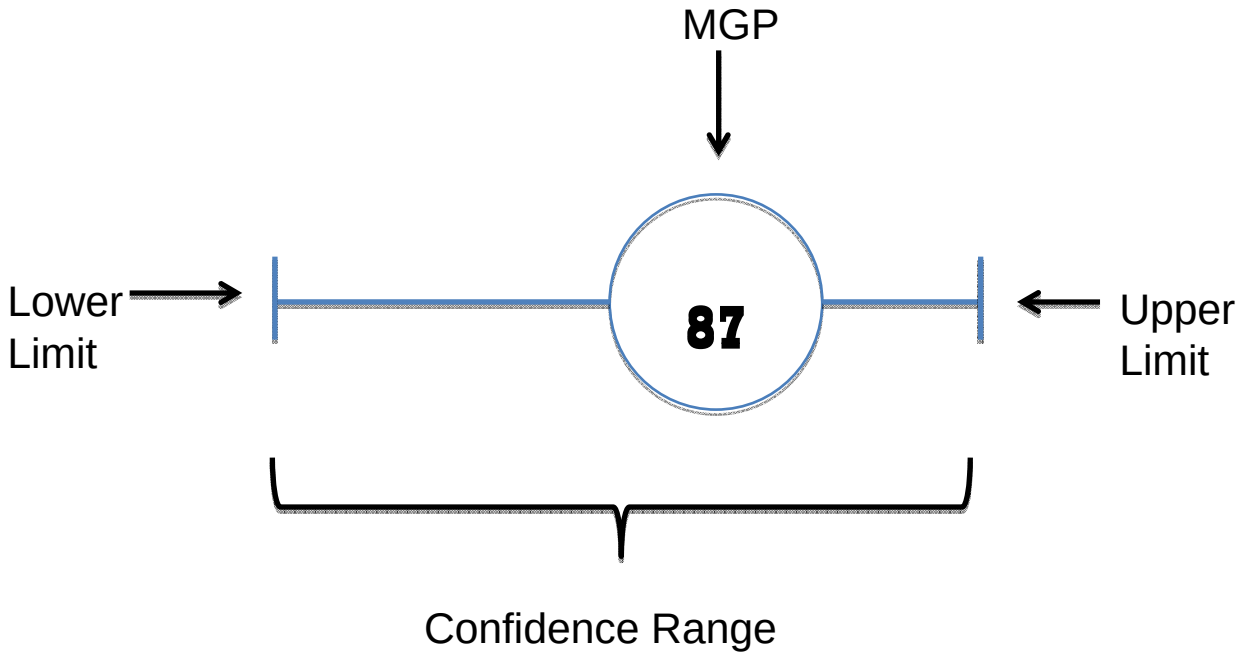
Step 1: 45+40+70+60+40=255

Step 2. 255/5=51

Ms. Smith's mean growth percentile (MGP) is 51, meaning on average her students performed better than 51 percent of similar students.

A principal's performance is measured by finding the mean growth percentile for all students in the school.

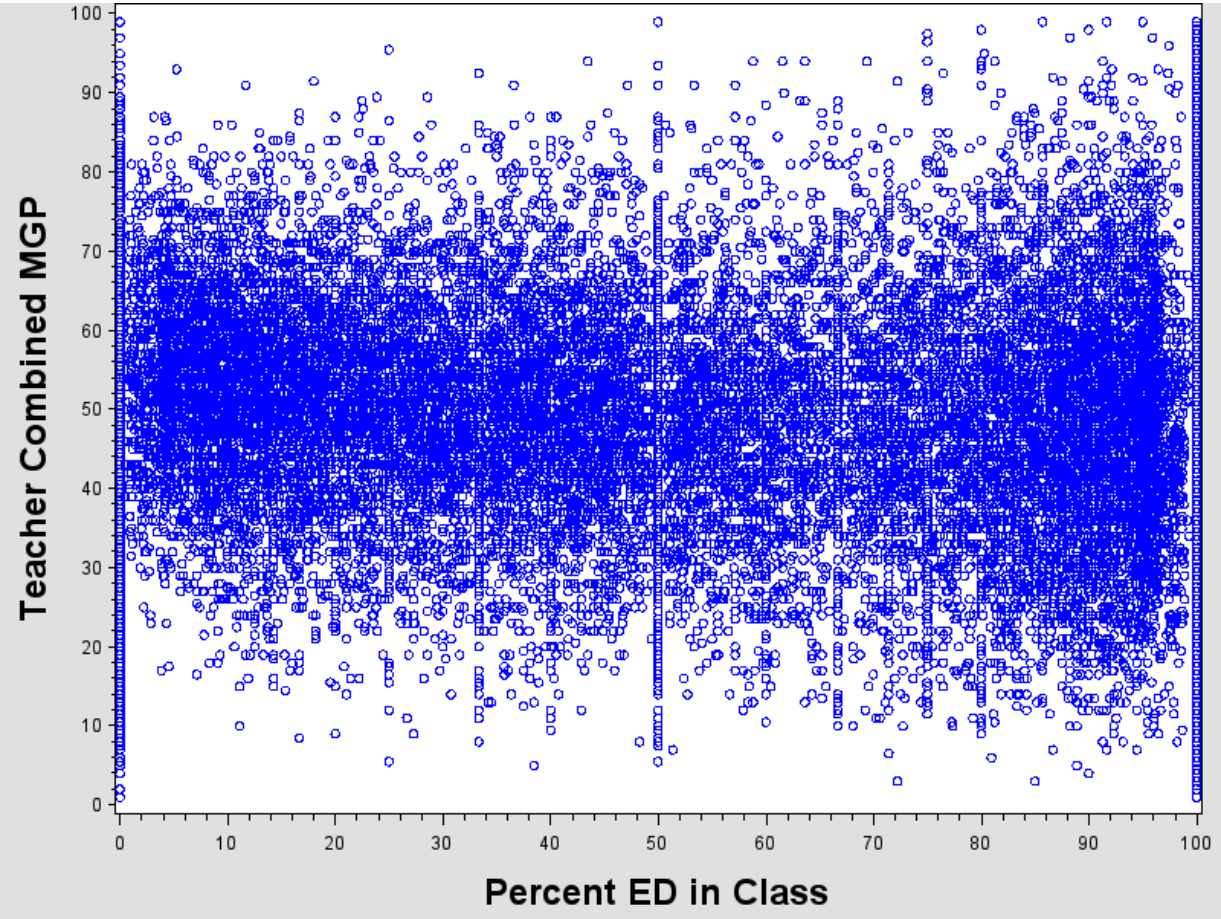
MGPs and Statistical Confidence



The most likely MGP for this educator is 87. However NYSED will provide the upper and lower limits of a 95 percent confidence range, meaning we can be 95 percent confident that an educator's "true" MGP lies within that range.

An educator's confidence range depends on a number of factors, including: number of student scores included in their MGP and the variability of student performance in their classroom.

Initial Results: Teacher MGPs accounting for Economic Disadvantage



Taking student-level characteristics into account helps ensure educators with many students with those characteristics have a fair chance to achieve high or low MGPs.

NOTE: Results using available sample 2010-2011 data.

Growth Model TAC Discussion Summary

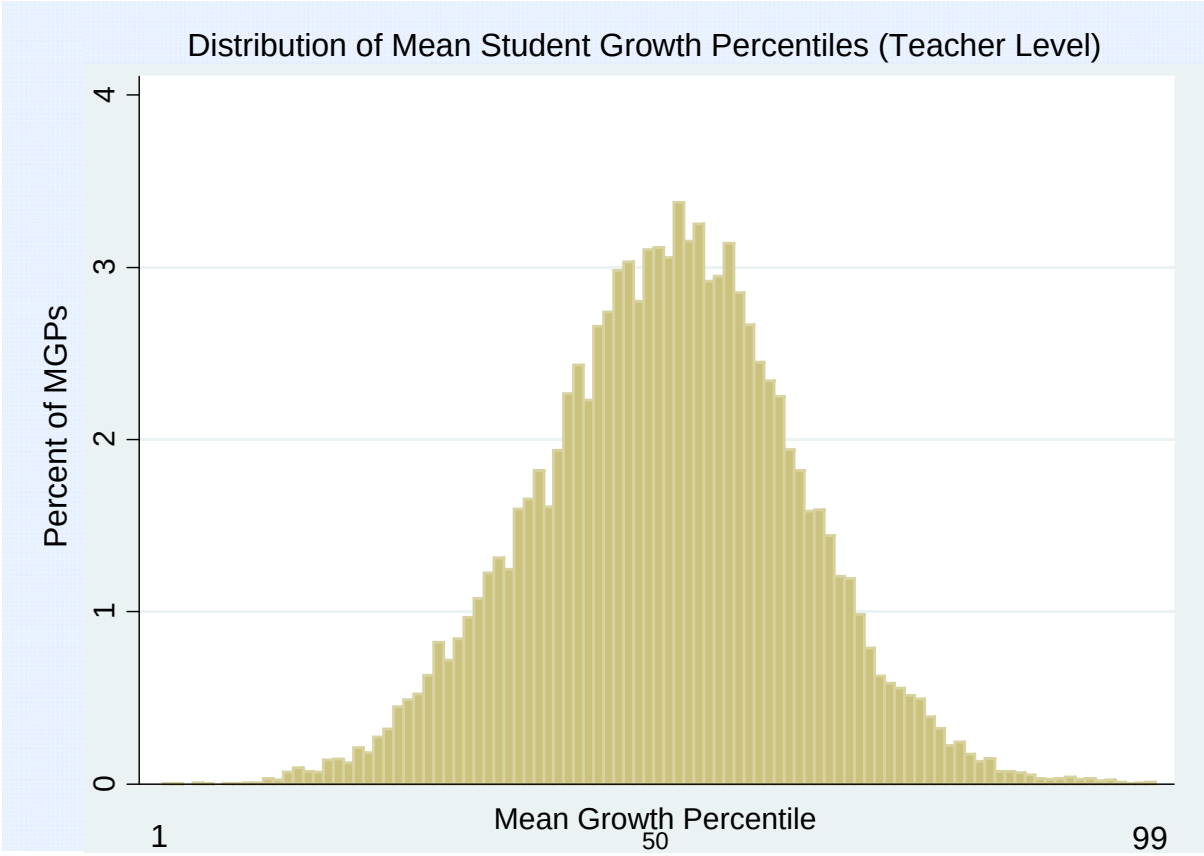
- The TAC concluded that the model was well constructed and suitable for implementation going forward.
- TAC members offered recommendations for improvement and AIR made changes to the analysis based on these recommendations.
- Recommendations:
  - Use the mean SGP rather than the median SGP when aggregating;
  - Modify the method used to calculate standard error in computing mean growth percentiles (MGPs) to account for the clustering of students in different classrooms for a teacher;
  - Modify how to best account for student scores at the tails of the test distribution to avoid statistical anomalies in teacher results caused by test scaling.

Growth Ratings and Score Ranges

Growth Rating	Description	Growth Score Range (2011-12)
Highly Effective	Well-above state average for similar students	18-20
Effective	Results meet state average for similar students	9-17
Developing	Below state average for similar students	3-8
Ineffective	Well-below state average for similar students	0-2

Distribution of 10-11 Teacher Level MGPs

For illustrative purposes only



NOTE: Beta results using available 2010-2011 data.

HEDI classification approach for Teachers

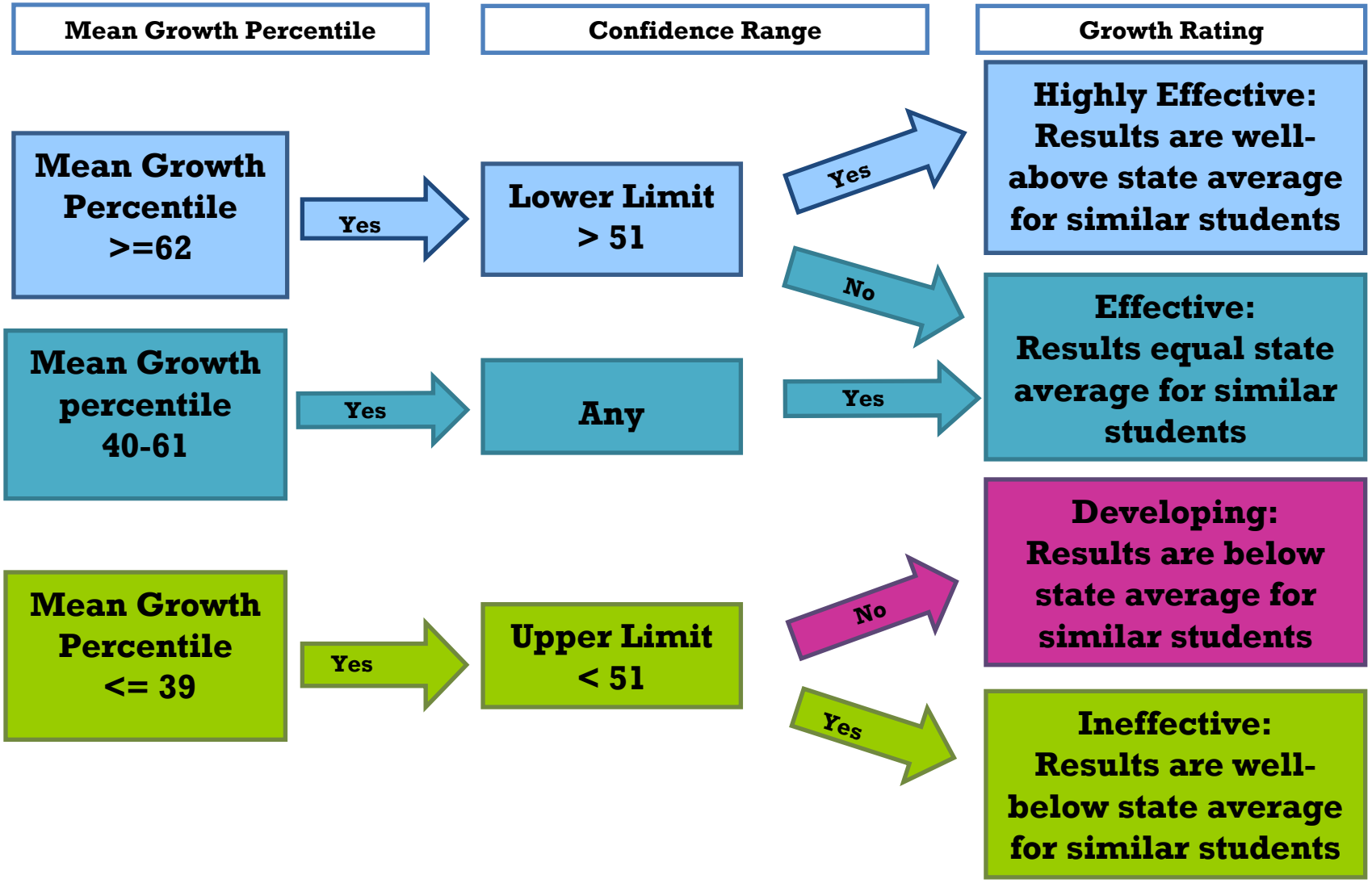
(using 10-11 sample data)

- Effective requires MGPs within 1 standard deviation of the average MGP of 51.
  - MGPs between 40 and 61 will earn Effective ratings.

- Well Above Average (Highly Effective) requires:
  - MGP of 62 or higher
  - AND confidence range above 51. (If not, rating is Effective)

- Well Below Average (Ineffective) requires
  - MGP of 39 or lower
  - And Confidence Range must be less than 51. (If not rating is Developing.)

From MGPs to Growth Ratings



Illustrative results: Teachers

(10-11 Sample Data)

Rating & 11-12 Points	Number of Teacher MGPs	Percent of Teacher MGPs
Highly Effective 18-20	1618	7%
Effective 9-17	16,681	76%
Developing 3-8	2015	10%
Ineffective 0-2	1419	7%

Points available within each HEDI category will be assigned based on educator MGP

HEDI classification approach for Principals
(using 10-11 sample data)

Same methodology as for Teachers. Slightly different cut scores

- Effective requires MGPs within 1 standard deviation of the average MGP of 50.
  - MGPs between 43 and 57 will earn Effective ratings.
- Well Above Average (Highly Effective) requires:
  - MGP of 58 or higher
  - AND confidence range above 50. (If not, rating is Effective)
- Well Below Average (Ineffective) requires
  - MGP of 42 or lower
  - And Confidence Range must be less than 50. (If not rating is Developing.)

Illustrative results: Principals

(10-11 Sample Data)

Rating & Points (2011-12)	Number of Principal MGPs	Percent of Principal MGPs
Highly Effective 18-20	345	10%
Effective 9-17	2696	75%
Developing 3-8	318	9%
Ineffective 0-2	241	7%

Points available within each HEDI category will be assigned based on educator MGP

Data – What to Expect When

Test scores finalized and teacher linkage data final submission

Growth scores provided to districts

Online reporting system available

Data – What to Expect in August

Data Elements (for teachers and schools)

- Unadjusted mean growth percentiles (Unadjusted MGPs)
- Adjusted mean growth percentiles (Adjusted MGPs and upper and lower limits based on confidence range for these adjusted MGPs)
- Percent of students above the median
- Number of student scores included
- Growth rating (HEDI)
- Growth score (0-20)

Breakdowns (by teacher and school)

- MGPs by subject, grade (not HEDI)
- Overall MGPs for subgroups – ELL, SWD, Economic Disadvantage, High- and Low-Achieving (levels 1 and 4)

engage ny

Our Students. Their Moment.