

Accountability 2.0 Next-Generation Design & Performance

Richard J. Wenning
rwennig99@gmail.com
www.twitter.com/rwennig

www.schoolview.org



This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](https://creativecommons.org/licenses/by-nc-sa/3.0/).

Overview

- The performance imperative
- Understanding academic performance
- Designing a coherent system
 - Themes: fair, useful, engaging, insightful, transparent
- Selecting a growth model
- Differentiated accountability & support

Accountability Complexity

- Accountability for educator effectiveness now layered onto systems for student, school, district, state & federal accountability
- Better when these multiple layers are aligned to support the business we are in

3

Our Business

- Maximize student progress toward & attainment of college and career readiness
 - Bright line: **all kids ready by exit**
 - Requires a definition of readiness & the content & performance standards leading there
 - Requires measurement system that determines how well students are progressing toward & reaching the destination

4

Policy Perspective on Growth

- **Why is measuring student growth so important?**
 - NCLB (Accountability 1.0) had right intent but...
 - AYP metric not useful for school performance management
 - Incentives focused on short-term increases in percent proficient, on “bubble” kids, invited moral hazard
 - Instead of long-term effectiveness and progress for all kids toward college & career readiness
 - ESEA waivers & reauthorization provides opportunity to get the measures & incentives right

5

Next Generation Performance

- Dramatic, not incremental improvements required for students that need to catch up to become college & career ready (CCR)
 - **From a system where most students that start behind stay behind to a system where most catch up**
- Implies that our accountability systems should provide information that fuels a consensus for change & capacity for improvement

6

Desired System: Accountability 2.0

- Coherent system focused on learning and building student, educator, school, district, state and federal performance management capacity.
 - Maximize local ownership of high quality information to drive insight and action
- We should ensure educator effectiveness conversation not stuck in Accountability 1.0

7

Consequential Validity

- Henry Braun (2008)
 - *Assessment practices and systems of accountability are consequentialy valid if they generate useful information and constructive responses that support one or more policy goals (Access, Quality, Equity, Efficiency) within an education system, without causing undue deterioration with respect to other goals.*

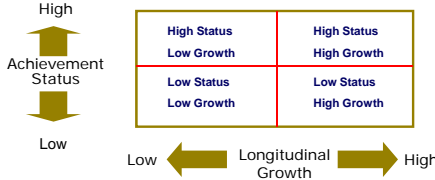
8

Marshaling a Consensus for Change

There is a difference between retrospectively identifying fault and blame-worthiness and a prospective strategy for corrective actions and building a consensus for a vision of change.

Christopher Edley (2006)

Understanding School Performance



Coherent Systems Serve Multiple Purposes



What Models?

- What statistical models of longitudinal student growth will promote the most coherence and alignment in our accountability system?

Some Framing Ideas

- *We understand best those things we see emerge from their very beginnings.*
--Aristotle
- *All Models are wrong but some are useful.*
--George E. P. Box
- *It is better to have an approximate answer to the right question than a precise answer to the wrong question.*
--John Tukey

13

Questions Set the Table

- Growth models address specific questions
 - Different techniques are good at answering different questions
 - Different questions lead to different conversations which lead to different uses and outcomes
 - Starting with the right questions simplifies development and motivates the proper use of the growth model results

14

Policy Q's: Educator Effectiveness

- What questions do we want to answer about growth rates of students associated with educators?
 - What matters and when? Student growth rates as evidence of effectiveness or ineffectiveness?
- How many categories of effectiveness and ineffectiveness are important?
 - Which categories are consequential and for what?
- What body of evidence will be used to evaluate and infer teacher value-added?
- Normative and criterion-referenced growth?

15

How much growth did a student make and is it enough?

- **Describing** growth versus **ascribing** responsibility
 - The Colorado Growth Model began by separating the description of growth from discussions of responsibility/accountability
 - Incorporating growth into accountability followed from the accepted description of growth
 - The description of growth facilitated stakeholder engagement and investigations of responsibility for good/bad growth
 - That in turn led to greater stakeholder support for particular forms of accountability

16

Describing Student Growth

- Discussing student growth, even with a vertical scale, is not a simple task
- Growth and change require context. Consider, for example, height:
 - A child might grow 4 inches between ages 3 and 4
 - 4 inches is a well understood quantity
 - The 4 inch increase becomes meaningful only when understood alongside the growth of other 3 to 4 year olds
- Student growth percentiles were developed to provide a norm-referenced basis for describing student growth

17

Who/What is Responsible for Student Growth?

- Some analyses of student growth attempt to determine the amount of student progress that can be attributed to the school or teacher
 - Called value-added analyses, these techniques attempt to estimate the teacher/school contribution to student academic growth
- Value added is an inference – a causal conclusion drawn from the data
- All growth models can be used for value-added purposes

18

Colorado Growth Model Asks...

- What is? ⇨ How much growth did a child make in one year?
- What should be? ⇨ How much growth is enough to reach college & career readiness?
- What could be? ⇨ How much growth have other students made with the same starting point?

19

Student Growth Percentiles

- Should we be surprised with a child's current achievement given their prior achievement?
 - Student growth percentiles answer this question
- Consider a low achieving student with 90th percentile growth and a high achieving student with 10th percentile growth
 - The low achieving student grew at a rate exceeding 90 percent of similar students
 - The high achieving student grew at a rate exceeding just 10 percent of similar students
 - The low achiever's growth is more exemplary (probabilistically) than the high achiever's
- Judgments about the adequacy of student growth require external criteria together with standard setting

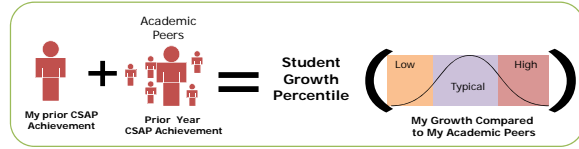
20

Establishing Growth Standards Based Upon Growth Norms

- The most common adequacy criterion is judging growth toward an achievement goal (i.e., growth-to-standard)
- Results from student growth percentile analyses can be used to calculate percentile growth trajectories for each student
- These trajectories indicate what future rates of growth will lead to and are used to make adequacy judgments
- This growth-to-standard approach was approved as part of Colorado's successful application to the Growth Model Pilot Program

21

Understanding Student Growth Percentiles

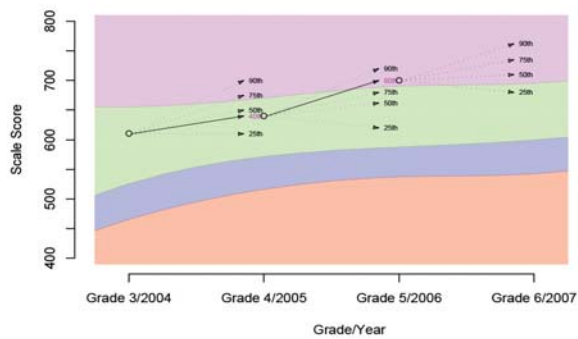


What is Student Adequate Growth Percentile (AGP)?



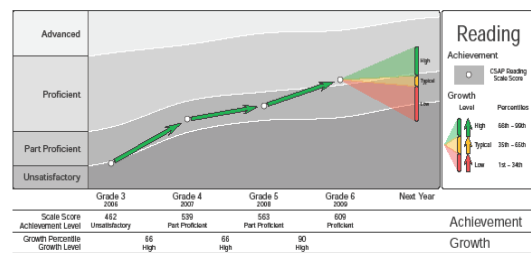
22

Individual Student CSAP Growth Chart: Reading 2004 to 2007



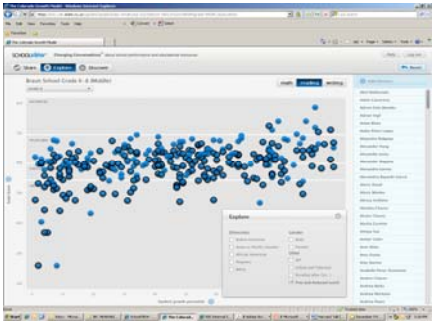
23

One Student's Growth Percentiles



24

Students in a Grade in a School

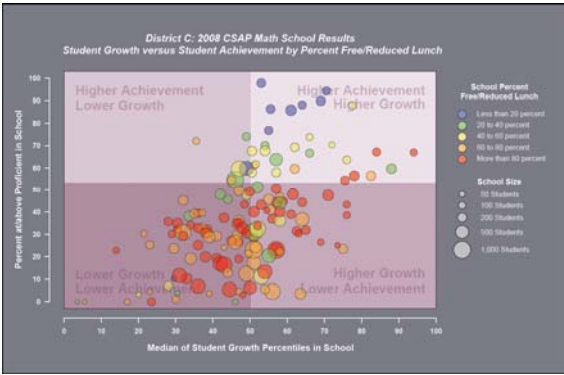


25

Views of Schools within a District



26



27

View of Districts



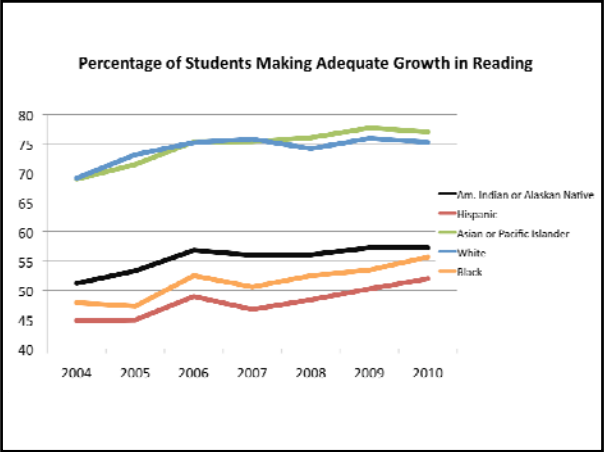
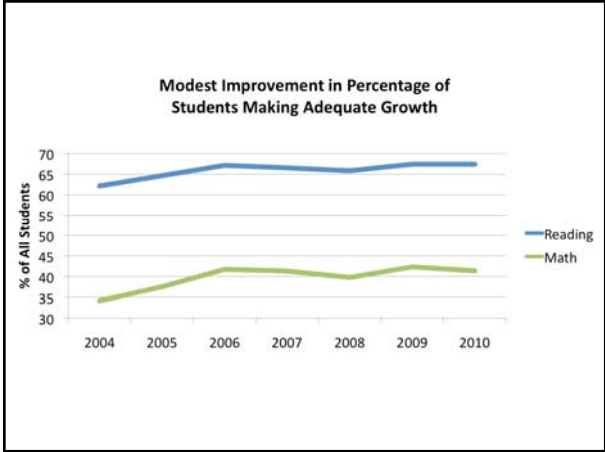
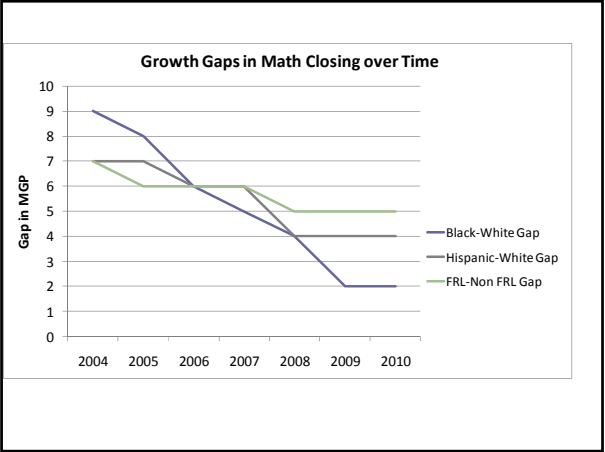
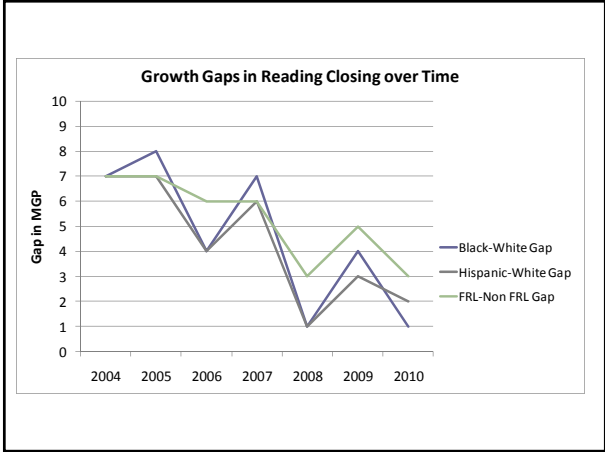
28

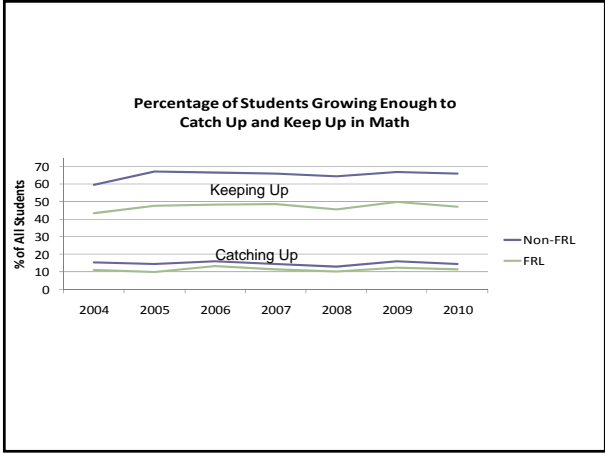
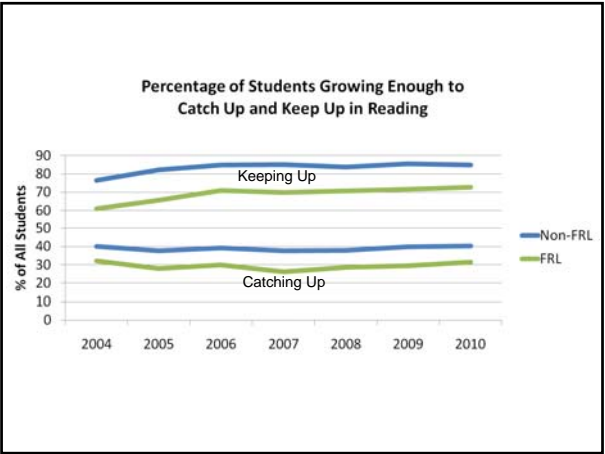
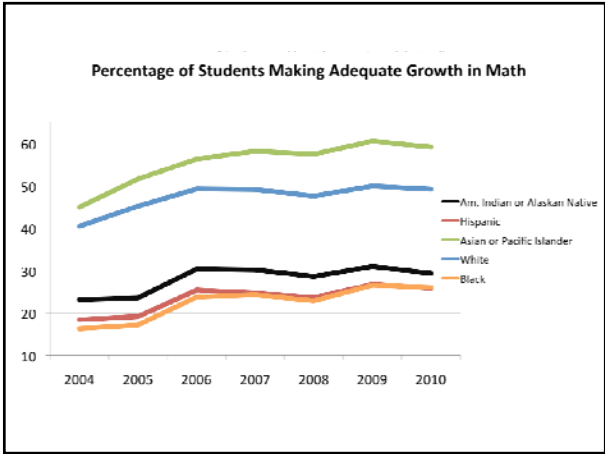
District Performance Framework Report 2010 - 2011 Level: All Levels

Performance Indicator	Rating/Rate	% of Points Earned out of Points Possible*
Student Achievement	Meets	98% (24.6 out of 25 points)
Student Growth	Meets	98% (24.6 out of 25 points)
Student Achievement Gap	Meets	98% (24.6 out of 25 points)
Professional Development	Meets	98.2% (24.6 out of 25 points)
Student Achievement**	Meets	98.2% (24.6 out of 25 points)
TOTAL	Meets	98.2% (24.6 out of 25 points)

- ### 15 States with MOU
- 15 states have signed MOU to use the Colorado Growth Model growth percentile methodology and SchoolView display tools:
 - Arizona, Colorado, Georgia, Indiana, Kansas, Massachusetts, Missouri, Nevada, New Hampshire, New York, Rhode Island, Virginia, Washington, West Virginia, Wisconsin
 - Pursue cloud-based analytics and visualizations
 - Creative Commons Share Alike Noncommercial Use: <http://creativecommons.org/licenses/>

- ### Colorado Growth Model Shows...
- Growth not correlated with how close a student is to proficient
 - Normative growth gaps narrowing among student groups
 - Most low achieving students not growing fast enough to catch up
 - Many proficient students not keeping up
 - Many high achieving schools show low growth and many low achieving schools show high growth





- Differentiated Accountability & Support - Key Components**
1. Key Performance Indicators
 2. Multi-Measure Framework
 3. Incentives for Change & Innovation
 4. Planning Process
 5. Service Mix
 6. Service Delivery
 7. Evaluation & Validation
 8. Rollout Strategy - Communications, Stakeholder Engagement, Training

Key Performance Indicators (KPIs)

- Establish KPI's and a multi-measure performance framework used for District, School, and educator accountability purposes.
 - Growth, Status, College & Career Readiness, Gaps & others...

41

Multi-Measure Framework

- Develop a multi-measure framework with measures, metrics, and targets for each big indicator
 - **Use the framework evidence to identify schools for Reward, Focus, Priority & other state categories**
- Balance normative and criterion-referenced growth & status evidence
 - Take note of variance in state assessment cutpoints by subject
 - Consider different normative & criterion-referenced weightings for teacher, school, district, state purposes

42

Multi-Measure Framework, cont.

At least two functions:

- **Improvement** - diagnostic feedback to support a solid planning process
- **Accountability** - summative evaluation with a set of performance categories that describe overall performance across KPIs & signal rewards (money, autonomy) and consequences (intervention)

43

Incentives for Change and Innovation

Rewards, sanctions, and disclosure

- Recognition and financial awards for high growth schools & incentives to replicate
- State authority to close schools
- Public access to engaging, insightful information about student, school, district & state performance
 - Shine light on best, worst by demographics and type of school & encourage productive social collaboration

44

Planning Process

- Develop a unified planning process based on the feedback from the multi-measure framework
 - Requires a robust qualitative review component
- Promote focused statewide inquiry into evidence, root causes, planning, and improvement

45

Service Mix

- Determine the differentiated service mix for tiers of schools based on the performance categories
 - Key support for all tiers is building solid district, school, educator performance management capacity (incorporates standards and assessments & cuts across federal program silos)
 - Service mix for middle tier?
 - Intervention mix for Gap schools? Measures matter a great deal in diagnosing the problem (status vs. growth gaps)
 - Intensive intervention for bottom 5% (Transformation, Turnaround, Replacement – consider grade span)

46

Service Delivery Strategy

- Role of SEA central (delivery across silos)
 - Regardless of local control context, foundation is quantitative & qualitative review of performance & practice with a consistent planning & evaluation process
- Role of regional delivery structures (education service agencies)?
- Role of Third Parties (EMOs, CMOs, Consultants) & SEA due diligence?

47

Evaluation Strategy

- Multi-measure framework, implementation benchmarks, qualitative reviews provide formative & summative feedback on success of support & interventions
- Key validation of measures:
 - extent of regular, constructive, and coherent use in discourse & practice across system levels
 - observed improvement in what different growth rates obtain in proficiency and CCR @ exit
- Establish a third-party evaluation process to compliment internal review of evidence

48

Rollout Strategy Considerations

- Plan to bring all stakeholders along, establishing ownership, setting expectations that the SEA & they can deliver on
- Rollout of evidence: Is there time for sequence of no, low, then high stakes implementation?
- Sequence of statewide & local communications & training

49