Commissioner’s Performance-Based Accountability Task Force: Indicators and Growth

DEB WISWELL & TIM KURTZ
NH DOE
SCOTT MARION
CENTER FOR ASSESSMENT
NOVEMBER 30, 2009
Before discussing the indicators and student growth percentiles, we want to quickly review three common approaches for calculating indicator values.

This will be related to your homework...so yes, there will be a quiz!
Growth, Improvement, & Status

- **Status** = A point-in-time measurement, e.g., 72% of the students are proficient in math
- **Improvement** = Is generally the change in status measures when students are NOT matched, e.g., 5% more students are proficient this year compared to last year
- **Growth** = Is based on determining changes over time based on following the same students, e.g., our 5th grade students grew at a rate of X compared to where they were in 4th grade.
AYP Task Force members (including Emma Rous) provided ratings

Used median values to try to summarize value ratings

Many people “wanted it all”
Medians/SD for “readiness” indicators

- Drop out rate (or decrease in)
- Graduation Rate
- College/ career readiness measure
- Competency Completions
- SAT/ ACT/ Accuplacer (number who take? Score?)
- # of State scholars
- Applications to college or other postsecondary
- Diploma type (number who earn)
- AP Courses
- Course taking patterns

Commissioner’s Performance Based Accountability System Task Force. 11/30/09
Medians/SD for “test-based” indicators

- NECAP scores/ Subgroups
- SAT/ ACT/ Accuplacer (number who take? Score?)
- NWEA (or other adaptive test) scores
- Standardized Tests (ITBS, CAT, Terra Nova)
- Curriculum assessments (e.g., tests)
- Local Common Assessments
- DIBLES
- Report card grades

Commissioner's Performance Based Accountability System Task Force. 11/30/09
HS-only indicators

- Drop out rate (or decrease in)
- Graduation Rate
- College/ career readiness measure
- Competency Completions
- SAT/ ACT/ Accuplacer (number who take? Score?)
- # of State scholars
- # of extracurricular activities provided – Participation?
- Applications to college or other postsecondary
- Diploma type (number who earn)
- AP Courses
- Course taking patterns

std deviation
Median

Commissioner's Performance Based Accountability System Task Force. 11/30/09
Values for already-collected indicators

- Drop out rate (or decrease in)
- Graduation Rate
- NECAP scores/ Subgroups
- SAT/ ACT/ Accuplacer (number who take? Score?)
- # of State scholars
- NWEA (or other adaptive test) scores
- Standardized Tests (ITBS, CAT, Terra Nova)
- # of extracurricular activities provided – Participation?
- Applications to college or other postsecondary
- Diploma type (number who earn)
- AP Courses
- Relative attendance, disaggregated
- DIBLES and others

Commissioner's Performance Based Accountability System Task Force. 11/30/09
Values for indicators not collected

- College/ career readiness measure
- Competency Completions
- # of State scholars
- Curriculum assessments (e.g., tests)
- # of extracurricular activities provided – Participation?
- Parent perceptions
- Student survey/ perceptions (climate)
- Applications to college or other postsecondary
- Local Common Assessments
- Course taking patterns
- Interventions
- Report card grades

Commissioner's Performance Based Accountability System Task Force. 11/30/09
Reminder!

- Remember, this component of the Adequacy accountability system is designed to focus on PERFORMANCE.
- We already have a system designed to focus on INPUTS.
- Performance generally refers to outputs of the system (e.g., graduation rate) and not on process indicators such as teacher actions or parental support.
  - We’re not saying that those things aren’t important. They are simply beyond the scope of our charge here.
**Question #1:** Which of these do you REALLY want to collect (1=not important; 5=critically important). Be parsimonious!

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Importance Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local Common (w/in district) Assessments</td>
<td>K-8</td>
</tr>
<tr>
<td>Completion of Competency Exams</td>
<td></td>
</tr>
<tr>
<td>College/ career readiness measure (other than SAT/ACT)</td>
<td></td>
</tr>
</tbody>
</table>

Commissioner's Performance Based Accountability System Task Force. 11/30/09
**Question #2:** Which of these *already-collected* indicators do you REALLY want to include (1=not important; 5=critically important)? Let’s focus on minimizing redundancy!

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Importance Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early (K-2) reading measures</td>
<td>K-8</td>
</tr>
<tr>
<td>Attendance</td>
<td></td>
</tr>
<tr>
<td>AP Courses (#enrolled/#passing)</td>
<td></td>
</tr>
<tr>
<td>Applications to college or other postsecondary</td>
<td></td>
</tr>
<tr>
<td>Standardized Tests (NWEA, ITBS, CAT, Terra Nova)</td>
<td></td>
</tr>
<tr>
<td>SAT/ ACT/ Accuplacer (number who take? Score?)</td>
<td></td>
</tr>
<tr>
<td>NECAP scores/ Subgroups</td>
<td></td>
</tr>
<tr>
<td>Diploma type (number who earn each type)</td>
<td></td>
</tr>
<tr>
<td>Graduation Rate</td>
<td></td>
</tr>
<tr>
<td>Drop out rate (or decrease in)</td>
<td></td>
</tr>
</tbody>
</table>

Commissioner's Performance Based Accountability System Task Force. 11/30/09
Question 3: So what’s the final list?

- Elementary/Middle School?

- High School?
What questions can student growth percentiles be used to address?

What are student growth percentiles and percentile growth projections/trajectories?

What role do the results from growth percentile analyses play in accountability system determinations?

What role do the results from growth percentile analyses play in program evaluations?
Relevant Questions for Parents

- Yen (2007), from a state survey of parents, teachers and administrators, compiled a list of frequently voiced questions/concerns by various stakeholder groups.

- Common Parent Questions
  - Did my child make a year’s worth of progress in a year?
  - Is my child growing appropriately toward meeting state standards?
  - Is my child growing as much in Math as Reading?
  - Did my child grow as much this year as last year?
Relevant Questions for Teachers

- Did my students make a year’s worth of progress in a year?
- Did my students grow appropriately toward meeting state standards?
- How close are my students to becoming Proficient?
- Are there students with unusually low growth who need special attention?
Relevant Questions for Administrators

- Did the students in our district/school make a year’s worth of progress in all content areas?
- Are our students growing appropriately toward meeting state standards?
- Does this school/program show as much growth as that one?
- Can I measure student growth even for students who do not change proficiency categories?
- Can I pool together results from different grades to draw summary conclusions?
Descriptive Questions

- Note that the questions put forward by stakeholders are primarily descriptive.
- The questions are only peripherally associated with causality.
- High stakes accountability has transformed questions about student growth into questions about responsibility/cause: Teacher and School Effectiveness such as the case with Value-Added Modeling (VAM).
- We argue that the place to begin is with description and a model that supports such uses.
Accountability system results can have value without making causal inferences about school quality, solely from the results of student achievement measures and demographic characteristics. Treating the results as descriptive information and for identification of schools that require more intensive investigation of organizational and instructional process characteristics are potentially of considerable value. Rather than using the results of the accountability system as the sole determiner of sanctions for schools, they could be used to flag schools that need more intensive investigation to reach sound conclusions about needed improvements or judgments about quality [Linn, 2008, p. 21, emphasis added].
Measuring student growth, even with a vertical scale, is not a simple task. Some believe a vertical scale simplifies the task of measuring student growth. Even with an interval (or ratio) scale, growth is not easy to interpret.

Consider, for example, a child’s height.
- A child might grow 4 inches between ages 3 and 4.
- 4 inches is a well understood quantity.
- The 4 inch increase becomes really meaningful only when understood alongside the growth of other 3 to 4 year olds.

Student growth percentiles were developed to provide a normative context for describing student growth.
Real data are a little noisier
Should we be surprised with a child’s current achievement given their prior achievement?

• Given a student’s prior scale scores and the associated conditional density, their current scale score corresponds to a percentile of that conditional distribution.
  ○ In other words, comparing a student’s growth to students with the same prior score history reveals how well that student performed relative to other students with the same history.

• This percentile is the student’s growth percentile.

• Growth percentiles are closely related to estimating the probability of observing a student’s current achievement taking account of their past achievement:
  ○ Pr(Current Achievement|Past Achievement).

• As such, growth percentiles describe the rarity of a student’s current achievement conditional upon their prior achievement.
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 judgmental and/or external criteria.
Remember for every starting point (score history), there is a distribution of scores in Year 2 (or Year n). In other words, for every starting point, there will be a range of growth percentiles from 1 to 99. Therefore, growth percentiles are completely unrelated to initial status (i.e., score history). This means that growth percentiles are an extremely fair way to describe student growth all along the achievement continuum.
Aggregating Student Growth Percentiles

- For program and instructional uses within a school, it is critical to have access to each student’s growth percentile.
- For school accountability purposes, it is useful to have a summary measure.
- Often we use an “average” or the mean.
- However, for several reasons, the median—or middle value—is more appropriate for student growth percentiles.
What’s the Median?

<table>
<thead>
<tr>
<th>Left Side</th>
<th>Right Side</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>50</td>
<td>90</td>
</tr>
</tbody>
</table>

**Median = ?**

**Commissioner’s Performance Based Accountability System Task Force / 11/30/09**
Mountain View School District: 2007 CSAP Reading School Results
Student Growth versus Student Achievement by Free/Reduced Lunch Percentage

Median of Student Growth Percentiles in School

Percentage of Proficient/Advanced Students in School

School Percent
Free/Reduced Lunch
- Less than 20 percent
- 20 to 40 percent
- 40 to 60 percent
- 60 to 80 percent
- More than 80 percent

School Size
- 50 Students
- 100 Students
- 200 Students
- 500 Students
- 1,000 Students
The Colorado Reporting System

- The public website

https://cdeapps.cde.state.co.us/growth_model_public/

- And the private website:

https://cdeapps.cde.state.co.us/growth_model_demo/
Homework

- Please complete and email to Deb your ratings (by December 7th) on the following two slides:
Homework Question #1: Please indicate whether you think we should use, if possible and valid, the indicator for growth (g), improvement (i), status (s), or multiple (m).

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Growth/Improve/Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local Common (w/in district) Assessments</td>
<td>K-8</td>
</tr>
<tr>
<td>Competency Completions</td>
<td></td>
</tr>
<tr>
<td>College/ career readiness measure (other than SAT/ACT)</td>
<td>HS</td>
</tr>
</tbody>
</table>
Homework Question #2: Please indicate whether you think, if possible and valid, we should use the indicator for growth (g), improvement (i), status (s), or multiple (m).

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Growth/Improve/Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early (K-2) reading measures</td>
<td>K-8</td>
</tr>
<tr>
<td>Attendance</td>
<td></td>
</tr>
<tr>
<td>AP Courses (#enrolled/#passing)</td>
<td>HS</td>
</tr>
<tr>
<td>Applications to college or other postsecondary</td>
<td></td>
</tr>
<tr>
<td>Standardized Tests (NWEA, ITBS, CAT, Terra Nova)</td>
<td></td>
</tr>
<tr>
<td>SAT/ ACT/ Accuplacer (number who take? Score?)</td>
<td></td>
</tr>
<tr>
<td>NECAP scores/ Subgroups</td>
<td></td>
</tr>
<tr>
<td>Diploma type (number who earn each type)</td>
<td></td>
</tr>
<tr>
<td>Graduation Rate</td>
<td></td>
</tr>
<tr>
<td>Drop out rate (or decrease in)</td>
<td></td>
</tr>
</tbody>
</table>

Commissioner's Performance Based Accountability System Task Force. 11/30/09