Developing Connecticut’s Growth Model for the Smarter Balanced Summative Assessments in English Language Arts (ELA) and Mathematics

November 2016
Contents

Introduction .................................................................................................................................................. 3
Achievement versus Growth......................................................................................................................... 3
Three Ways to Understand Change in Performance .................................................................................. 3
What is a Growth Model ............................................................................................................................... 4
The Smarter Balanced Vertical Scale ............................................................................................................ 5
Determining the Growth Categories ............................................................................................................. 5
Establishing Ambitious Yet Achievable Growth Targets ............................................................................ 7
Final Growth Target Tables ........................................................................................................................... 8
Outcome Measures ....................................................................................................................................... 9
Growth Models and Value-Added Models ................................................................................................. 10
Conclusion ................................................................................................................................................... 11
Reference .................................................................................................................................................... 11
Introduction

This paper describes the development of Connecticut’s Growth Model for the Smarter Balanced Summative Assessments in English Language Arts (ELA) and Mathematics. It applies to students in grades 4 through 8. This growth model provides ambitious yet achievable individual student growth targets for all students. The aggregate results from this growth model will be reported publicly and used as a key component (i.e., Indicator 2- Growth) of Connecticut’s Next Generation Accountability System for districts and schools.

Achievement versus Growth

Before diving deeper into the growth model, let’s first understand the differences between achievement and growth. Here’s a simple definition of achievement.

- **Achievement** or proficiency or status is a one-time snapshot measurement of a student’s academic performance in a subject area like ELA or Math. It is an indicator of how well a student or a group of students performed on the standards assessed by the test at a specific point in time.
- **Growth** on the other hand is about the change in achievement scores for the *same student* between two or more points in time.

Three Ways to Understand Change in Performance

To further understand the concept of growth, let’s contrast three ways in which educators commonly understand change in student performance.

1. **Achievement Change** simply compares student achievement for the same grade across years. For example, a superintendent may say that the proficiency rate of students in grade 4 in our district has increased from 50% in one year to 53% in the next year, an improvement of 3 percentage points. While that is technically accurate, this approach is actually comparing the performance of two different groups of fourth graders. The difference in performance between the two groups may be due to the fact that the groups are different to begin with; maybe the
higher performing group of fourth graders just started off higher. This approach is really just the
starting point for understanding change in performance.

2. In the “Rough Cohort” Change approach, for example, a superintendent may compare the
proficiency rate of this year’s fourth graders to that of last year’s third graders. If your district
experiences little student mobility and almost all students are promoted from one grade to the
next each year, most of the students will be the same across years. However, if your district
experiences high student mobility, a greater percentage of students across the two years will be
different.

3. The Matched Student Cohort Change (or Growth) compares the achievement of the same
student from one grade in year 1 to the next higher grade in year 2. This is generally considered
the gold standard for growth because there are no mismatched students; only those students
who are matched across years are included in the calculation. The matched approach allows us
to quantify the amount of growth achieved by the same students from near the end of one
grade, to the end of the next grade – a good measure of curriculum and instructional
effectiveness.

What is a Growth Model?

While growth describes the change in achievement for the same student over two or more points in
time, a growth model according to Castellano and Ho (2013) “is a collection of definitions, calculations,
or rules that summarizes student performance over two or more time points and supports
interpretations about students, their classrooms, their educators, or their schools.” In effect, a growth
model can help to set appropriate student achievement targets, monitor student growth in achievement
toward those targets, and identify students who are not growing at an adequate rate.

Castellano and Ho (2013) describe a few different growth models. These include the Gain Score Model,
the Categorical Model, the Growth-to-Standard Model, the Student Growth Percentile Model, and the
Multivariate Model. Different models require different measures as their foundation and enable
different interpretations. For example, the Student Growth Percentile model uses a normative approach
and evaluates a student’s growth relative to the growth achieved by his/her academic peers. On the
other hand, a growth-to-standard model utilizes a vertical scale and evaluates a student’s growth relative to a fixed criterion for gain on that scale.

### The Smarter Balanced Vertical Scale

Connecticut’s growth model is based on the Smarter Balanced vertical scale scores for ELA and Mathematics. This vertical scale spans the grades from 3 through 8. The vertical scale scores are derived directly from a linear transformation of the Item Response Theory (IRT) proficiency estimates with fixed highest and lowest obtainable scale scores for each grade. ELA and Math scales range from around 2100 to 2800. Each vertical scale score is mapped into one of the four achievement levels per grade. The achievement level designations are Level 1-**Not Met**, Level 2-**Approaching**, Level 3-**Met**, and Level 4-**Exceeded**.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Subject</th>
<th>Lowest scale score</th>
<th>SS Cut between Levels 1 and 2</th>
<th>SS Cut between Levels 2 and 3</th>
<th>SS Cut between Levels 3 and 4</th>
<th>Highest scale score</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>ELA</td>
<td>2114</td>
<td>2367</td>
<td>2432</td>
<td>2490</td>
<td>2623</td>
</tr>
<tr>
<td>4</td>
<td>ELA</td>
<td>2131</td>
<td>2416</td>
<td>2473</td>
<td>2533</td>
<td>2663</td>
</tr>
<tr>
<td>5</td>
<td>ELA</td>
<td>2201</td>
<td>2442</td>
<td>2502</td>
<td>2582</td>
<td>2701</td>
</tr>
<tr>
<td>6</td>
<td>ELA</td>
<td>2210</td>
<td>2457</td>
<td>2531</td>
<td>2618</td>
<td>2724</td>
</tr>
<tr>
<td>7</td>
<td>ELA</td>
<td>2258</td>
<td>2479</td>
<td>2552</td>
<td>2649</td>
<td>2745</td>
</tr>
<tr>
<td>8</td>
<td>ELA</td>
<td>2288</td>
<td>2487</td>
<td>2567</td>
<td>2668</td>
<td>2769</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade</th>
<th>Subject</th>
<th>Lowest scale score</th>
<th>SS Cut between Levels 1 and 2</th>
<th>SS Cut between Levels 2 and 3</th>
<th>SS Cut between Levels 3 and 4</th>
<th>Highest scale score</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Math</td>
<td>2189</td>
<td>2381</td>
<td>2436</td>
<td>2501</td>
<td>2621</td>
</tr>
<tr>
<td>4</td>
<td>Math</td>
<td>2204</td>
<td>2411</td>
<td>2485</td>
<td>2549</td>
<td>2659</td>
</tr>
<tr>
<td>5</td>
<td>Math</td>
<td>2219</td>
<td>2455</td>
<td>2528</td>
<td>2579</td>
<td>2700</td>
</tr>
<tr>
<td>6</td>
<td>Math</td>
<td>2235</td>
<td>2473</td>
<td>2552</td>
<td>2610</td>
<td>2748</td>
</tr>
<tr>
<td>7</td>
<td>Math</td>
<td>2250</td>
<td>2484</td>
<td>2567</td>
<td>2635</td>
<td>2778</td>
</tr>
<tr>
<td>8</td>
<td>Math</td>
<td>2265</td>
<td>2504</td>
<td>2586</td>
<td>2653</td>
<td>2802</td>
</tr>
</tbody>
</table>

### Determining the Growth Categories

Connecticut’s growth model uses the matched student cohort change approach and can be thought of as a growth-to-standard approach. It is based on the Smarter Balanced vertical scale. The model establishes ambitious yet achievable vertical scale score targets for each student.

The amount of growth achieved by students performing at different points on the vertical scale can vary greatly. Generally, students at higher levels of achievement show smaller amounts of growth. Therefore, the CSDE utilized an approach to divide each of the nationally established Smarter Balanced
achievement levels into two; this approach was similar to the one used in the past with the Connecticut Mastery Test (CMT).

In order to segment the achievement levels, the CSDE computed average standard errors of measurement (SEM) for students performing at each cut point. The figures below illustrate a band of one average SEM around the mean for students performing at the cut score in both ELA and Math for all grades. Ideally, the new cut points should be out of the standard error range to ensure that scores in different levels are reliably different.

ELA Cut Points and their SEM’s by Grade

Math Cut Points and their SEM’s by Grade
At one SEM, the errors bands across achievement levels in a grade did not overlap. This provided an opportunity to place the new cut points at the gaps between the average SEMs of the original adjacent cut point scores. Each achievement band was divided into two – low and high – yielding a total of eight categories.

**Establishing Ambitious Yet Achievable Growth Targets**

The eight categories served as the foundation for descriptions of growth. Several data cleaning steps were employed to ensure the quality of the analyses. For example, only students who had scale scores in both 2014-15 (year 1) and year 2015-16 (year 2) and who were promoted to the next grade from year 1 to year 2 were retained in the analyses.

The first step was to determine the actual amount of growth achieved by Connecticut students from 2014-15 to 2015-16 within each of the eight categories. Each student was assigned to one of the eight categories based on the student’s year1 score. For each student, the growth amount was calculated by subtracting the year1 score from the year2 score.

\[
growth \ amount = Score_{year2} - Score_{year1}
\]

In addition to the growth amount, the standard error of the growth amount was also computed for each student. Psychometric theory tells us that a test score is an estimate of a student’s achievement and contains a certain amount of measurement error. When calculating growth, we are comparing test scores from two tests, each of which has error. The standard error of the growth amount takes into account the error in both scores and is calculated as

\[
SE \ of \ the \ growth \ amount = \sqrt{SE(Score_{year1})^2 + SE(Score_{year2})^2}
\]

The standard error of the growth amount was taken into account to help inform decisions about the growth targets.

The percentiles of the growth amounts within each of the eight categories were then calculated. The purpose of studying the distribution of the growth amount was to determine a growth standard that is ambitious (i.e., achieving the targets annually put students on a path to higher levels of achievement in future years) and achievable (i.e., the targets were achieved by a reasonable percentage of students).
For example, the 75th percentile of the growth amount in a category indicates that 25 percent of the students in that category achieved this amount of growth or more from year 1 to year 2. The full range of the distribution in each category was examined thoroughly but special focus was placed on the mid-to-high ranges (e.g., 50th, 60th, 70th, 75th percentiles). The selected growth amounts were also translated into trajectories to see if students at different starting achievement categories reached higher categories in future years if they achieved their respective targets in each grade.

The growth amounts at selected points of the distribution were also compared to the average of the standard errors of the growth amounts in each category. We wanted to choose a growth amount target that exceeded the average standard errors of the growth amount in most categories.

This model was reviewed by Connecticut’s Technical Advisory Committee, which is a group of psychometric experts from around the country. It was also discussed with local educators.

The final decision was to set the growth targets at a point where 40% of the students within each category attained those targets from 2014-15 to 2015-16. Though a percentile distribution was used to determine the ambitious yet achievable growth amounts, these amounts are now established as a fixed criterion for at least the next few growth cycles.

**Final Growth Target Tables**

The final growth target tables for ELA and Mathematics are presented below. Here is an example to illustrate how to determine the growth target amount for a student. If a Grade 3 student earns a Smarter Balanced ELA vertical scale score of 2350 in the first year, this places the student in the High Level 1 category in Grade 3 (highlighted below). By the end of grade 4, this student will be expected to grow 71 points from 2350, or in other words, achieve a vertical scale score of at least 2421. Note that sometimes students achieving their growth target may not advance from one category in one grade to the next higher category in the next grade. This is not a categorical growth model but one based purely on the vertical scale scores.
### ELA Achievement Level Ranges and Growth Targets

<table>
<thead>
<tr>
<th>Grade in Yr.</th>
<th>Level</th>
<th>Level 1: Not Met</th>
<th>Level 2: Approaching</th>
<th>Level 3: Met</th>
<th>Level 4: Exceeded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 - LOW</td>
<td>2 - HIGH</td>
<td>3 - LOW</td>
<td>4 - HIGH</td>
</tr>
<tr>
<td>3</td>
<td>Range</td>
<td>2114-2330</td>
<td><strong>2331-2366</strong></td>
<td>2367-2399</td>
<td>2400-2431</td>
</tr>
<tr>
<td></td>
<td>Target</td>
<td>82</td>
<td><strong>71</strong></td>
<td>70</td>
<td>69</td>
</tr>
<tr>
<td>4</td>
<td>Range</td>
<td>2131-2378</td>
<td>2379-2415</td>
<td>2416-2444</td>
<td>2445-2472</td>
</tr>
<tr>
<td></td>
<td>Target</td>
<td>82</td>
<td>69</td>
<td>69</td>
<td>64</td>
</tr>
<tr>
<td>5</td>
<td>Range</td>
<td>2201-2405</td>
<td>2406-2441</td>
<td>2442-2471</td>
<td>2472-2501</td>
</tr>
<tr>
<td></td>
<td>Target</td>
<td>69</td>
<td>56</td>
<td>55</td>
<td>48</td>
</tr>
<tr>
<td>6</td>
<td>Range</td>
<td>2210-2417</td>
<td>2418-2456</td>
<td>2457-2493</td>
<td>2494-2530</td>
</tr>
<tr>
<td></td>
<td>Target</td>
<td>73</td>
<td>58</td>
<td>53</td>
<td>47</td>
</tr>
<tr>
<td>7</td>
<td>Range</td>
<td>2258-2438</td>
<td>2439-2478</td>
<td>2479-2515</td>
<td>2516-2551</td>
</tr>
<tr>
<td></td>
<td>Target</td>
<td>69</td>
<td>50</td>
<td>49</td>
<td>44</td>
</tr>
<tr>
<td>8</td>
<td>Range</td>
<td>2288-2446</td>
<td>2447-2486</td>
<td>2487-2526</td>
<td>2527-2566</td>
</tr>
</tbody>
</table>

### Math Achievement Level Ranges and Growth Targets

<table>
<thead>
<tr>
<th>Grade in Yr.</th>
<th>Level</th>
<th>Level 1: Not Met</th>
<th>Level 2: Approaching</th>
<th>Level 3: Met</th>
<th>Level 4: Exceeded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 - LOW</td>
<td>2 - HIGH</td>
<td>3 - LOW</td>
<td>4 - HIGH</td>
</tr>
<tr>
<td>3</td>
<td>Range</td>
<td>2189-2351</td>
<td>2352-2380</td>
<td>2381-2408</td>
<td>2409-2435</td>
</tr>
<tr>
<td></td>
<td>Target</td>
<td>77</td>
<td>61</td>
<td>59</td>
<td>60</td>
</tr>
<tr>
<td>4</td>
<td>Range</td>
<td>2204-2381</td>
<td>2382-2410</td>
<td>2411-2447</td>
<td>2448-2484</td>
</tr>
<tr>
<td></td>
<td>Target</td>
<td>51</td>
<td>38</td>
<td>40</td>
<td>44</td>
</tr>
<tr>
<td>5</td>
<td>Range</td>
<td>2219-2419</td>
<td>2420-2454</td>
<td>2455-2491</td>
<td>2492-2527</td>
</tr>
<tr>
<td></td>
<td>Target</td>
<td>43</td>
<td>46</td>
<td>45</td>
<td>44</td>
</tr>
<tr>
<td>6</td>
<td>Range</td>
<td>2235-2434</td>
<td>2435-2472</td>
<td>2473-2512</td>
<td>2513-2551</td>
</tr>
<tr>
<td></td>
<td>Target</td>
<td>49</td>
<td>41</td>
<td>38</td>
<td>36</td>
</tr>
<tr>
<td>7</td>
<td>Range</td>
<td>2250-2438</td>
<td>2439-2483</td>
<td>2484-2525</td>
<td>2526-2566</td>
</tr>
<tr>
<td></td>
<td>Target</td>
<td>58</td>
<td>35</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>8</td>
<td>Range</td>
<td>2265-2456</td>
<td>2457-2503</td>
<td>2504-2544</td>
<td>2545-2585</td>
</tr>
</tbody>
</table>

### Outcome Measures

Using the growth target set for each student entering grades 4 through 8, two outcome measures are assigned to each student:

1. **Growth Rate**: This is a binary measure indicating whether a student met the growth target (i.e., value=1 or yes) or not (i.e., value=0 or no)

Developing Connecticut’s Growth Model for the Smarter Balanced Summative Assessments in English Language Arts (ELA) and Mathematics
2. **Percentage of Target Achieved (PTA):** The Percentage of Target Achieved is how much of the growth target was achieved by the student. It is calculated as follows:

\[
PTA = \frac{\text{Growth Amount}}{\text{Growth Target}}
\]

The growth rate is not a continuous measure. Students *nearly* meeting the target will be deemed to not have met the target, even if they missed the target by just 1 vertical scale score point. On the contrary, the Percentage of Target Achieved is a continuous measure. Students get credit for any growth up to and even 10 percent beyond the target.

The two growth outcome measures are aggregated for schools, districts, or student groups. This results in two measures: growth rate and average PTA. The growth rate is the percentage of students that met their target, while the average PTA is the average percentage of the growth target that was achieved across all students. The growth rate is simpler to understand while the average percentage of target achieved is more nuanced.

The CSDE will report both measures but will include the more precise, average percentage of target achieved in the district and school accountability model. The PTA for a student is capped at 110%; students earning more than 110% of the target will be deemed to have achieved 110% of the target. This ensures that unusually high student level growth do not unduly skew the PTA statistic. Also, the bottom is set at 0; students who evidence negative growth are set to 0 PTA.

**Growth Models and Value-Added Models**

The terms “growth model” and “value-added” are often used interchangeably. A value-added model (VAM) is only one of several types of models that measure student growth. Connecticut’s approach is indeed a growth model but it is not a value-added model; neither are targets adjusted nor are growth results evaluated using some expectation of student achievement that is based on student characteristics or demographics. Connecticut’s model does not set different targets for different students. All students at a prior achievement range have the same growth expectation.
Unlike in a value-added model, there is no arcane, statistical calculation that is done to quantify the effects of teachers, leaders, schools or districts on student growth. Under Connecticut’s model, the calculations are transparent. Anyone with authorized access to student test scores from year 1 and year 2 can determine if those students achieved their target, and how much of the target they achieved.

**Conclusion**

To summarize, Connecticut’s model is:

- **Criterion-referenced** because there is an objective, fixed growth target for each student. A student’s growth measure does not depend on how other students achieved or grew.
- **Continuous** because all growth counts; there are no “golden bands.” It is not a value table or a categorical growth model where only movement from one category or level to another is rewarded. There is no incentive in this system to focus on getting a small group of students over some preset proficiency bar; instead the message here is that all growth achieved by all students counts.
- **Familiar** because it uses an approach similar to that used with the CMT
- **Transparent** because local districts and schools can replicate the results; there are no “black-box” adjustments to the growth results.
- **Collaborative** because the transparency allows for conversation and reflection among educators.
- **Fair** because it excludes “partial-year” students; only those students who were enrolled in the same district or school on October 1st and at the time of testing are included in the calculations.
- **Achievable** because it is based on the actual growth achieved by Connecticut students.
- **Ambitious** because the model encourages growth above target.

**Reference**