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

CSDE

CONNECTICUT STATE DEPARTMENT OF EDUCATION

November 2016

Contents

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**.

Lowest, Highest, and Achievement Level Cut Scores for Smarter Balanced ELA and Math

Grade	Subject	Lowest scale score	SS Cut between Levels 1 and 2	SS Cut between Levels 2 and 3	SS Cut between Levels 3 and 4	Highest scale score		Subject	Lowest scale score	SS Cut between Levels 1 and 2	SS Cut between Levels 2 and 3	SS Cut between Levels 3 and 4	Highest scale score
3	ELA	2114	2367	2432	2490	2623		Math	2189	2381	2436	2501	2621
4	ELA	2131	2416	2473	2533	2663		Math	2204	2411	2485	2549	2659
5	ELA	2201	2442	2502	2582	2701		Math	2219	2455	2528	2579	2700
6	ELA	2210	2457	2531	2618	2724		Math	2235	2473	2552	2610	2748
7	ELA	2258	2479	2552	2649	2745		Math	2250	2484	2567	2635	2778
8	ELA	2288	2487	2567	2668	2769		Math	2265	2504	2586	2653	2802

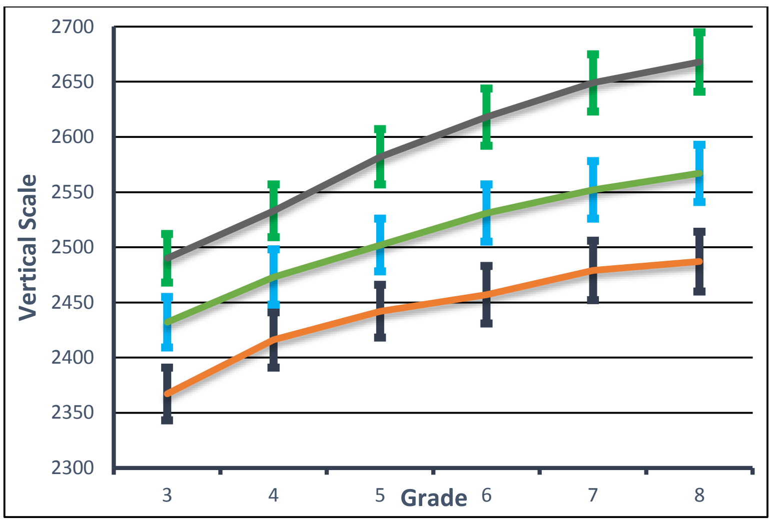
Determining the Growth Categories

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.

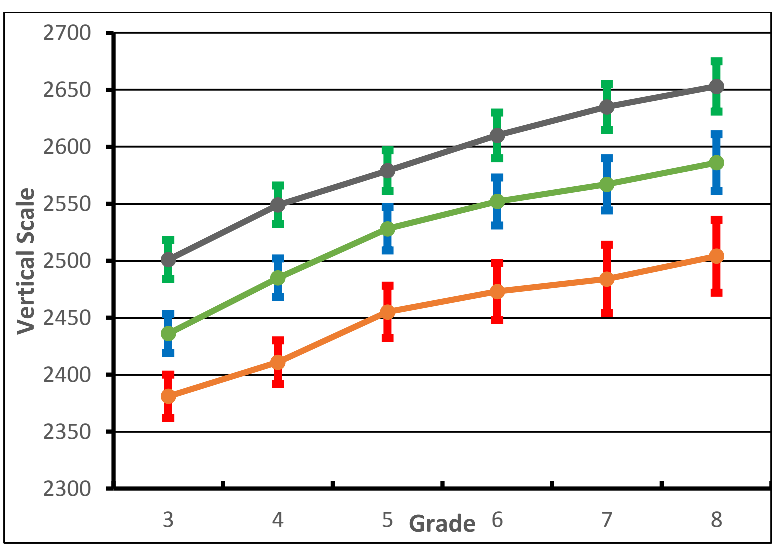
The amount of growth achieved by students performing at different points on the vertical scale can vary greatly. Generally, students at higher levels of achievement show smaller amounts of growth. Therefore, the CSDE utilized an approach to divide each of the nationally established Smarter Balanced achievement 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 generally calculated as

$$SE\ of\ the\ growth\ amount = \sqrt{SE(Score_{year1})^2 + SE(Score_{year2})^2 - covariance(Score_{year1}, Score_{year2})}$$

To help inform decisions about the growth targets, the CSDE adopted a conservative approach and excluded the covariance from this calculation. Therefore, when comparing the empirical growth amounts achieved from 2014-15 to 2015-16, the CSDE was able to compare that growth to the maximum possible standard error across the two assessments.

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

Grade in Yr. 1	Level	Level 1: Not Met		Level 2: Approaching		Level 3: Met		Level 4: Exceeded	
		1 - LOW	2 - HIGH	3 - LOW	4 - HIGH	5 – LOW	6 - HIGH	7 - LOW	8 - HIGH
3	Range	2114-2330	2331-2366	2367-2399	2400-2431	2432-2460	2461-2489	2490-2522	2523+
	Target	82	71	70	69	68	64	60	45/maintain
4	Range	2131-2378	2379-2415	2416-2444	2445-2472	2473-2502	2503-2532	2533-2568	2569+
	Target	82	69	69	64	58	55	49	34/maintain
5	Range	2201-2405	2406-2441	2442-2471	2472-2501	2502-2541	2542-2581	2582-2619	2620+
	Target	69	56	55	48	43	39	30	16/maintain
6	Range	2210-2417	2418-2456	2457-2493	2494-2530	2531-2574	2575-2617	2618-2656	2657+
	Target	73	58	53	47	44	38	33	21/maintain
7	Range	2258-2438	2439-2478	2479-2515	2516-2551	2552-2600	2601-2648	2649-2687	2688+
	Target	69	50	49	44	40	31	20	12/maintain
8	Range	2288-2446	2447-2486	2487-2526	2527-2566	2567-2617	2618-2667	2668-2703	2709+

Math Achievement Level Ranges and Growth Targets

Grade in Yr. 1	Level	Level 1: Not Met		Level 2: Approaching		Level 3: Met		Level 4: Exceeded	
		1 - LOW	2 - HIGH	3 - LOW	4 - HIGH	5 - LOW	6 - HIGH	7 - LOW	8 - HIGH
3	Range	2189-2351	2352-2380	2381-2408	2409-2435	2436-2468	2469-2500	2501-2526	2527+
	Target	77	61	59	60	59	57	56	47/maintain
4	Range	2204-2381	2382-2410	2411-2447	2448-2484	2485-2516	2517-2548	2549-2574	2575+
	Target	51	38	40	44	46	47	43	37/maintain
5	Range	2219-2419	2420-2454	2455-2491	2492-2527	2528-2553	2554-2578	2579-2605	2606+
	Target	43	46	45	44	42	41	41	44/maintain
6	Range	2235-2434	2435-2472	2473-2512	2513-2551	2552-2580	2581-2609	2610-2639	2640+
	Target	49	41	38	36	36	36	38	31/maintain
7	Range	2250-2438	2439-2483	2484-2525	2526-2566	2567-2600	2601-2634	2635-2664	2665+
	Target	58	35	31	31	36	37	38	35/maintain
8	Range	2265-2455	2457-2503	2504-2544	2545-2585	2586-2619	2620-2652	2653-2685	2686+

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)

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:

$$PTA = \frac{Growth\ Amount}{Growth\ Target}$$

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

- Castellano, K. E., & Ho, A. D. (2013). A Practitioner's Guide to Growth Models. Council of Chief State School Officers.
- Thompson, T. D. (2008, December). Growth, precision, and CAT: An examination of gain score conditional SEM. In annual meeting of the National Council on Measurement in Education, New York, NY.