Connecticut's Growth Model for the

English Language Proficiency Assessments

November 2017

Contents

Introduction

This paper describes the development of Connecticut's Growth Model for the English Language Proficiency assessment. It applies to students in grades 1 through 12. 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 separate component within Indicator 2 (i.e., Academic Growth) of Connecticut's Next Generation Accountability System for districts and schools. The approach used to develop this growth model is very similar to the one used to develop the growth model for the Smarter Balanced assessments.

Achievement versus Growth

Before diving deeper into the growth model, let's first understand the differences between achievement and growth. Here are simple definitions.

- **Achievement** or proficiency or status is a one-time snapshot measurement of a student's academic performance in a subject area (e.g., reading, listening, speaking). 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 percentage of students meeting mastery in Grade 4 in the district has increased from 30% in one year to 33% 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 second 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 average reading vertical scale scores on the ELP assessment 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, more of the students will be the same across years. However, students who exited English learner (EL) status in third grade would be excluded from the fourth grade ELP assessment, thus exacerbating the differences among the two groups.

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 based on a vertical scale allows us to quantify the amount of growth achieved by the same students from one grade, to the next – 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 LAS Links Vertical Scale

Connecticut uses the LAS Links assessments to measure English language proficiency. These assessments have been used in Connecticut for over a decade. In 2013-14, Connecticut transitioned from using forms A/B to forms C/D. Unlike forms A/B which were more life skills oriented, the newer forms C/D measure an English learner's ability to interact with academic context-based vocabulary.

Forms C and D are offered in each of the following five grade bands: K-1, 2-3, 4-5, 6-8, and 9-12. Within each band and form, there is a separate test booklet for each of the following four subject areas or domains: reading, writing, listening, and speaking. For example, in the grade 2-3 band, there is a Form C for Reading and a Form D for Reading. The Form C and D tests within a grade band are parallel forms; moreover, results from both forms in all grade bands and in all domains are reported on domain-level vertical scales. See table below which presents the lowest obtainable scale score (LOSS) and the highest obtainable scale score (HOSS) for each domain scale.

Vertical Scale Score Ranges for Each Domain

Part 1: Domain-Level Scales								
Grades	**Speaking**		**Listening**		**Reading**		**Writing**	
	LOSS	**HOSS**	**LOSS**	**HOSS**	**LOSS**	**HOSS**	**LOSS**	**HOSS**
K – 1	300	580	300	530	240	550	200	630
2 – 3	350	600	310	560	300	610	270	640
4 – 5	360	635	350	640	360	680	290	680
6 – 8	365	645	360	680	380	690	300	710
9 – 12	370	650	370	730	390	715	310	720

These vertical scales allow for the evaluation of growth through the measurement of change in vertical scale score across forms and across grades.

In addition to domain-level vertical scales, five composite scores are calculated "as the unweighted average of the student's scale scores from corresponding domains, and the results are then truncated to the integer part for reporting purposes" (CTB, 2013). The following two tables present the domain scales that are included in each of the five composite scores.

Correspondence between Domain and Composite Scales

Composite Scales	Base Scales			
	Listening	Speaking	Reading	Writing
Overall	X	X	X	X
Oral	X	X		
Comprehension	X		X	
Literacy			X	X
Productive		X		X

This table presents the LOSS and HOSS for those composite scales across the grade bands.

Vertical Scale Score Ranges for the Composite Scales

Part 2: Composite Scales

Grades	Overall		Oral		Comprehension		Literacy		Productive	
	LOSS	HOSS	LOSS	HOSS	LOSS	HOSS	LOSS	HOSS	LOSS	HOSS
K – 1	260	572	300	555	270	540	220	590	250	605
2 – 3	307	602	330	580	305	585	285	625	310	620
4 – 5	340	658	355	637	355	660	325	680	325	657
6 – 8	351	681	362	662	370	685	340	700	332	677
9 – 12	360	703	370	690	380	722	350	717	340	685

Note. LOSS = Lowest Obtainable Scale Score; HOSS = Highest Obtainable Scale Score

Within each domain and composite area, the scales are categorized into five achievement levels with Levels 4 and 5 representing the desired levels of achievement. These five levels are established through four cut scores. These cut scores were set originally in 2005 and revalidated in 2013 after the deployment of forms C and D. The cut scores for each of the domain scales as well as the five composite scores are presented on the following two pages.

Form C/D Cut Scores for Composite Vertical Scales

Note. SK=Speaking. LI=Listening. RD=Reading. WR=Writing. OV=Overall.
OR=Oral. CO=Comprehension. LT=Literacy. PR=Productive.

Proficiency Level Mastery	SK	LI	RD	WR	OV	OR	CO	LT	PR
Kindergarten									
1	430	426	355	347	389	430	386	351	388
2	461	444	381	417	425	461	412	399	439
3	496	470	421	488	468	487	449	454	492
4	550	520	475	516	515	526	467	495	533
Grade 1									
1	432	432	360	355	394	432	390	357	393
2	462	450	385	435	433	463	416	410	448
3	496	476	423	489	471	490	452	456	492
4	551	521	479	535	521	530	486	507	543
Grade 2									
1	443	442	435	425	436	443	448	430	434
2	473	462	472	475	470	470	473	473	474
3	509	492	499	504	501	495	495	501	506
4	557	536	547	544	546	540	531	545	550
Grade 3									
1	443	447	436	428	438	444	452	432	435
2	474	468	474	484	475	471	482	479	479
3	509	504	504	529	511	505	500	516	519
4	558	546	549	560	553	548	533	554	559
Grade 4									
1	449	457	468	434	452	450	485	451	441
2	475	484	504	498	490	478	506	501	486
3	510	525	535	533	525	514	526	534	521
4	559	581	588	584	578	575	563	586	571
Grade 5									
1	449	458	470	435	453	452	491	452	442
2	475	490	505	499	492	485	509	502	487
3	511	528	536	538	528	516	531	537	524
4	559	584	590	585	579	580	573	587	572
Grade 6									
1	451	462	501	447	465	455	499	474	449
2	476	489	529	498	498	481	514	513	487
3	512	532	559	548	537	518	540	553	530
4	560	586	608	591	586	575	574	599	575
Grade 7									
1	451	463	502	447	465	460	500	474	449
2	477	492	530	498	499	485	517	514	487
3	513	533	560	548	538	521	546	554	530
4	560	588	608	592	587	580	576	600	576

Connecticut's Growth Model for the English Language Proficiency Assessments

Proficiency Level Mastery	SK	LI	RD	WR	OV	OR	CO	LT	PR
Grade 8									
1	451	467	502	448	467	465	501	475	449
2	477	498	532	499	501	492	519	515	488
3	514	535	561	548	539	525	553	554	531
4	560	590	608	593	587	582	579	600	576
Grade 9									
1	452	471	507	449	469	465	512	478	450
2	478	509	545	500	508	490	534	522	489
3	515	546	581	549	547	525	567	565	532
4	560	625	632	594	602	561	597	613	577
Grade 10									
1	452	475	508	449	471	468	514	478	450
2	478	511	546	500	508	495	536	523	489
3	516	550	582	549	549	527	569	565	532
4	560	627	633	594	603	566	606	613	577
Grade 11									
1	452	480	509	450	472	471	515	479	451
2	479	516	547	501	510	497	540	524	490
3	517	555	583	550	551	530	575	566	533
4	560	630	634	595	604	567	608	614	577
Grade 12									
1	452	481	510	451	473	472	515	480	451
2	479	517	548	502	511	500	540	525	490
3	518	560	584	550	553	531	575	567	534
4	560	633	635	596	606	569	610	615	578

Note. SK=Speaking. LI=Listening. RD=Reading. WR=Writing. OV=Overall. OR=Oral. CO=Comprehension. LT=Literacy. PR=Productive.

What to Measure for Growth

The Connecticut State Department of Education (CSDE) began by planning for a growth model based on the overall composite score. However, it became clear early in the process that if only the overall composite was valued, then students may be able to compensate for low performance in one or more areas with high performance in others. The CSDE then studied the feasibility of measuring growth in each of the four domains of reading, writing, listening, and speaking. However, there were some limitations with this approach: the domain-level tests were comparatively short i.e., around 20 score points for listening, around 30 each for reading and writing, and around 40 for speaking. This results in lower precision and greater standard error of measurement (SEM). In addition, these tests had few

score points within levels 2 and 3; in some instances, listening had as few as two score points in a level while reading had only four. Therefore, the Oral and Literacy composites were considered.

The Oral composite score is an unweighted average of the listening and speaking domains. The Literacy composite score is an unweighted average of the reading and writing domains. These composite scores are inherently more reliable than the individual domain-level scores because they are based on student performance on a larger number of items. There are 46-49 items per composite area in Kindergarten and around 60 items per composite area in all other grades. Also, if growth in these two composite scores could be evaluated separately, then we will send the clear message that ELs need to grow in both their oral and literacy skills and cannot compensate one with higher achievement in the other.

Establishing Ambitious Yet Achievable Growth Targets

The amount of growth achieved by students performing at different points on a vertical scale can vary greatly. Generally, students at higher levels of achievement show smaller amounts of growth. Therefore, the observed achievement of Connecticut's ELs from 2013-14, 2014-15, and 2015-16 ELP assessments (i.e., a total of 48,432 matched scores) were evaluated to determine the growth that was evidenced at each starting achievement level in both the oral and literacy composite areas. The CSDE decided not to subdivide the achievement levels into two halves (as was the case with the Smarter Balanced growth model) because of the limited number of score points in some of the domain level tests.

The first step was to determine the actual amount of growth achieved by Connecticut students from 2013-14 to 2014-15 and from 2014-15 to 2015-16 within each of the five achievement levels in each grade (1-12). Each student was assigned to one of the five levels based on the student's year1 score. For each student, the growth amount was calculated by subtracting the year1 score from the year2 score.

$$growth\ amount = Score_{year2} - Score_{year}$$

The percentiles of the growth amounts within each of the five achievement levels 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 percentiles). The selected growth amounts were also translated into trajectories to see if students at different starting achievement levels reached higher levels in future years if they achieved their respective targets in each grade. The maximum number of years to English language proficiency was set at five years.

After review of the data, the growth targets for Oral were set at the 60th percentile of observed growth. They were sufficiently robust to ensure that any student, regardless of their starting achievement level, would reach Level 4 in Oral in five years or less. It was also determined that for students who had reached an Oral achievement level of 4 or 5, their growth expectation was to remain in levels 4 or 5 in the subsequent year.

The growth targets for Literacy were set at the 70th percentile of observed growth with two exceptions:

- Targets for students in first grade, whose Literacy achievement level in kindergarten was either a Level 2 or 3, were based on the observed 60th percentile of growth. These slightly lower targets still ensured that ELs would attain Literacy proficiency within the five-year timeline.

- Targets for ninth grade students whose Literacy achievement was either level 2 or 3 in eighth grade are the same growth amount as the observed 70th percentile of growth from ninth to tenth grade. These more ambitious targets were necessary to overcome the observed decline in ninth grade Literacy scores.

The Literacy targets were also sufficiently robust to ensure that any student, regardless of their starting achievement level, would reach Level 4 in five years or less. As with Oral, it was determined that for students who had reached a Literacy achievement level of 4 or 5, their growth expectation was to remain in levels 4 or 5 in the subsequent year.

These growth targets were also greater than the combined estimated SEM. Since SEMs were only available for the domain scores (e.g., Reading, Listening) and not the composite scores (e.g., Oral, Literacy), the CSDE estimated composite SEMs at several points along the vertical scale for each grade. These combined SEMs were estimated at the LOSS, half-way in level 1, the level 1/2 cut, and the level 2/3 cut).

The combined SEMs were computed as follows:

1. Determine the test information from each domain level test at a particular point on the vertical scale.

$$\frac{1}{SEM^2}$$

2. Combine the test information from the two domains at that particular point on the vertical scale

$$\frac{1}{SEM_1^2} + \frac{1}{SEM_2^2}$$

3. Compute the SEM for the composite as the inverse of the combined test information

$$\frac{1}{\sqrt{\dfrac{1}{SEM_1^2} + \dfrac{1}{SEM_2^2}}}$$

When calculating growth, we are comparing composite 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

$$SEM \ of \ the \ growth \ amount = \sqrt{SEM(Score_{year})^2 + SEM(Score_{year2})^2}$$

This standard error of the growth amount was taken into account to help inform decisions about the growth targets. Though percentiles were used to determine the ambitious yet achievable growth amounts, these target amounts will now serve as a fixed criterion for at least the next few growth cycles.

Final Growth Target Tables

The final growth target tables for Oral and Literacy are presented on the following page. Note that students who achieve the expected growth targets will reach the desired level of proficiency (Levels 4 or 5) in five years or less. The growth trajectory for a hypothetical kindergarten student with an oral score of 335, who subsequently achieves the expected targets, is presented below. As is evident, this student reaches Level 4 in four years.

ELP Growth Table (ORAL)						
Grade in Yr. 1		Level 1	Level 2	Level 3	Level 4	Level 5
K	Range	300-429	430-460	461-486	487-525	526-555
	Target	61	36	28	Maintain	Maintain
1	Range	300-431	432-462	463-489	490-529	530-555
	Target	68	33	25	Maintain	Maintain
2	Range	330-442	443-469	470-494	495-539	540-580
	Target	59	32	26	Maintain	Maintain
3	Range	330-443	444-470	471-504	505-547	548-580
	Target	79	32	28	Maintain	Maintain
4	Range	355-449	450-477	478-513	514-574	575-637
	Target	82	44	32	Maintain	Maintain
5	Range	355-451	452-484	485-515	516-579	580-637
	Target	76	38	25	Maintain	Maintain
6	Range	362-454	455-480	481-517	518-574	575-662
	Target	78	46	31	Maintain	Maintain
7	Range	362-459	460-484	485-520	521-579	580-662
	Target	73	43	32	Maintain	Maintain
8	Range	362-464	465-491	492-524	525-581	582-662
	Target	60	30	14	Maintain	Maintain
9	Range	370-464	465-489	490-524	525-560	561-690
	Target	60	28	22	Maintain	Maintain
10	Range	370-467	468-494	495-526	527-565	566-690
	Target	63	33	25	Maintain	Maintain
11	Range	370-470	471-496	497-529	530-566	567-690
	Target	53	31	24	Maintain	Maintain
12	Range	370-471	472-499	500-530	531-568	569-690

ELP Growth Table (LITERACY)						
Grade in Yr. 1		Level 1	Level 2	Level 3	Level 4	Level 5
K	Range	220-350	351-398	399-453	454-494	495-590
	Target	111	75	68	Maintain	Maintain
1	Range	220-356	357-409	410-455	456-506	507-590
	Target	96	68	61	Maintain	Maintain
2	Range	285-429	430-472	473-500	501-544	545-625
	Target	75	49	40	Maintain	Maintain
3	Range	285-431	432-478	479-515	516-553	554-625
	Target	80	51	44	Maintain	Maintain
4	Range	325-450	451-500	501-533	534-585	586-680
	Target	87	53	40	Maintain	Maintain
5	Range	325-451	452-501	502-536	537-586	587-680
	Target	82	46	31	Maintain	Maintain
6	Range	340-473	474-512	513-552	553-598	599-700
	Target	80	47	37	Maintain	Maintain
7	Range	340-473	474-513	514-553	554-599	600-700
	Target	75	48	39	Maintain	Maintain
8	Range	340-474	475-514	515-553	554-599	600-700
	Target	59	38	26	Maintain	Maintain
9	Range	350-477	478-521	522-564	565-612	613-717
	Target	63	38	26	Maintain	Maintain
10	Range	350-477	478-522	523-564	565-612	613-717
	Target	60	36	25	Maintain	Maintain
11	Range	350-478	479-523	524-565	566-613	614-717
	Target	60	33	23	Maintain	Maintain
12	Range	350-479	480-524	525-566	567-614	615-717

Connecticut's Growth Model for the English Language Proficiency Assessments

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.

Outcome Measures

Using the growth target set for each student, 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:

$$PTA = \frac{Growth\ Amount}{Growth\ Target}\ x\ 100$$

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 and districts. This results in two measures: growth rate and average PTA. The growth rate is the percentage of students who met their targets, 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 does 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 Smarter Balanced.
- **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.

- CTB/McGraw-Hill (2013). LAS Links Second Edition, Forms C and D, Technical Manual, Monterey, CA.

- 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.