Measures of Student Progress for STAAR Assessments

Over the past several years, the Texas legislature has required a method for measuring annual improvement in student achievement. These requirements include House Bill 1 (in 2005), Senate Bill 1031 (in 2007), and House Bill 3 (in 2009). As Texas moves to the State of Texas Assessments of Academic Readiness (STAAR) decisions will need to be made about method(s) best suited to measuring progress and annual improvement in student achievement under the new program. The purpose of this document is to describe the features of the growth models that have been approved by the United States Department of Education (USDE) and are currently in use.

This paper describes each of the three types of progress measures under consideration. There are brief descriptions of each of the models followed by a table summarizing their features. The three models that are described represent the general types of growth models approved by the USDE for states to use in Adequate Yearly Progress calculations. There are many variations on these general types of models that can be implemented in practice.

- **Student Growth Percentile (SGP) Models—**these models use statistical regression and provide two measures: a measure that compares each student's growth with that of his or her academic peers (i.e., growth percentiles) and a projection value that indicates whether the student is on track to meet a future proficiency goal. SGP models are statistically complex and lack transparency; it is not possible for stakeholders to replicate results. An advantage of growth percentile models is that they provide output in a form that is intuitive to stakeholders. The normative growth data, or student growth percentile information, can be likened to the growth percentiles that parents receive on their children. For example, a parent may be told that her son is in the 80th percentile for height or that he is taller than 80 percent of similarly aged males. Furthermore it is possible to say that other boys at the 80th percentile grow to be just under 6 foot at maturity. All parents have experienced these charts, so it is familiar to look at the "growth charts" and make sense of them. A disadvantage of growth percentile models is that they are computationally burdensome and require a lot of time and resources to produce results. The Colorado Growth Model is an example of a growth percentile model. Sample output from the Colorado Growth Model is presented in Figure 1.

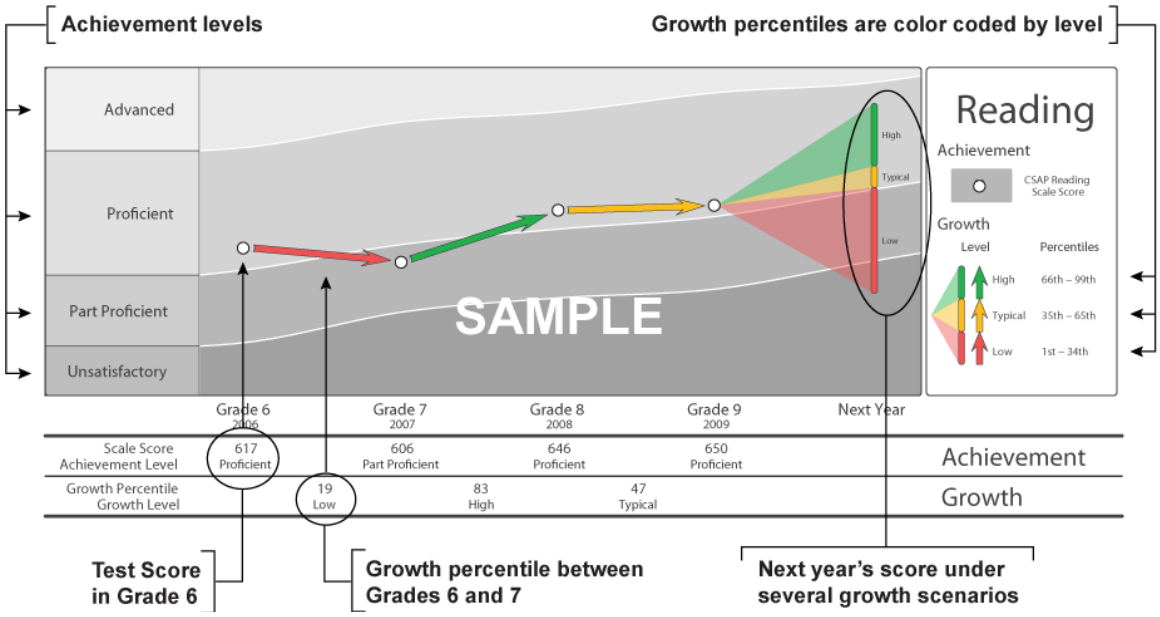

Figure1. Colorado Growth Model Student Report Example

- **Growth to Proficiency Models**—these models provide yearly growth targets for students based on the distance between their current performance and a future proficiency goal. Students' growth is then evaluated based on a comparison of their observed performance to their target performance.

  Calculations for these models are straightforward and fully transparent, which is one of their advantages. A disadvantage of growth to proficiency models is that they are not as accurate as SGP models. The definition of expected growth is based on logic and not on empirical data. For example, if a student needs to gain 60 points in three years to reach the future goal, the growth target each year is set to the logical 20 points. The 20-point yearly target is not based on student data or previous performance; instead, it is based on the logical argument that increasing the score by 20 points each year will result in the student meeting the future goal in three years. These models fit well with instruction purposes because they are tailored to the individual student, are easy to understand, and the growth is based only on the subject of interest. Figure 2 depicts a hypothetical growth to proficiency model

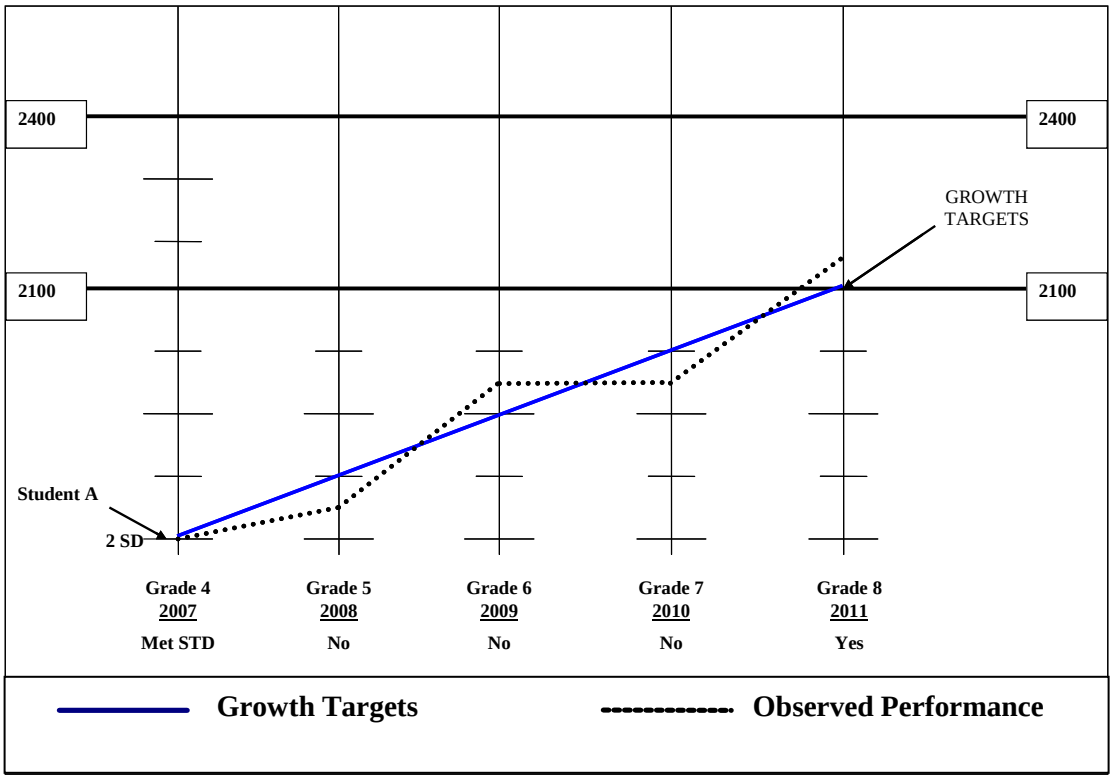

Figure 2. Growth to Proficiency Model Example

- **Value/Transition Tables—**these models evaluate student growth by first subdividing performance standards into subcategories and then setting growth targets based on the number of subcategories students must traverse to meet a future goal. Students are classified based on their transition across subcategories. Value tables subdivide performance standards into subcategories based on scale score changes. Transition tables subdivide performance standards into descriptive rather than numerical subcategories.

  These models are very similar to growth to proficiency models, differing primarily in the method of setting targets and presentation of results. Calculations for these models are straightforward and fully transparent, which is one of their advantages. A disadvantage of these models is that they are not as accurate as prediction models. These models fit well with instruction purposes because they are tailored to the individual student. Figures 3 and 4 present value and transition table examples respectively.

Year One	Year Two							
	Unsatisfactory Academic Performance		Satisfactory Academic Performance			Advanced Academic Performance		
	Low	High	Low	Mid	High	Low	Mid	High
Low Unsatisfactory Academic Performance	38	43	48	53	58	63	68	73
High Unsatisfactory Academic Performance	35	40	45	50	55	60	65	70
Low Satisfactory	33	38	43	48	53	58	63	68
Mid Satisfactory	30	35	40	45	50	55	60	65
High Satisfactory	28	33	38	43	48	53	58	63
Low Advanced	18	23	28	33	38	43	48	53
High Advanced	15	20	25	30	35	40	45	50

Figure 3. Value Table Example: Numbers in the Table Represent Vertical Scale Score Changes across Two Years

Year One	Year Two						
	Low Unsatisfactory Academic Performance	High Unsatisfactory Academic Performance	Low Satisfactory Academic Performance	Mid Satisfactory Academic Performance	High Satisfactory Academic Performance	Low Advanced Academic Performance	High Advanced Academic Performance
Low Unsatisfactory	Maintained	Slightly Improved	Slightly Improved	Improved	Improved	Significantly Improved	Significantly Improved
High Unsatisfactory	Slightly Regressed	Maintained	Slightly Improved	Slightly Improved	Improved	Significantly Improved	Significantly Improved
Low Satisfactory	Slightly Regressed	Slightly Regressed	Maintained	Slightly Improved	Slightly Improved	Significantly Improved	Significantly Improved
Mid Satisfactory	Regressed	Slightly Regressed	Slightly Regressed	Maintained	Slightly Improved	Significantly Improved	Significantly Improved
High Satisfactory	Regressed	Regressed	Slightly Regressed	Slightly Regressed	Maintained	Improved	Significantly Improved
Low Advanced	Significantly Regressed	Significantly Regressed	Significantly Regressed	Regressed	Regressed	Maintained	Slightly Improved
High Advanced	Significantly Regressed	Significantly Regressed	Significantly Regressed	Significantly Regressed	Regressed	Slightly Regressed	Maintained

Figure 4. Transition Table Example

Features of Growth to Proficiency, Value Table/Transition Table, and Student Growth Percentile Progress Measures

	Student Progress Measures		
	Growth to Proficiency	**Value Tables / Transition Tables**	**Student Growth Percentiles (SGP)**
States Using Model	AL, AZ, AR, FL, MO, NC	DE, IA, MN, MI, TX 1%	CO, GA, HI, ID, IN, MA, NH, NJ, NV, NY, RI, VA, WV
Meets Legislative Requirements	State: Yes Federal: Yes	State: Yes Federal: Yes	State: Yes Federal: Yes
Transparency	▪ Easy to understand ▪ Easy to compute ▪ Easily replicated by stakeholders	▪ Easy to understand ▪ Easy to compute ▪ Easily replicated by stakeholders	▪ SGP are easy to understand ▪ Regression is difficult to understand ▪ Very difficult to compute ▪ Stakeholder cannot replicate
Underlying Assumptions	▪ Individual student growth rates remain constant across years	▪ Individual student growth rates remain constant across years	▪ Individual student growth rates remain constant across years ▪ The growth rates of one cohort of students generalizes to different cohorts of students
Strengths	▪ Transparency ▪ Adapts well to instruction ▪ Can be used with vertical and horizontal scales	▪ Transparency ▪ Adapts well to instruction ▪ Can be used with vertical and horizontal scales	▪ Incorporates desirable trajectory and prediction model elements ▪ Growth percentiles are an intuitive measure to stakeholders ▪ Can be used with vertical and horizontal scales
Weaknesses	▪ Not as accurate as prediction models ▪ Expect the same amount of growth each year ▪ Not based on empirical data	▪ Not as accurate as prediction models ▪ Not based on empirical data ▪ Some states do not provide student-level reports based on these models	▪ Computationally intensive and time consuming
Types of Outcome Measures Available	▪ Change scores ▪ Yes/No "Met Growth" predictions	▪ Descriptions of growth trajectories ▪ Yes/No "Met Growth" predictions ▪ Value score	▪ Growth percentiles ▪ Yes/No "Met Growth" predictions
Reporting Considerations	▪ Growth targets can be established in 2012 ▪ Growth results can be reported on time from 2013 forward	▪ Growth targets can be established in 2012 ▪ Growth results can be reported on time from 2013 forward	▪ First growth results can be late reported in 2013 ▪ Growth results must be late reported from 2014 forward

Timeline for Evaluating Measures of Student Progress for STAAR

	Steps	Timeline
1)	Identify the models to evaluate as part of the measures of student progress research study for the STAAR program	June 2011
2)	Discuss required elements of the measures of student progress research study with TEA	July 2011
3)	Review legislative requirements	July 2011
4)	Discuss the measures of student progress research study plan and possible analyses with the TTAC	Aug 2011
	a) Conduct growth to proficiency analyses	Sept 2011
	b) Conduct transition table analyses	Oct 2011
	c) Conduct student growth percentile model analyses	Nov-Dec 2011
	d) Summarize outcomes across models	Dec 2011–Jan 2012
5)	Consider having multiple indicators of student progress and which indicators to include in addition to the traditional growth model	Sept 2011
6)	Discuss application of student progress measures to STAAR Alternate and STAAR Modified	Nov 2011
7)	Present summary of the measures of student progress study findings	Jan 2012
8)	Discuss application of student progress measures to STAAR L	Feb 2012
9)	2012 Accountability Workbook to USDE	Jan–Feb 2012
10)	Present measures of student progress study findings to TTAC	Mar 2012
11)	Discuss application of progress measures to evaluating educator effectiveness	May 2012
12)	Initial draft of the progress measures **preliminary** report	June 2012
13)	Obtain educator and expert input (DAC meeting)	June 2012
14)	Approval of wave 1 (spring 2013) measures of student progress	July 2012
15)	Final draft of the progress measures **preliminary** report	July 2012
16)	Update the progress measures **preliminary** report for use as **final** report	Sept 2012–May 2013
17)	2013 Accountability Workbook to USDE	Dec 2012–Feb 2013
18)	Obtain educator and expert input (e.g., DAC, TTAC, ELL focus groups, special education teacher review committees)	Jan–June 2013
19)	Implement and report wave 1 measures of student progress for the STAAR program	Spring 2013
20)	Initial draft of the progress measures final report	May 2013
21)	Submit progress measure plans to Peer Review (Phase II submission)	May 2013
22)	Reevaluate student progress measure plans based on two years of STAAR data	May–July 2013
23)	Approval of wave 2 (fall 2013 and spring 2014) measures of student progress	July 2013
24)	Final draft of the progress measures **final** report	Aug 2013
25)	2014 Accountability Workbook to USDE	Sept–Nov 2013 or Dec 2013–Feb 2014
26)	Implement and report wave 2 measures of student progress for the STAAR program	Fall 2013 Spring 2014
27)	Submit final progress measure decisions to Peer Review (Phase III submission)	Dec 2013