California Department of Education

Executive Office

SBE-002 (REV. 11/2017)

memo-pptb-amard-jun18item01

# **MEMORANDUM**

**DATE:** June 20, 2018

**TO:** MEMBERS, State Board of Education

**FROM:** TOM TORLAKSON, State Superintendent of Public Instruction

**SUBJECT:** Update on the Development of a Growth Model: Incorporating a Growth Model into the California Accountability System.

## Summary of Key Issues

In 2016, the State Board of Education (SBE) and the California Department of Education (CDE) began work on the possible inclusion of a student-level growth model in California’s new accountability system and the California School Dashboard (Dashboard). Working closely with the Technical Design Group (TDG) and staff at the Educational Testing Service (ETS), the CDE explored three different student growth models:

* Change in Distance to Met (CDTM)
* Conditional Percentile Rank of Gain (CPR), and
* Residual Gain (RG).

Based on the data simulations conducted on all three models, the SBE directed the CDE, at the May 2018 SBE meeting, to focus on the RG model to address the following questions:

1. What information does the RG model provide compared to the information that the current Distance from Level 3 (DF3) provides?
2. How would the RG model fit into the current accountability five-by five structure?

The CDE has since engaged several stakeholder groups in discussions on the RG model and sought their feedback on the usefulness of the model and its incorporation in the Dashboard.

## Summary of Previous State Board of Education Discussion and Action

At the May 2018 SBE meeting, CDE presented the three student-level growth model options (cited above) for consideration (<https://www.cde.ca.gov/be/ag/ag/yr18/documents/may18item02.docx>).

The SBE approved further exploration of the RG mode and requested additional information on the model (as detailed in the previous section) the July and September 2018 SBE meetings.

In March 2018, the SBE reviewed proposed revisions for the 2018 Dashboard, including the inclusion of a student-level growth model (<https://www.cde.ca.gov/be/ag/ag/yr18/documents/mar18item01.docx>).

In February 2018, the SBE received simulation results, conducted by ETS, for three potential student-level growth models (<https://www.cde.ca.gov/be/pn/im/documents/memo-pptb-amard-feb18item01.docx>).

In June 2017, the SBE received an update on the continued work toward a growth model, including three potential student growth models to be considered for simulations (<https://www.cde.ca.gov/be/pn/im/documents/memo-asb-adad-jun17item03.doc>).

In January 2017, the SBE discussed criteria for selecting a growth model to be used for school and district accountability (<https://www.cde.ca.gov/be/ag/ag/yr17/documents/jan17item02.doc>).

In December 2017, the SBE received an update on the ongoing development of the growth model
(<https://www.cde.ca.gov/be/pn/im/documents/memo-pptb-amard-dec17item03.docx>).

In June 2016, the CDE provided a progress update and clarified key issues related to the design of a school- and district-level accountability model, as opposed to reporting individual student-level growth and performance (<https://www.cde.ca.gov/be/pn/im/documents/memo-dsib-amard-jun16item01.doc>).

In February 2016, the SBE received an overview of student-level growth models that could be used to communicate Smarter Balanced Summative Assessment results (<https://www.cde.ca.gov/be/pn/im/documents/memo-dsib-amard-feb16item01.doc>).

### Comparing Residual Gain to Distance from Level 3

#### What Does the RG Model Tell Us?

RG is a relative measure of growth that compares a student’s actual score to a predicted score. The calculation of an expected score is created using a student’s prior year scores on the Smarter Balanced Summative Assessments for English language arts/literacy (ELA) and mathematics, as well as the scores of all other students in the same grade. The RG score is the difference between a student’s observed score and his/her expected score. A positive score indicates that a student met and/or exceeded his/her growth prediction.

The gain scores from all students are averaged to provide an overall RG score, which tells us (1) how many students met their predicted test score, and (2) how far above or below they were from their predicted score. It also tells us whether a school, LEA, or student group had high or low student growth in a relative comparison to all students.

#### What Doesn’t the RG Model Tell Us?

While the RG model allows us to compare schools, local educational agencies (LEAs), and student groups, it does not show how student growth compares to proficiency targets or changes in DF3; whether achievement gaps are closing; or if schools, LEAs, or student groups are on track to reach proficiency.

#### What does Distance from Level 3 (DF3) Tell Us?

The current Academic Indicator uses the DF3 methodology for measuring both academic status and progress. The status measure shows: (1) how much improvement is needed to bring the average student in the school, LEA, or student group to the Standard Met level (Level 3) in ELA or mathematics, or (2) the extent to which the average student exceeds Level 3. The change measure reflects the difference between current and prior year Status.

#### Comparing DF3 Change to RG Growth

Table 1 compares DF3 and RG results for student groups, using 2016 and 2017 CAASPP results. Data are presented as: (1) a numeric value of residual gain (shown as Average Academic Growth), and (2) a percent of students meeting their predicted target.

#### Table 1

| **Student Group** | **AverageAcademic Status (DF3)** | **AverageAcademic Change (DF3)** | **AverageAcademic Growth (RG)** | **Percent Meeting Predicated Target (RG)** |
| --- | --- | --- | --- | --- |
| All Students | -17 | -0.5 | 0.0 | 51.1% |
| Socioeconomically Disadvantaged | -45.9 | -1.3 | -2.3 | 49.9% |
| English Learners | -50.8 | -1.6 | -2.5 | 49.4% |
| White | 15.1 | -0.5 | 1.8 | 52.0% |
| African American | -60.9 | -1.9 | -4.7 | 47.6% |
| Homeless | -62.1 | -4.2 | -3.2 | 49.1% |

Significant differences in academic performance for 2017 can be observed among student groups, as shown in the second column (Average Academic Status—DF3). When comparing results to the prior year, using both the DF3 and RG methodologies (third and fourth columns respectively), there is a consistent downward trend among student groups with the exception of White students. For the latter group, application of the DF3 methodology results in negative Change (-0.5), while application of the RG methodology results in positive growth (+1.8).

The average academic growth for the “All Students” group was 0.0, which means that the average student performed as predicted, while 2017 Status was slightly lower for the group than that of 2016.

The last column shows the percentage of students who met their predicted target. Applying the RG model, the values are all clustered around 50 percent, which means that about half of all students, in all student groups, including the “All Students” group, met their predicted target.

### Incorporating Residual Gain into the Five-by-Five Model

Application of the RG model sometimes produces outcomes that are inconsistent with those obtained through the five-by-five model. An example is provided in Figure 1, which plots a school’s academic performance in ELA. In 2016, the school received a Status of +42.7 points, which means that the average student score was 42.7 points above the minimum threshold for Level 3 (Standard Met) on the Smarter Balanced Summative Assessments for ELA. In 2017, the school’s Status increased by 4.7 points, moving the school up in Status (Very High) and over in Change (Increased), and from a Green to Blue performance level (color).

However, if we were to apply the RG methodology, the school would show a decline in growth (i.e,-5.8). Although the school’s performance was higher in 2017, earning it a higher Status level, the school did not grow as much as predicted. This example demonstrates that the RG model does not always operate the same on the five-by-by grid as the other indicators.

#### Figure 1. Displaying Change and Growth on the Five-by-Five Grid for the ELA Academic Indicator



### Stakeholder Feedback on the use of the RG Growth Model

The CDE engaged multiple stakeholder groups in discussions on the student growth model in an effort to understand whether the RG model provided added value to the accountability system and whether the data it provided met an unmet need.

While stakeholders expressed a strong desire for data showing whether students were making growth from one year to the next, most did not find that the RG model produced actionable data. Concern was expressed that a relative measure does not provide grade-specific growth needed year over year. Stakeholders also found the RG model complex and difficult to communicate, specifically with respect to how schools, LEAs, and/or student groups could attain a higher Status from—one year to the next while, at the same time, receiving a negative score for growth.

### Technical Feedback on the Growth Model

As requested by the SBE, ETS provided an update to their February 2018 report on the growth models, using the most recent testing data from the 2017 CAASPP administration, “Updated Analysis of the Residual Gain Model 2017-18 Plan” (*Attachment 1*). The report found that, while the RG model performed statistically similar in both years, there was low year-to-year stability within the outcomes, which they described as a, “known point of concern and can compromise their use in evaluating schools, or LEAs, or holding them accountable for their students’ progress (e.g., Lash et al., 2016).” The report specifically cited that, “the year-to-year instability in the growth measures means that schools with low growth one year might have notably higher growth the next year and vice versa for schools with high growth. Such volatility can make it difficult for educational leaders to use the growth data for driving decisions, as decisions made one year might be contradicted with the next year’s growth data.” (*Attachment 1,* pg. 16) ETS also noted that cross-time (year-to-year) “stability is slightly higher at the school-by-grade or LEA-by-grade levels, but still not at levels that would likely be desirable for high-stakes decisions.” (*Attachment 1,* pg. 17)

In reflecting upon this feedback from the report, CDE thinks that the RG growth model could possibly be a useful tool for identifying outliers that consistently demonstrate high or low growth over multiple years. However, based on the data provided in the ETS memo, less than 10 percent of schools fall into this category (*Attachment 1*), leaving 90 percent of schools with data that may be difficult to interpret from year-to-year.

At the June 2018 Technical Design Group meeting, the CDE presented the report findings, along with feedback from stakeholders on the RG Model. Based on the ETS findings and the general dissatisfaction expressed by stakeholders for the RG model (e.g., it does not provide the information that the field is requesting), the TDG unanimously recommended to delay the inclusion of the RG model into the accountability system. They indicated that more years of data are needed to determine how to reduce the volatility of growth over time and suggested that CDE staff explore whether other growth models can provide the information that the field is requesting. In addition, the TDG strongly recommended against providing RG growth data for informational purposes due to the risk of misinterpretation and possible use of these volatile scores for high-stakes decisions.

#### Options

The CDE will need guidance from the SBE at the July 2018 SBE meeting on how to proceed with the possible incorporation of the RG growth model in the Dashboard. There are three possible options:

1. Incorporate the growth model in the 2018 Dashboard. This would require CDE staff to determine if refinements to the model are needed and to explore whether: (1) the current DF3 Change methodology should be replaced with RG scores, or (2) a new state indicator should be created.
2. Provide the RG growth data on the CDE web page for information only. This would also require CDE staff to determine if refinements to the model should be made and to explore options for displaying the data.
3. Delay the implementation of the RG growth model. Once more data is available, CDE staff would work with the TDG and ETS to explore if using more years of data would reduce volatility in the RG model or if another growth model might be a more appropriate option.

## Attachment(s)

* Attachment 1: Updated Analysis of the Residual Gain Model 2017–18 Plan
(18 Pages)