

Initial Observations on STAR 2011: Results: Blah Despite a Hidden Fudge Factor

D. J. McRae, Ph.D.
08/15/11

The Standardized Testing And Reporting (STAR) program results were released today. For the past several years, I have made a practice of recording some of my initial observations. Here is this year's version.

Overview for STAR California Standards Tests (CSTs) Results

In general, the 2011 results may be characterized as *blah*, or perhaps just *not very exciting*. We now have ten years of gain scores for the STAR California Standards Test (CST) program, and I have evolved to calculating an average annual gain statistic for the core E/LA and Math tests [grades 2-11 for E/LA and 2-7 for Math] for which annual gains are apples-to-apples comparisons. This average annual gain statistic looks somewhat like a GPA, with numbers in the 3.0 to 4.0 range amenable to a characterization of *good* to *very good*. The notion that average annual gains in the 3.0 to 4.0 percent range tend to be what “on track” large scale instructional and assessment programs yield has been noted by other educational measurement specialists and is quite consistent with my observations from 40 years of looking at K-12 assessment program data.

Using the data for CST percents *proficient and above* from the SPI press release (Tables 1 and 6), one may compute average annual gains for STAR results for the last ten years and assign “Grade Point Averages” as follows:

<u>Year</u>	<u>E/LA</u>	<u>Math</u>	<u>Average Gain</u>	<u>Grade</u>
2002	2.0	1.5	1.75	C-
2003	2.4	5.5	3.95	A
2004	0.3	1.2	0.75	D-
2005	4.5	5.1	4.80	A++
2006	1.8	3.2	2.50	C+
2007	1.4	0.5	0.95	D
2008	2.7	2.3	2.50	C+
2009	4.3	4.2	4.25	A
2010	2.6	2.3	2.45	C+
2011	1.6	2.5	2.05	C

[Note: Data for 2002 are found in earlier year SPI press releases.]

Given this 40,000 foot view of the STAR landscape, the 2011 results may be characterized as *very average*, clearly not in the same category as the very good results in 2003, 2005, and 2009, but rather closer to the poor results in 2002, 2004, and 2007.

Introduction of the California Modified Assessments (CMAs) – A Hidden Fudge Factor

Beginning in 2008, California phased in a new STAR test for selected Special Education students, a test designed to provide greater accessibility to the content of the STAR exams for these students. Special Education students who scored *far below basic* or *below basic* on a STAR CST at the previous grade level are eligible to take the CMA with assignment of tests to be administered determined by the students' IEP team. In 2008, about 40,000 Special Education students grades 3 to 5 took CMAs rather than CSTs. In 2009, CMAs for grades 6-8 were introduced and about 100,000 Special Education students grades 3 to 8 took CMAs rather than CSTs. In 2011, following a 2-year phase in for CMAs for grades 9-11, the 4-year gradual introduction of this test was completed with roughly 185,000 Special Education students taking CMAs rather than CSTs.

There are two issues to note regarding the introduction of CMAs to replace the more rigorous CSTs for selected Special Education students: (1) The CST percent *proficient and above* averages reported by the SPI press release are artificially inflated, and (2) the number of students taking the CMAs far exceed initial plans for this test, and are likely to increase to an alarming percentage of the entire group of Special Education students in California.

Inflated CST Percentages. When students who have not scored *proficient or above* are removed from the CST percent calculations, the numerators remain the same but the denominators decrease, thus artificially increasing the average *proficient and above* percentages – in effect, the STAR CST gain data reported by the SPI are artificially inflated by systematically eliminating scores for lower scoring students. This is what happens for the average percentages above for 2008 through 2011, providing a hidden fudge factor artificially inflating the data included in the SPI press release.

Now, the fine print in the press release includes a note that “With the inclusion of the CMA in the STAR program, caution may be needed when interpreting STAR results at the district and school levels, depending on the number of students who were assessed using the CMA.” Unfortunately, the SPI and CDE staff do not follow their own caution when they report the statewide results.

It is relatively easy to re-calculate the average percent *proficient and above* gains reported by the SPI to remove the hidden inflationary fudge factor. This can be done by simply adding the number of students taking CMAs to the denominators used for the percent *proficient and above* for each year and then recalculating to yield an “adjusted gain” score for the year. When these calculations are done, the adjusted annual gains are as follows:

<u>Year</u>	<u>Unadjusted Gain</u>	<u>Grade</u>	<u>Adjusted Gain</u>	<u>Grade</u>
2008	2.50	C+	2.00	C
2009	4.25	A	3.35	B+
2010	2.45	C+	2.00	C
2011	2.05	C	1.60	C-

The cumulative gain for the past four years reported by SPI press releases is 11.25 percentage points. However, if the CMAs had not been introduced, the cumulative gain would have been 8.95 percentage points. Thus, the SPI reported cumulative gains included an “inflation” factor of 2.30 percentage points, or an artificial inflation of **26 percent** over the past four years, due to the introduction of the CMAs to the STAR program.

I do not suggest that the CMAs should not be administered to selected Special Education students. Introduction of the CMAs has provided increased accessibility to STAR exams for many Special Education students, and that is a positive event. Rather, the point of this observation is that the introduction of the CMAs has generated an artificial boost to the STAR CST gain scores reported by the SPI. At the least, if a suitable adjustment for inflation cannot be incorporated, a caveat should accompany the SPI press releases indicating that some portion of the gains reported are indeed not true achievement gains but rather artificial due to changes in the overall STAR program. In addition, when CMA scores are included in California Academic Performance Index (API) scores, the CMA scores (which reflect lower achievement levels than the CSTs) should be weighted such that API gains are not artificially inflated. Last year, I estimated that API gain scores for elementary and middle schools were inflated by roughly 40 percent, and strongly suggested to CDE staff and the State Board of Education that adjustments for API calculations be made to address these artificial gains. Such adjustments have not been made as yet. With the availability of 2011 statewide STAR scores, including an increased use of CMAs for the Special Education population, it will be possible to re-calculate the CMA inflation effect for the 2011 Growth API scores scheduled for release in late August.

Finally, I might note that the US Department of Education has also noted that the introduction of so-called 2-percent tests based on modified achievement standards “obscures an accurate portrait of the academic needs of America’s students with disabilities.” [ED press release, 3/15/11] In a speech to the American Association of People with Disabilities, US Education Secretary Arne Duncan declared that students with disabilities should be judged with the same accountability system as everyone else. Sec Duncan has indicated that tests based on modified achievement standards will not be included in the so-called “next generation” tests designed to measure the Common Core standards that have now been adopted by many states, and indeed the work of both assessment consortia funded by the federal government to develop Common Core assessment systems does not include the development of tests like California’s CMA tests.

Increasing CMA Usage. When the CMA tests were proposed in 2007, CDE staff assured the State Board of Education that the tests would not affect more than 2 percent of the total population of students taking STAR, or roughly 20 percent of the Special Education population. That assurance was repeated as recently as early 2010. And, indeed, federal rules for the

inclusion of CMA scores in the federal accountability system (Annual Yearly Progress, or AYP) limits inclusion of CMA scores to 2 percent or less for students scoring *proficient and above* on a CMA. However, this limitation does not apply to the effect of CMA on either the STAR results released today by the SPI/CDE, nor on the calculation of scores for California’s statewide accountability system, the Academic Performance Index (API). In fact, by 2011 more than 4 percent of our total STAR population take CMAs, reflecting more than 40 percent of the Special Education population in California.

A look at the numbers of Special Education students taking CMAs over the past four years is instructive. The numbers below are the numbers of Special Education students taking the E/LA CMAs for each year as the CMAs were introduced in a phased fashion:

<u>Year</u>	<u>Grades 3-5</u>	<u>Grades 6-8</u>	<u>Grades 9-11</u>
2008	38,578	Not Avail	Not Avail
2009	54,021	47,215	Not Avail
2010	63,922	63,709	11,379*
2011	70,591	74,191	39,169

[Note: CMAs for Grades 9-11 were phased in over two years for budget reasons, and hence the 2010 data reflect only the CMA E/LA data for Grade 9; the 2011 data reflect CMA E/LA data for Grades 9, 10, and 11.]

The data released today show that in 2011 CMAs were administered to almost 184,000 Special Education students in California grades 3-11. That number is 4.4 percent of the total number of CSTs and CMAs administered grades 3-11, far greater than the advertised 2 percent. When one looks just at the Special Education students, for grades 3-11 just under 50 percent were administered CSTs, just over 40 percent were administered CMAs, and just less than 10 percent were administered CAPAs. However, as one can readily determine from the data above, it takes several years for CMAs to be fully introduced, and the numbers are still increasing even for the CMAs introduced 2 and 3 years ago. For Grades 4 through 8, already the CMAs are administered to more than 5 percent of the total population for each grade level, with several grades approaching 6 percent. And for grades 4 through 8, for every grade more CMAs are administered than CSTs to Special Education students.

The reason for the above participation data is extremely clear – the CMAs are based on modified achievement standards, which translates essentially to lower performance standard levels than the comparable levels for the CSTs. In plain English, the CMAs are easier tests than the CSTs. We do not know just how much easier the CMAs are, and that is a major flaw in the entire effort to introduce CMAs in California – it is possible to generate estimates for the comparability of CMA scores to CST scores, but that work has not been part of the CMA test development effort. Thus, we are flying blind when it comes to knowing exactly what a CMA *proficient* score means when translated to the CST scale of measurement. This work needs to be done before we can use CMAs with integrity for applications like contributions to API calculations. And for specific policy issues like, for instance, using STAR scores as an alternative means for the CAHSEE graduation requirement, the comparability of CMA scores to CST scores is a required piece of information for policymakers.

Finally, I might note a concern voiced last year regarding an “uneven” implementation of the CMA portion of STAR across school districts in California. Based on 2010 STAR results, last year I produced charts for percentages of Special Education students administered CMAs for 53 local districts in three counties, and the results were alarming – in some good sized districts, more than 70 percent of Special Education students were administered CMAs while in others less than 20 percent were administered CMAs. Since CMAs can serve to artificially increase API scores, differential implementation of CMAs for Special Education students can be one way to “game” the API system and artificially boost API gains. With the release of the STAR 2011 results today, it will be possible to update these data in the near future.

Nitty Gritty Program Participation Numbers

Each year, I look at several nitty gritty program participation numbers that become available with release of STAR results

Algebra I for Grade 8. Since 1997, California has had a goal that 8th graders take Algebra I. In 1997, an estimated 16 percent of 8th graders took Algebra I. In 2002, when the Algebra I end-of-course tests were first administered, 32 percent of 8th graders were enrolled in Algebra I courses. In 2009, the STAR results showed 59% of students took Algebra I by grade 8, reflecting a steady increase of roughly 4 percentage points per year. For 2010, this percentage increased to 64%, and in 2011 the percentage increased again to 67%. [Note: In recent years, the percentage of 7th graders taking Algebra I has become notable -- in 2011, almost 8 percent of 7th graders took the STAR Algebra I CST. The “by 8th grade” percentages above include the 7th grade numbers from the previous year.] These participation data show that California is making steady and commendable progress toward a long range goal of having all students take Algebra I in grade 8. Furthermore, the STAR data shows the percent *proficient* has not been adversely affected by the increased numbers and percentages of students enrolled in Algebra I: in 2002, the percent *proficient* was 39% while in 2010 the percent *proficient* was roughly 50%. These data tend to contradict any suggestion that actual achievement in Algebra I has been lowered by the increased numbers of participants.

English Learner Data. The overall numbers of English Learners, the numbers of ELs in bilingual programming, and the numbers of ELs taking the primary language tests are interesting data that become available each year with STAR results. The overall number of ELs tested by the STAR program has decreased a bit in recent years, from 1,129,000 in 2008 to 1,014,000 this year. I’m not sure why this is the case, but the decrease is at least curious. Could it be that California is finally redesignating ELs at rates higher than incoming EL rates? Or could it be that these data reflect our current and recent overall economic conditions, with more ELs leaving California than before?

The reported number of students in bilingual programming for grades 2-11 was about 35,000 this year, compared to about 40,000 last year, 45,000 in 2009, and 47-48,000 in 2007-8. As a percent of the total number of English Learners in these grades, the percent in bilingual programming slipped to 3.4 percent compared to 3.7 percent last year and 4.2 percent for each the previous three years. This percentage dropped dramatically, from an estimated 30 percent in 1998 (when Prop 227 was approved) to roughly 10 percent in 2000, and has been dribbling downward for the last 10 years to a leveling off in the 4 percent range. The percentage of all students (not just ELs) in bilingual programming is now only $\frac{3}{4}$ of one percent.

The number of students taking primary language tests [the *Standards-based Tests in Spanish (STS)*] has declined substantially the last four years, from roughly 68,000 in 2007 to about 58,000 students in 2008, to 48,000 in 2009, to 43,000 in 2010, and only 40,000 in 2011. Of these total numbers of ELs taking STS, less than 25,000 were coded as students in bilingual programming. One might want some further analysis why the number of ELs taking primary language tests is considerably less than the number in bilingual programming – one speculation I would offer is that LEAs are not administering the STSs to all students in bilingual programming since the STS scores are not used for accountability purposes. [Note that the STSs have not been designed to yield comparable results to the counterpart CSTs, and hence may not be used for accountability system calculations.]

Integrated / Coordinated CSTs. For several years I've noted that most of the Integrated or Coordinated CSTs at the secondary grade level for Math and Science have been very sparsely used. This year, the three Integrated Math CSTs were administered to only 17,000 students, about 0.4 percent of the total number of Math CSTs given grades 8-11. In addition, only the Coordinated Science I CST was given to a substantial number of students (roughly 55,000), while the remaining three Coordinated Science CSTs were administered to less than 6,000 students or only about 0.1 percent of the number of students grades 9-11. It has been more than 10 years since these Integrated / Coordinated CSTs were developed, in part based on advocacy from teachers and curriculum specialists with favorable views of integrated / coordinated approaches to teaching high school math and science. It is abundantly clear that the current Integrated / Coordinated Sciences CSTs either do not serve this advocacy well, or that the advocacy simply is not widespread enough to support continuation of the costs associated with developing and maintaining these CSTs. It is past time to review the STAR practice of continuing to administer the Integrated / Coordinated CSTs.