Summary of Responses to the Inclusion of Growth

The following is a summary of the responses provided to ADE by members of the ATAC regarding the inclusion of a growth measure.

There is a general consensus that having growth in the 9-12 model is not strongly favored under the Menu of Assessments. However, it is not a complete consensus. One member of ATAC suggested we need to take more time and exhaust all our resources before reaching a decision regarding growth. Another member would like a growth measure included in the FY19 system regardless of potential issues. Several members who raised concerns about growth wanted to revisit a growth measure following a year pause for additional information that would be necessary to address the potential challenges around growth in Menu.

Members of ATAC raised a number of concerns with a growth measure in a Menu of Assessments environment. Under Menu, growth calculations are challenging and subject to a number of important statistical and methodological considerations in order to be representative of student's growth over time. The growth measurement possibilities under Menu may lead to unsatisfactory outcomes such as statistically invalid measures, an inability to meaningfully differentiate between school performance, repetition of other indicators, introduction of construct biases, potential for erroneous mathematical representation of student development over time (e.g. sample size variability and assessment dependent variability in scores), and presenting a distraction from the intended developmental outcomes that students in high school should be acquiring (college and career preparedness). In addition, any growth measures that would be calculated are only one measure in a larger picture of accountability. Some of the other indicators available may be more beneficial for the time being, until more methodological developments or modeling potentials are determined in the future. The three models without growth provided by the ATAC emphasize the various possibilities available that can better capture the overall intentions and goals of 9-12 accountability without the issues that including growth brings to the table with marginal differences in regard to the relationship to FRL and school size.

Further, there were several comments on how the growth indicator is more purposeful at the K-8 levels where individual student's academic progress from year to year is a useful tool for teachers to understand the academic needs of students and the variation that exists in their development.

Index

Introduction to the Analysis: The following pages include descriptive statistics and letter grade distributions for three new models being currently explored by the ATAC.

In the following few short sections we briefly explain methodologies used for some of the analysis included below and some needed descriptions of the subgroups "free and reduced lunch (FRL)" and "school size".

Letter Grade Cut Scores

The letter grades in the analysis below follow cut scores based upon standard deviations ("SD" in the tables below) around the mean. To help frame our conversations we included what the FY18 letter grade model would look like had standard deviations of FY18 data been used to set the letter grade cut scores. Due to the methodology of setting letter grades based on standard deviations, the letter grade of the mean is always a C as the maximum value of a C is equal to the mean letter grade within these guidelines. All models use cut scores based on the standard deviation of their data. Below is an explanation of the standard deviation methodology. These cut scores are in no way a recommendation from the ATAC on letter grade cut scores and were simply used as a consistent baseline for comparison.

Letter Grade	Relationship to the Mean
A	Above one standard deviation from the mean
B	Between .01 of the mean and one standard deviation above
C	The mean and within one standard deviation below the mean
D	Between one standard deviation and two standard deviations from the mean
F	Below two standard deviations from the mean

Means and Medians:

We also included a comparison of the letter grades using the median score. The median is the score that is the exact midpoint in the data, where 50 percent of the population is above the median value and 50 percent is below the median value. We included data on the median score to more clearly illustrate the impact of the different models on school performance in the model. Slight rounding was done in the tables provided below for cleaner visualization.

Correlations:

A correlation coefficient is a measure of how closely two variables are in relation to one another. The number that is reported is the strength and direction of the relationship. The range for a correlation coefficient is -1 to +1. How strong of a relationship is problematic? Generally, a correlation coefficient between -.70 and +.70 is viewed as acceptable.

How to read correlation coefficient strength:

- **Exactly –1.** A perfect negative linear relationship
- **–0.70.** A strong negative linear relationship
- **–0.50.** A moderate negative linear relationship
- **–0.30.** A weak negative linear relationship

- **0.** No linear relationship
- **+0.30.** A weak positive linear relationship
- **+0.50.** A moderate positive linear relationship
- **+0.70.** A strong positive linear relationship
- **Exactly +1.** A perfect positive linear relationship

Growth Score:

At the December 17th Arizona State Board of Education meeting the State Board of Education asked ADE and the ATAC to develop, analyze, and present a model with a growth indicator. Included in this analysis are two approaches to the inclusion of a growth measure. Approach 1 is based on previous conversations in prior ATACs. Approach 2 was designed by ADE Accountability and Research staff. These approaches are included despite a lack of research supporting the inclusion of growth in an accountability system utilizing multiple tests. Further, there are statistical issues that bring into question the validity of such a system and such comparison. Lastly, growth would have to shift from being at the student level to the school level considering that under the menu of assessments environment students would not be taking state tests annually. Growth is not required for 9-12 under federal law. During the transition to menu, does the benefit outweigh the scrutiny and validity of the entire model?

Approach 1:

Student's current year performance level (1-4) will be compared to their most recent previous year level (with the last three years) regardless of assessment taken.

- CY FAY students with a valid test score (ELA and Math to be calculated separately)
  - CY Performance Level
  - PY Performance Level

Weighted

	CY Level 1	CY Level 2	CY Level 3	CY Level 4
PY Level 1	0	0	0	0
PY Level 2	1	0	0	0
PY Level 3	2	1	1	1
PY Level 4	3	2	2	1

Approach 2:

11th Grade Cohort Percent Change and Improvement (Growth)

This is a school level changed based on cohort to cohort comparison.

- CY FAY 11th Grade Cohort (2019 for FY 2018)
- PY FAY 11th Grade Cohort (2018 for FY 2017)
- All ELA and All Math State Assessments taken at that school using the last assessment taken at that school for that 11th grade cohort for both ELA and Math.
  - CY FAY 11th Grade Cohort (2019 for FY 2018) testing years 18, 17, 16
  - PY FAY 11th Grade Cohort (2018 for FY 2017) testing years 17, 16, 15
- $ELA\% = \left(\frac{CYELA}{}\right)\frac{(\#\ of\ Assessments\ in\ PL\ 3\&4)}{Total\ \#\ of\ Assessments} - \left(\frac{PYELA}{}\right)\frac{(\#\ of\ Assessments\ in\ PL\ 3\&4)}{Total\ \#\ of\ Assessments}$
- $Math\% = \left(\frac{CYMath}{}\right)\frac{(\#\ of\ Assessments\ in\ PL\ 3\&4)}{Total\ \#\ of\ Assessments} - \left(\frac{PYMath}{}\right)\frac{(\#\ of\ Assessments\ in\ PL\ 3\&4)}{Total\ \#\ of\ Assessments}$

School Size:

The school size groups used for correlation analysis are also used by other departments in the Arizona Department of Education. In the provided appendix we also look at correlation across raw student enrollment, finding similar relationships.

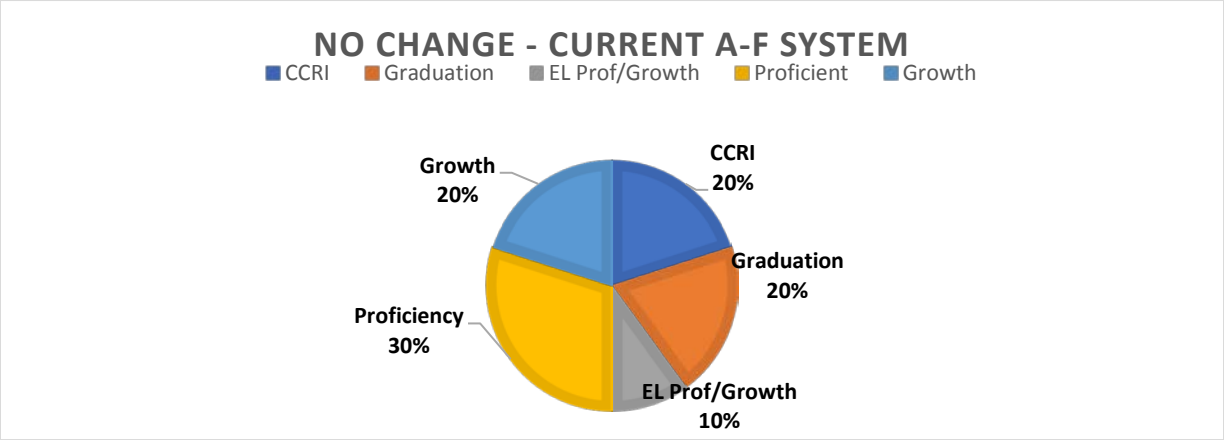
School Sizes
1=100 students or less
2=101 to 350
3=351 to 600
4=601 to 900
5=901 or more

Free or Reduced Lunch:

The free or reduced lunch groups used for correlation analysis are the same as those submitted to the State Board of Education previously for discussion around the A-F system.

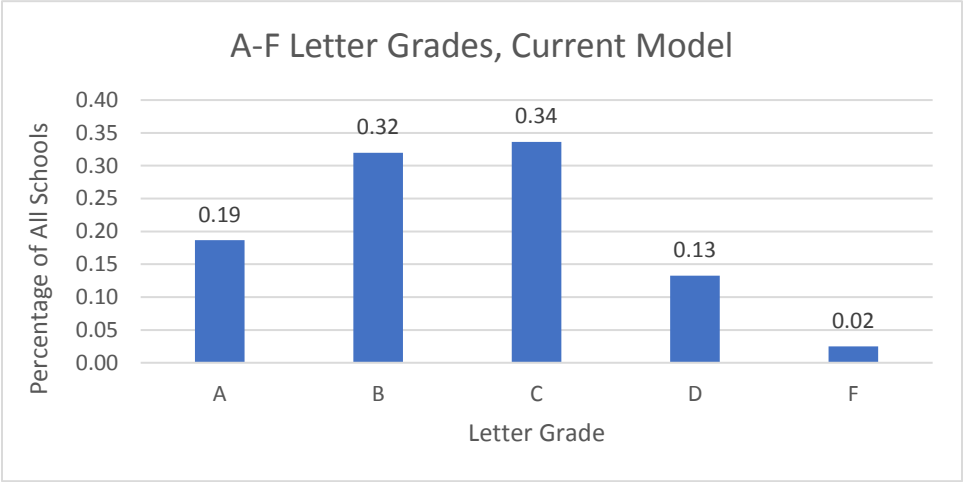
Free and Reduced Lunch
1=25% of students or less
2=26 to 50%
3=51 to 75%
4=76% or more

Current Model, Current Growth



Mean Percent Earned	Median Percentage Earned	Median Grade	Std. Dev	Minimum	Maximum
70.51	70.85	B	15.59	13.83	104.56

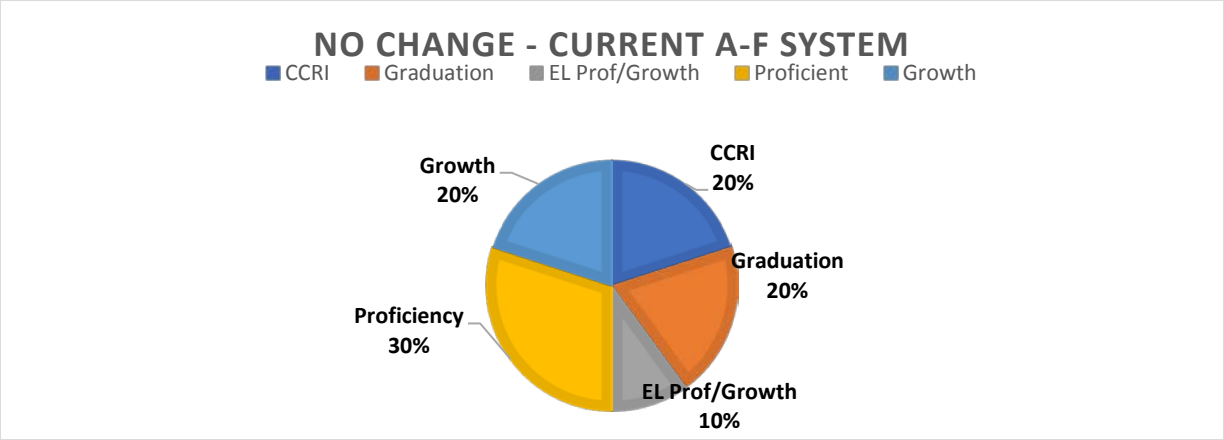
Current Model – SD Cut Scores	Min	Max	Percent
A	86.11		0.19
B	70.52	86.10	0.32
C	54.93	70.51	0.34
D	39.34	54.92	0.13
F	0.00	39.33	0.02



Correlations to Free and Reduced Lunch and School Size

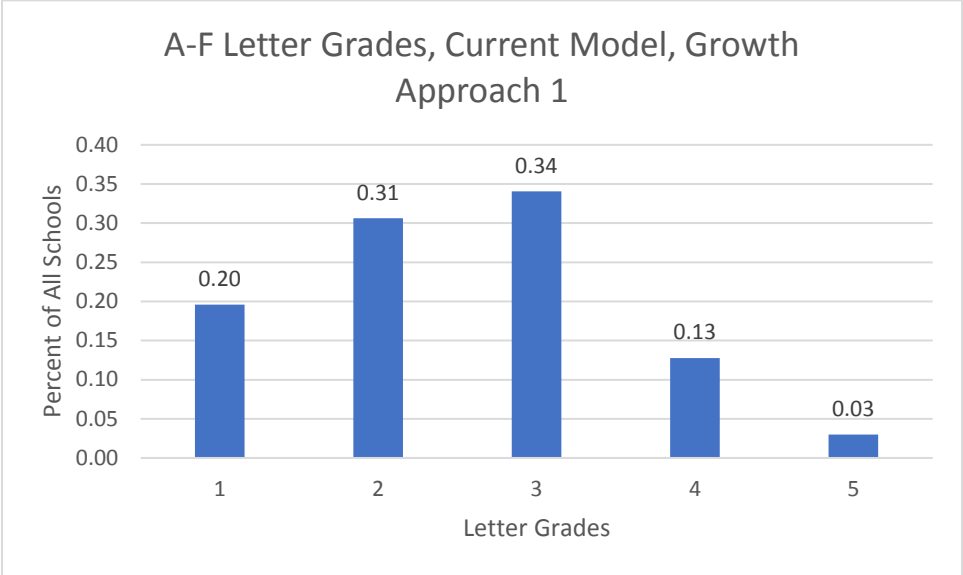
Model	FRL	School Size
Current	-0.37	0.11

Current Model, Growth Approach 1



Mean Percent Earned	Median Percentage Earned	Median Grade	Std. Dev	Minimum	Maximum
65.84	66.59	B	16.16	9.45	106.33

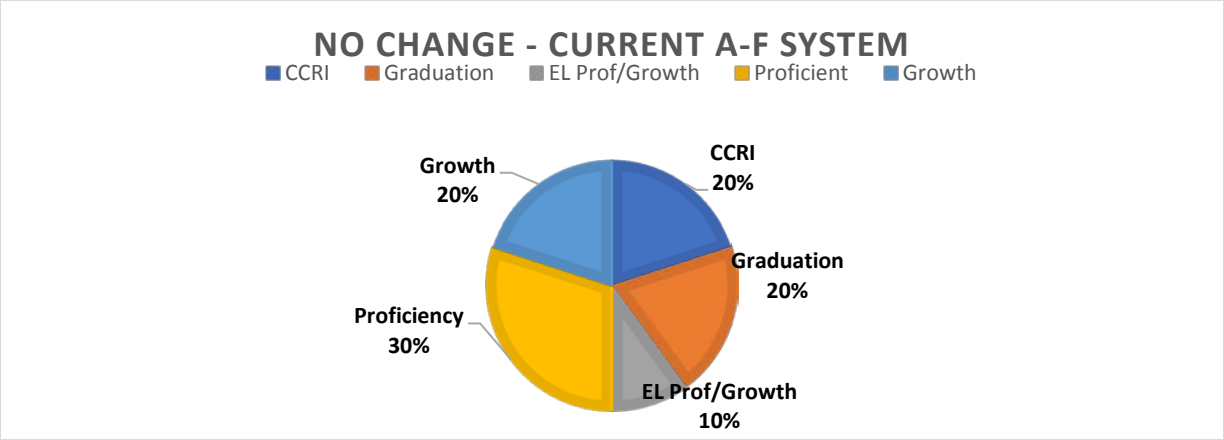
Current Model – SD Cut Scores	Min	Max	Percent
A	82.01	100	.20
B	65.85	82.00	.31
C	49.69	65.84	.34
D	33.53	49.68	.13
F	0	33.52	.03



Correlations to Free and Reduced Lunch and School Size

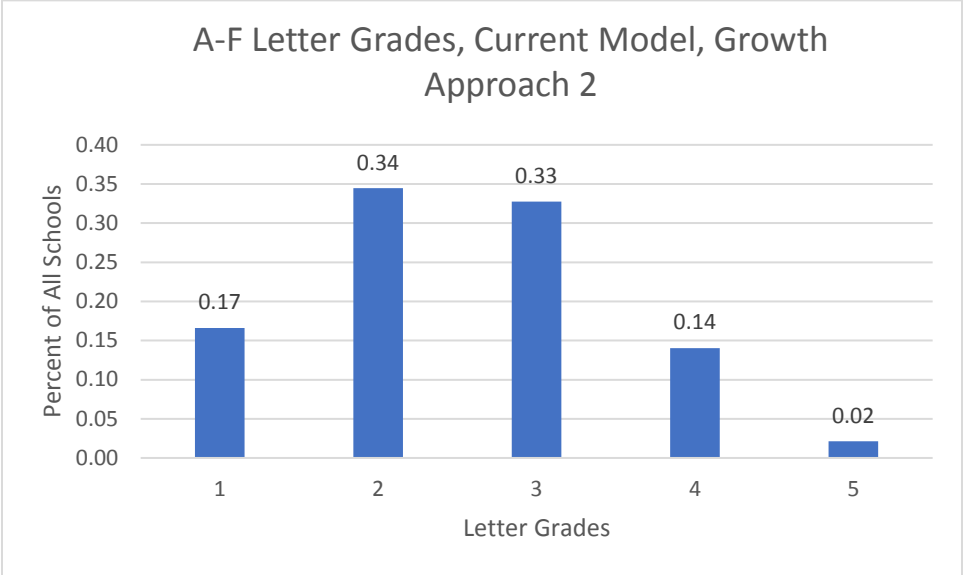
Model	FRL	School Size
Current	-0.36	0.18

Current Model, Growth Approach 2



Mean Percent Earned	Median Percentage Earned	Median Grade	Std. Dev	Minimum	Maximum
63.43	63.63	C	14.75	8.25	95.44

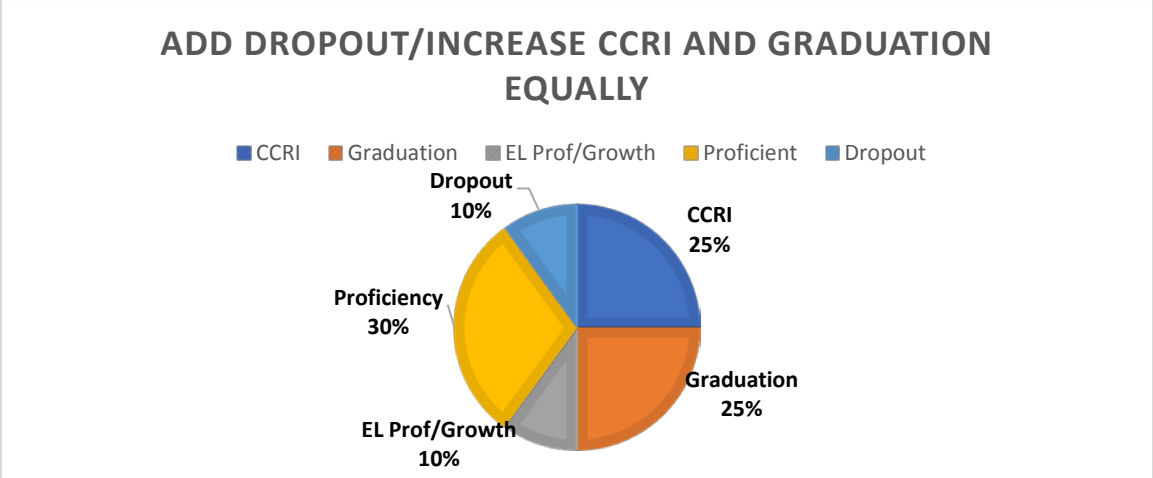
Current Model – SD Cut Scores	Min	Max	Percent
A	78.19	100	.17
B	63.44	78.18	.34
C	48.70	63.43	.33
D	33.95	48.69	.14
F	0.00	33.94	.02



Correlations to Free and Reduced Lunch and School Size

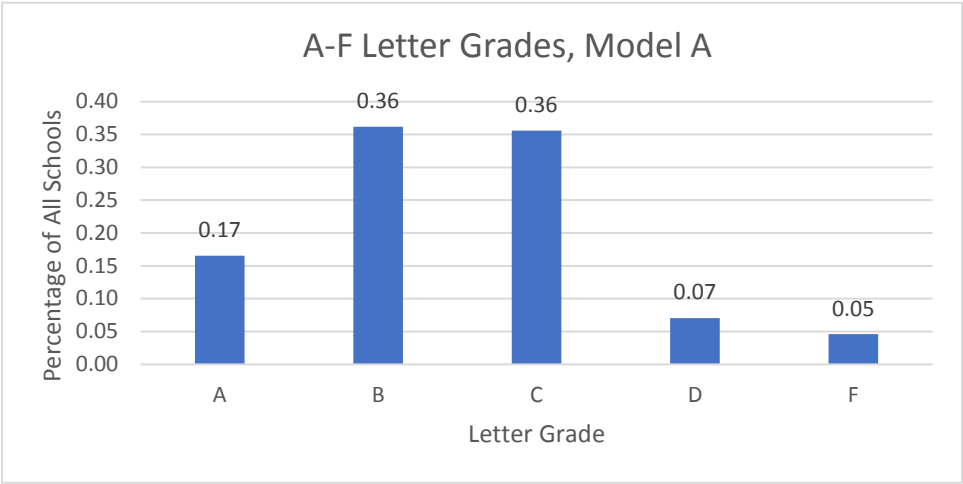
Model	FRL	School Size
Current	-0.34	0.13

Model A



Mean Percentage Earned	Median	Letter Grade of Median Score	Std Dev	Minimum	Maximum
72.09	72.85	B	15.77	16.75	106.89

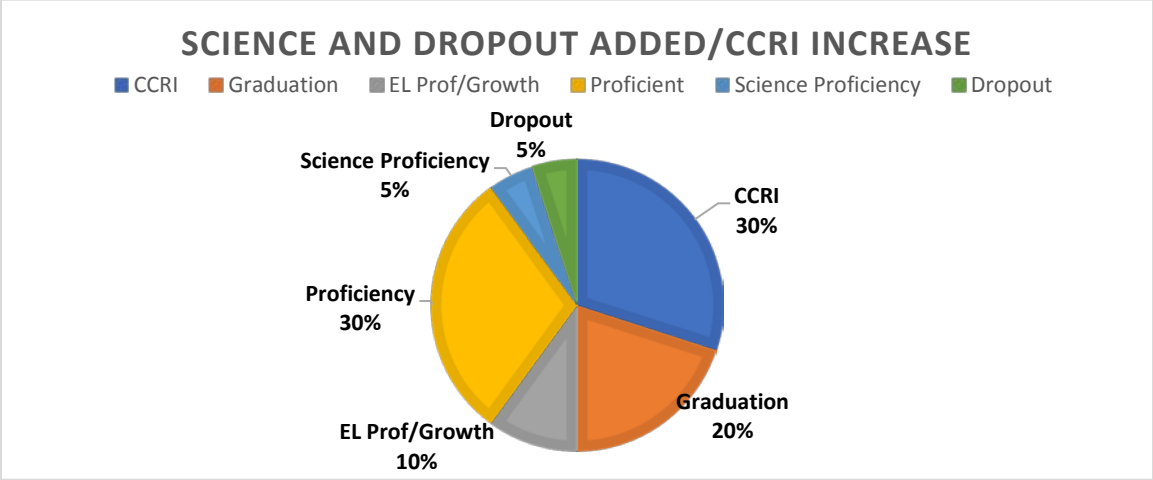
Model A – SD Cut Scores	Min	Max	Percentage
A	87.87		0.17
B	72.10	87.86	0.36
C	56.33	72.09	0.36
D	40.57	56.32	0.07
F	0.00	40.56	0.05



Correlations to Free and Reduced Lunch and School Size

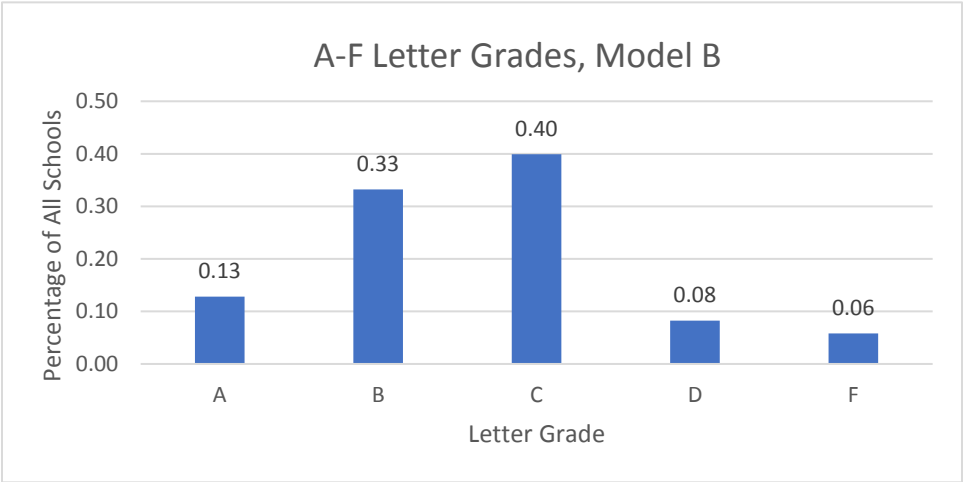
Model	FRL	School Size
Model A	-0.40	0.25

Model B



Mean Percentage Earned	Median	Letter Grade of Median Score	Std Dev	Minimum	Maximum
64.37	64.55	B	15.57	13.77	100.52

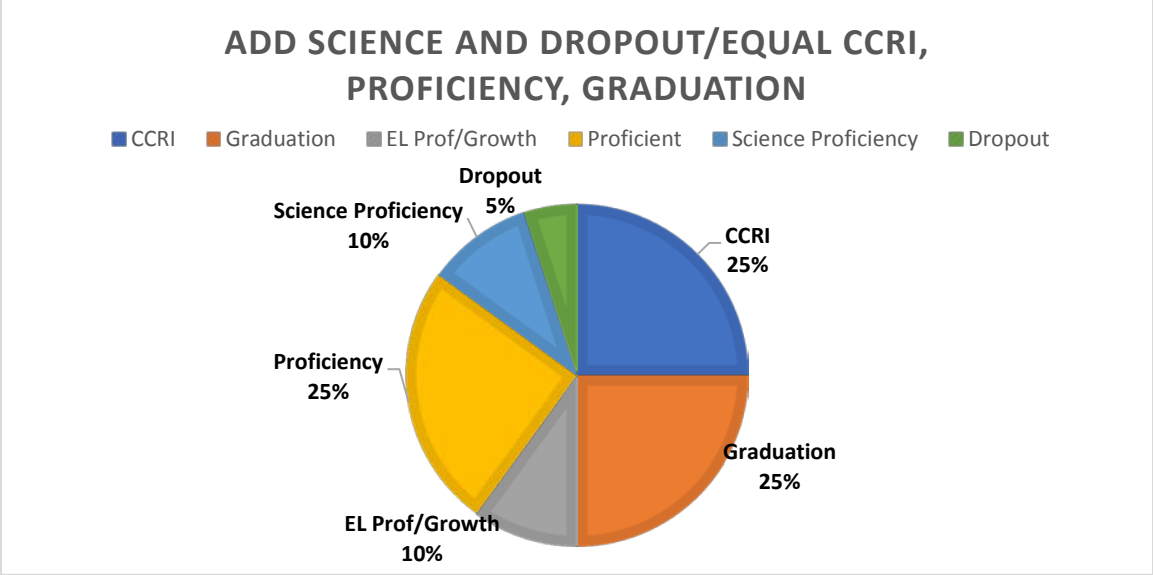
CP Model 4 – SD Cut Scores	Min	Max	Percentage
A	79.95	100	0.13
B	64.38	79.94	0.33
C	48.81	64.37	0.40
D	33.24	48.80	0.08
F	0.00	33.23	0.06



Correlations to Free and Reduced Lunch and School Size

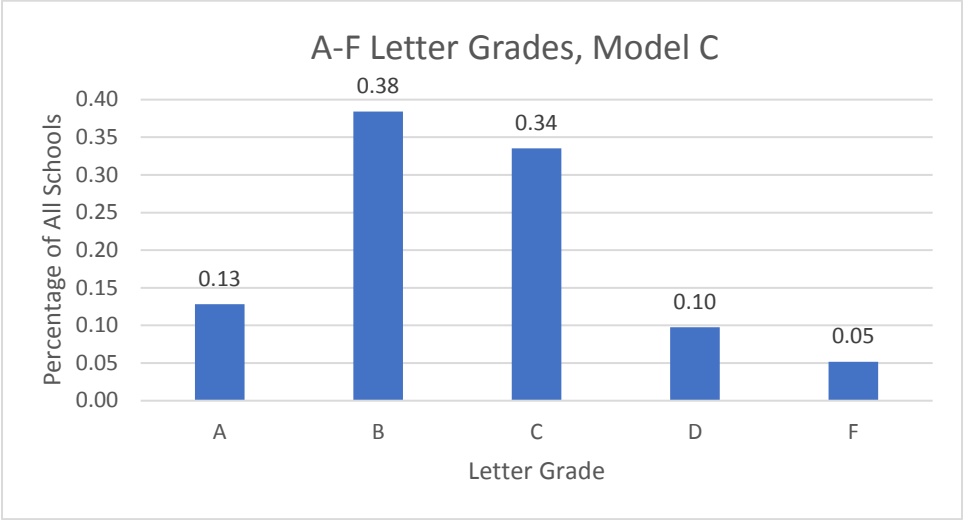
Model	FRL	School Size
Model B	-0.41	0.25

Model C



Mean Percentage Earned	Median	Letter Grade of Median Score	Std Dev	Minimum	Maximum
67.35	67.88	B	15.83	14.18	103.60

Model C – SD Cut Scores	Min	Max	Percentage
A	83.19	100	0.13
B	67.36	83.18	0.38
C	53.53	67.35	0.34
D	35.70	51.52	0.10
F	0.00	35.69	0.05



Correlations to Free and Reduced Lunch and School Size

Model	FRL	School Size
Model C	-0.41	0.26

One suggestion discussed with the ATAC was the suspension of a Growth indicator for one year while ATAC and ADE work to develop a Growth indicator that is not test dependent. The first question asked was how many schools would have increase and decreased their letter grades in FY18 if Growth had been removed compared to the existing FY18 A-F system. The amount of schools that increased are highlighted in blue in the table below while schools that decreased are highlighted in yellow. In total, 18 schools would have increased their performance and 17 would have decreased.

	A with Growth	B	C	D	F
A without Growth	**86**	**7**	**0**	**0**	**0**
B	**3**	**84**	**6**	**0**	**0**
C	**0**	**4**	**86**	**3**	**0**
D	**0**	**0**	**5**	**37**	**2**
F	**0**	**0**	**0**	**5**	**11**

Questions were immediately raised around the impact of free and reduced lunch (FRL) and school size regarding which schools increased and decreased. Given the small N-size of the population in question, statistical analysis that ATAC has previously requested would not be appropriate. In the tables below is some summary information about the different populations. The "FRL Group Average" column is based on the FRL size groupings used by ADE and the US Department of Education. The "AVG % FRL" is based on the reported percent of each school that either increased or decreased. **All schools in this population have an FRL percentage reported.** "AVG School Size" is the FY18 enrollment at each school. "School Size Group Average" is based on the five school size groupings used by ADE and used in ATAC analysis previously. "% Large Schools" is based on a measure requested by ATAC previously, which split school size into two groups, those above 351 students (large schools) and those below (small schools). ATAC asked for this measure due to concerns that five school size groupings might be masking the impact of school size and that collapsing size into two groups might better represent the impact of school size. Below these tables are two graphs showing the distribution of the data by FRL and school size.

School Sizes	**Free and Reduced Lunch**
1=100 or less students	1=25% or less
2=101 to 350	2=25.01% to 50%
3=351 to 600	3=50.01% to 75%
4=601 to 900	4=75.01% or more
5=901 or more	

	FRL Group Average	AVG % FRL	AVG School Size	School Size Group Average	% Large Schools
Increase	2.2	0.46	919	3.3	0.5
Decrease	2.3	0.47	592	2.80	0.47