Summary of Finding from an Analysis of the SBAC Interim Assessments' Impact on Student Outcomes on the SBAC Summative Assessment (SY 2017 – 2018)

Robin Clausen, PhD
GEMS Stakeholder Liaison and Research Analyst
Montana Office of Public Instruction

This analysis finds that there are positive significant mean differences when comparing student outcomes on the Smarter Balanced Assessment Consortium (SBAC) Interim Assessments and corresponding student outcomes on the SBAC Summative Assessment indicating growth between the two time periods the tests were taken and a possible positive impact of taking the SBAC Interim and subsequent summative scores. This is important as we encourage districts to implement the testing, or specific parts of the testing, we can state that for numerous parts of the assessments students showed growth between the two test periods. There are two kinds of Interim Assessments. The Interim Assessment Blocks Assessment (IAB) focus on specific topics, can be administered in one class period, and do not require that all blocks are taken in a grade level. The Interim Comprehensive Assessment measures the same content and standards as the SBAC Summative Assessment, takes 3-4 hours to administer, and provides information on achievement overall, including claims data. Both assessments can be taken as a standardized or non-standardized assessment. Both assessments provide data in the form of a scale score that is comparable with the scores on the SBAC Summative Assessment.

In Montana, the interim assessments were taken over a three month period prior to the summative assessment window. 63 school districts participated with varying degrees of intensity. The top 25 districts based on student participation on the IABs is presented in Table 1. There are three different measures of intensity: student count, number of ELA blocks taken, and number of math blocks taken. A relative intensity (number of blocks tested for either ELA or Math per student) can be inferred by comparing the total number of students with the ELA or Math opportunities taken (an opportunity is an incidence where a certain block was taken by a student). Not surprisingly, the larger districts factor in the top five. Kalispell had students at each grade level take the assessments. Billings focused on grades 3-5. No Billings students took the assessment in grades 6-8 (high school/11th grade represented too few students to measure in this analysis). Butte focused on Grades 3-6. Belgrade had students in all grades take the assessment, with the exception of Grade 5. Missoula County Public Schools had students take the assessment in grades 3-5 and 7 assessments.

Table 1: Student Count and Relative Intensity of Number of Blocks Taken

	Student Count	ELA Opportunities	Math Opportunities
Kalispell Public Schools	1902	3317	3806
Billings Public Schools	1527	750	2434
Butte Public Schools	1229	1628	2762
Belgrade Public Schools	983	1764	2084
Missoula Co Public Schls	926	851	1608
Hellgate Elementary	676	1678	1916
Ronan Public Schools	640	943	1655

	Student Count	ELA Opportunities	Math Opportunities
Libby K-12 Schools	501	963	1390
Evergreen Elementary	374	495	661
Bozeman Public Schools	307	323	443
Columbia Falls Pub Schls	297	12	518
Polson Public Schools	264	156	323
West Valley Elementary	234	366	254
Malta K-12 Schools	221	301	639
Dillon Elementary	214	411	412
Lewistown Public Schools	167	219	273
Big Timber Elementary	162	21	205
Frenchtown K-12 Schools	161	265	179
Monforton Elementary	157	55	385
Troy Public Schools	143	493	606
Fort Benton Public Schls	140	579	557
Lone Rock Elementary	107	127	241
Columbus Public Schools	103	60	281
Gallatin Gateway Elem	92	156	211
Huntley Project K-12 Schls	86	38	125

When looking at the remainder of the list we see a variety of small districts that took the tests with relatively high intensities. This trend continues through the rest of the 63 districts. When looking at the numbers by grade level we see that the assessments were predominantly taken in the lower grades by a few large schools. This is important since not all large schools in Montana took the assessments (Helena and Great Falls are the notable exceptions).

Table 2: Student Count and Relative Intensity by Grade Level

Grade	Student Count	ELA Opportunity	Math Opportunity
2	11	5	0
3	2841	3475	6191
4	2866	3775	6217
5	2248	3150	5408
6	1725	3455	3464
7	1515	2131	3026
8	1285	1473	2079
Total	12491	17464	26385

This list shows that continuing to network with schools with high populations will have the most impact on the numbers of students taking the tests. Also, by encouraging the large districts to implement the

assessments in all grade levels numbers will improve. Nonetheless, encouraging small districts to participate as well may increase the relative intensity of how many opportunities are taken.

Results

ELA

Student earned a scale score grade on the assessment. The scale score is framed so that it is comparable across grade levels and across assessments. The research question that informed the analysis is: 'Do students who take the SBAC Interim assessments preform significantly higher on the SBAC Summative Assessment? The null hypothesis assumes there is no impact of the SBAC Interim Assessment upon subsequent SBAC Summative Test Scores. What I was looking to achieve with the IAB analysis was to look at the effectiveness of each IAB in producing better student outcomes and the relative effectiveness of the IABs by grade level (whether one IAB is better than another in producing positive outcomes on the SBAC Summative Assessment.

This analysis incorporates a paired samples t-test which can be used when a student has taken two comparable assessments. A T-test is a statistical test that compares the significance of the difference between two means (averages). In this section I focus on the effectiveness of each IAB in producing positive student outcomes on the summative.

Provided in Table 3 are the student mean scores on the various assessments at the third grade level. A column for the mean difference is provided. A positive score means that from the time that a student took a certain IAB (or, taking the IAB impacted the student's performance), there was growth in the normalized scale score. The next step to interpretation of this data is to look at the statistical significance (Sig.). This analysis adopts the $p<.05$ standard common in the social sciences. Stated another way there is less than a 5% chance that the mean difference occurred by chance.

Table 3: 3rd Grade Growth IAB to SBAC Summative - ELA[1]

	N	Mean IAB	Mean Summative	Mean Diff	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
G3Edit	600	2411.29	2447.03	35.74	96.01	3.92	-9.12	599	0.000
G3Listen	542	2428.02	2433.85	5.83	75.36	3.24	-1.80	541	0.072
G3Opinion	73	2229.53	2442.97	213.44	146.17	17.11	-12.48	72	0.000
G3ReadInfo	164	2361.32	2415.51	54.19	111.49	8.71	-6.22	163	0.000
G3ReadLit	176	2352.26	2447.52	95.27	113.08	8.52	-11.18	175	0.000
G3Research	218	2425.14	2447.98	22.84	70.34	4.76	-4.79	217	0.000
G3Revision	235	2447.24	2471.40	24.15	80.59	5.26	-4.59	234	0.000
G3Vocab	1152	2424.86	2428.24	3.38	76.06	2.24	-1.51	1151	0.132
G3Writes[2]									

[1] For space reasons I only present T-test for Grades 3,5,7 (ELA) and Grades 4,6,8 (Math). Extended data tables are available upon request.

[2] No students took the Brief Writes block.

Most of the blocks had students who experienced statistically significant growth between the time they took the IAB and their ultimate summative score later in the year. For example, 'Edit' shows statistically significant growth, with a mean difference of 35.74 points[3]. On the other hand, 'Vocab' shows no significant trend with a modest mean difference of 3.38. Stated another way, students who took most of the IABs showed significant higher outcomes on the summative assessments (with the exception of 'Listen' and 'Vocab.'

There are fewer significant statistical differences at Grade 5. Outcomes for the 'ReadInfo' block exhibit significant trends. There is a mean difference of 29.97. On the other hand 'Vocab' continues to show insignificant trends, with a mean difference of -3.03.

Table 4: 5th Grade Growth IAB to SBAC Summative - ELA

	N	Mean IAB	Mean Summative	Mean Diff	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
G5Editing	509	2523.58	2532.88	9.30	115.96	5.14	-1.81	508	0.071
G5Listen	501	2515.72	2515.47	-0.25	109.17	4.88	0.05	500	0.959
G5Narative	82	2539.35	2531.12	-8.23	71.68	7.92	1.04	81	0.301
G5ReadInfo	325	2508.43	2538.39	29.97	134.36	7.45	-4.02	324	0.000
G5ReadLit	171	2508.24	2536.16	27.92	96.16	7.35	-3.80	170	0.000
G5Research	113	2454.73	2523.56	68.82	243.90	22.94	-3.00	112	0.003
G5Revision	236	2493.58	2542.14	48.55	106.50	6.93	-7.00	235	0.000
G5Vocab	938	2528.51	2525.49	-3.03	73.60	2.40	1.26	937	0.208

For Grade 7, most measures show significant difference between IAB mean and Summative mean. These differences are positive indicating growth from the time a student took the IAB and when they took the summative. It is also possible to infer that taking the IAB related significantly to stronger performance on the SBAC Summative Assessment. In grade 7, for example, the 'Editing' block was taken among 513 students, had a mean difference of 58.88 points favoring the higher SBAC summative score, and this difference is significant at the p<.000 level. In contrast to the other grade levels mentioned above, 'Vocab' showed positive significant difference (30.55) favoring higher scores on the SBAC Summative at the p<.000 level.

Table 5: 7th grade Growth IAB to SBAC Summative - ELA

	N	Mean IAB	Mean Summative	Mean Diff	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
G7Editing	513	2533.95	2592.83	58.88	117.97	5.21	-11.30	512	0.000
G7Explanatory	9	2450.11	2490.78	40.67	91.06	30.35	-1.34	8	0.217
G7Listen	273	2542.97	2560.21	17.24	82.34	4.98	-3.46	272	0.001

[3] For full names of the variables please refer to the last page of this document.

	N	Mean IAB	Mean Summative	Mean Diff	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
G7ReadInfo	297	2564.23	2572.09	7.86	79.58	4.62	-1.70	296	0.090
G7ReadLit	91	2475.57	2536.70	61.13	78.92	8.27	-7.39	90	0.000
G7Research	293	2552.23	2607.65	55.42	94.26	5.51	-10.06	292	0.000
G7Revision	185	2515.74	2565.16	49.43	77.78	5.72	-8.64	184	0.000
G7Vocab	369	2558.22	2588.77	30.55	74.30	3.87	-7.90	368	0.000
G7Writes	20	2539.60	2503.25	-36.35	122.07	27.30	1.33	19	0.199

Overall, there were 29 statistically significant difference indicating students who took these blocks scored higher on the summative assessment (out of 39 total blocks for all grade levels). Stated another way these blocks are effective in producing the desired effect, providing students exposure to the test and test taking that had an impact on subsequent scores.

Math

The process to interpret the Math data tables is the same as with the ELA. The results are mixed at the fourth grade level. 794 students took the 'Algebraic Thinking' assessment. There is a positive mean difference (13.70) that is significant at the p<.001 level. On the other hand, there were 2193 students who took the 'Number and Operations'. The mean difference was negative (-3.32 which was significant at the p=.05 level).

Table 6: 4th grade Growth IAB to SBAC Summative – Math

	N	Mean Interim	Mean Summative	Mean Diff	Std. Deviation	Std. Error Mean	T	df	Sig. (2-tailed)
G4M-Geometry	1201	2489.55	2494.22	4.67	114.52	3.30	-1.41	1200	0.158
G4M-Measurement and Data	303	2488.89	2493.57	4.67	75.02	4.31	-1.08	302	0.279
G4M-Number and Operations	2193	2494.09	2490.77	-3.32	79.27	1.69	1.96	2192	0.050
G4M-Fractions	1536	2493.55	2498.43	4.89	79.89	2.04	-2.40	1535	0.017
G4M-Algebraic Thinking	794	2487.09	2500.78	13.70	77.67	2.76	-4.97	793	0.000
G4M-Performance	35	2445.46	2483.63	38.17	92.80	15.69	-2.43	34	0.020

Results for the 6th grade level are positive. Results for 'Ratio and Proportions' are positive. 570 student took the block. There is a positive mean difference (14.75) favoring growth with the Summative Assessment that are significant at the p<.000. There is another noteworthy difference with 'Statistics and Probability'. 209 students took the block. There is a positive mean difference of 41.52 points that is

significant at the p<.000 level. There are two negative mean differences ('Expressions and Equations' and 'Number System.' These difference are insignificant indicating that a trend is not observable.

Table 7: 6th grade Growth IAB to SBAC Summative - Math

	N	Mean Interim	Mean Summative	Mean Diff	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
G6M-Expressions and Equations	853	2544.75	2541.58	-3.16	68.67	2.35	1.35	852	0.179
G6M-Geometry	337	2514.45	2542.87	28.42	120.81	6.58	-4.32	336	0.000
G6M-Number System	926	2545.83	2543.70	-2.13	64.63	2.12	1.00	925	0.317
G6M-Performance	17	2436.41	2564.06	127.65	97.97	23.76	-5.37	16	0.000
G6M-Ratio and Proportions	570	2528.86	2543.61	14.75	74.59	3.12	-4.72	569	0.000
G6M-Statistics and Probability	209	2514.26	2555.78	41.52	111.77	7.73	-5.37	208	0.000

Results at the 8th grade are mixed. There was 621 students who took the 'Expressions and Equations' IAB. The mean difference was negative (-9.66) and this is significant at the p<.05 level. 292 students took the 'Geometry' IAB. The mean difference was negative (-48.67) and is significant at the p<.001 level. There are two significant positive differences. 134 students took the 'Expressions and Equations II' block. The mean difference is positive (35.16) and is significant at the p<.001 level. 269 students took the 'Number System' block. The mean difference is positive (21.19) and this is significant at the p<.05 level.

Table 8: 8th grade Growth IAB to SBAC Summative -Math

	N	Mean Interim	Mean Summative	Mean Diff	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
G8M-Expressions and Equations	621	2569.67	2560.01	-9.66	98.17	3.94	2.45	620	0.014
G8M-EXpressions and Equations II	134	2557.03	2592.19	35.16	104.04	8.99	-3.91	133	0.000
G8M-Functions	293	2586.01	2577.36	-8.65	114.29	6.68	1.30	292	0.196

	N	Mean Interim	Mean Summative	Mean Diff	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
G8M-Geometry	292	2618.50	2569.83	-48.67	136.11	7.97	6.11	291	0.000
G8M-Number System	269	2558.53	2579.71	21.19	121.41	7.40	-2.86	268	0.005
G8M-Performance[4]									

Overall, 20 blocks had positive statistically significant differences between the time of the IAB and the Summative Assessment (out of 35 blocks). Generally, the blocks that had the most students participate proved to be the least effective in increasing scores. Students in the upper grades show the most insignificant trends, likely due to the relatively small n's of students that took the assessment.

Relative Effectiveness of the IABs

A Multivariate Analysis of Variance (MANOVA) test was used to gauge the relative effectiveness of different assessments compared to each other by grade level. MANOVA allows us to test hypotheses regarding the effect of one or more independent variables on two or more dependent variables. In this case the dependent variables are the IAB scale score and the SBAC Summative scale score. The independent variable is the assessment being taken at each grade level. The test is a comparison of means as seen in the first table. In the second table you'll find the F statistics, the significance, and the Partial Eta Squared (a measure of effect size).

This analysis is useful if you want to recommend which IABs schools offer to their students. Most of the differences proved to be significant, however did not approach a level of having a moderate or higher effect (low: .01 - .124; moderate: .125 - .249; high: >.25). There was only one difference that approached a moderate level when comparing IABs for each grade level.

G8 Math proved to have a moderate effect. When analyzing the mean differences 'Expressions and Equations II' and 'Number System' have a high mean differences.

Table 9: Mean Difference- Grade 8 Math

	N	Mean IAB	SD IAB	Mean Summative	SD Summative	Mean Diff
G8Expressions and Equations	620	2569.64	114.28	2559.99	94.00	-9.65
G8Expressions and Equations II	134	2557.03	159.62	2592.19	116.46	35.16
G8Functions	293	2586.01	146.98	2577.36	109.02	-8.65
G8Geometry	292	2618.50	154.99	2569.83	107.46	-48.67

[4] No student took this performance task.

	N	Mean IAB	SD IAB	Mean Summative	SD Summative	Mean Diff
G8NumberSystem	268	2558.67	145.75	2579.73	96.17	21.06

When you look at the significance and effect size of these differences, there is no practically significant trend among the IAB score, meaning that the mean differences exhibit a small effect on the basis of the scores on the IAB. However, there is a moderate effect among the same students when they take the SBAC Summative Assessment (.174), indicating that students who took these assessments scored higher on the SBAC Summative Assessment.

Table 10: MANOVA Analysis of Grade 8 Math IAB

	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
IAB	788736.09	6	131456.01	6.92	0.000	0.025
Summative	3621785.86	6	603630.98	58.35	0.000	0.179

Even though there was only one effect that approached the moderate level, it is still possible to recommend IABs on the basis of their internal effectiveness (large significant positive mean differences) instead of one IAB is better than another.

Interim Comprehensive Assessment

The Interim Comprehensive Assessments were taken by fewer student (ELA: 64 students and Math: 526). These represent completed tests and may not factor in tests that had been started but not finished. There was one significant trend. Eighth grade ICA had a mean difference of -65.17 and was significant at the p<.050 level. However, only six students took the assessment and it is difficult to make inferences due to the small number of students.

Table 11: ELA ICA

	N	Mean IAB	Mean Summative	Mean Diff	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
G3ICA	1	2277.00	2464.00	187.00					
G4ICA	6	2331.00	2406.00	75.00	106.96	43.67	-1.72	5	0.147
G5ICA	27	2464.52	2469.44	4.93	59.62	11.47	-0.43	26	0.671
G6ICA	17	2564.00	2543.12	-20.88	42.33	10.27	2.03	16	0.059
G7ICA	7	2589.71	2565.29	-24.43	33.56	12.68	1.93	6	0.102
G8ICA	6	2597.67	2532.50	-65.17	41.38	16.89	3.86	5	0.012

More students took the Math ICA. For Grade 6 there was a positive mean difference of 18.26 that was significant at the p<.000 level. For Grade 7 there was a positive mean difference of 19.98 that was significant at the p<.000 level. For Grade 8 a similar trend. There was a positive mean difference of 19.99 that was significant at the p<.000 level. All of these findings show promise that as more students take the ICA there will undoubtedly be positive statistical differences.

Table 12: Math ICA

	N	Mean Interim	Mean Summative	Mean Diff	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
ICA-G3M	22	2400.73	2394.50	-6.23	44.21	9.42	0.66	21	0.516
ICA-G4M	3	2396.67	2372.33	-24.33	35.53	20.51	1.19	2	0.357
ICA-G5M	28	2479.79	2483.89	4.11	58.81	11.11	-0.37	27	0.715
ICA-G6M	143	2521.55	2539.81	18.26	48.29	4.04	-4.52	142	0.000
ICA-G7M	107	2546.31	2566.29	19.98	51.25	4.95	-4.03	106	0.000
ICA-G8M	223	2552.10	2572.09	19.99	56.75	3.80	-5.26	222	0.000

Conclusion

There are many statistically significant results when comparing the growth of students who take the IAB or ICA. Growth between the two time periods can be attributed to the treatments in each individual school. Taken as a group, the impact of taking the Interim Assessments on student outcomes on the Summative Assessment is also apparent. In a sense the test is both a measure and a treatment. Taking the assessments and subsequent growth in scores on the Summative assessment undoubtedly occurred in many areas. This impact could be analyzed in future analyses by comparing two years of summative scores with the IAB in a given year.

Table 13: Full Variable Names - IAB

Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Read Literary Texts Block Achievement Category	Read Literary Texts Block Achievement Category	Read Literary Texts Block Achievement Category	Read Literary Texts Block Achievement Category	Read Literary Texts Block Achievement Category	Read Literary Texts Block Achievement Category
Read Informational Texts Block Achievement Category	Read Informational Texts Block Achievement Category	Read Informational Texts Block Achievement Category	Read Informational Texts Block Achievement Category	Read Informational Texts Block Achievement Category	Read Informational Texts Block Achievement Category
Editing Block Achievement Category	Editing Block Achievement Category	Editing Block Achievement Category	Editing Block Achievement Category	Editing Block Achievement Category	Edit/Revise Block Achievement Category
Language and Vocabulary Use Block Achievement Category	Language and Vocabulary Use Block Achievement Category	Language and Vocabulary Use Block Achievement Category	Language and Vocabulary Use Block Achievement Category	Language and Vocabulary Use Block Achievement Category	Brief Writes Block Achievement Category
Revision Block Achievement Category	Revision Block Achievement Category	Revision Block Achievement Category	Revision Block Achievement Category	Revision Block Achievement Category	Listen/Interpret Block Achievement Category
Brief Writes Block Achievement Category	Brief Writes Block Achievement Category	Brief Writes Block Achievement Category	Brief Writes Block Achievement Category	Brief Writes Block Achievement Category	Research Block Achievement Category
Listen/Interpret Block Achievement Category	Listen/Interpret Block Achievement Category	Listen/Interpret Block Achievement Category	Listen/Interpret Block Achievement Category	Listen/Interpret Block Achievement Category	Explanatory Performance Task Block Achievement Category
Research Block Achievement Category	Research Block Achievement Category	Research Block Achievement Category	Research Block Achievement Category	Research Block Achievement Category	Mathematics Number of Blocks Tested
Opinion Performance Task Block Achievement Category	Narrative Performance Task Block Achievement Category	Narrative Performance Task Block Achievement Category	Argument Performance Task Block Achievement Category	Explanatory Performance Task Block Achievement Category	Mathematics OppNumber
Geometry Block Achievement Category	Operations and Algebraic Thinking Block Achievement Category	Operations and Algebraic Thinking Block Achievement Category	Ratio and Proportional Relationships Block Achievement Category	Ratio and Proportional Relationships Block Achievement Category	Expressions and Equations II Block Achievement Category
Operations and Algebraic Thinking Block Achievement Category	Number and Operations in Base 10 Block Achievement Category	Number and Operations in Base 10 Block Achievement Category	Geometry Block Achievement Category	Geometry Block Achievement Category	Functions Block Achievement Category
Number and Operations in Base 10 Block Achievement Category	Fractions Block Achievement Category	Fractions Block Achievement Category	Expressions and Equations Block Achievement Category	Expressions and Equations Block Achievement Category	Geometry Block Achievement Category

Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Fractions Block Achievement Category	Geometry Block Achievement Category	Geometry Block Achievement Category	Number System Block Achievement Category	Number System Block Achievement Category	Mathematics Performance Task Block Achievement Category
Measurement and Data Block Achievement Category	Measurement and Data Block Achievement Category	Measurement and Data Block Achievement Category	Statistics and Probability Block Achievement Category	Statistics and Probability Block Achievement Category	Number System Block Achievement Category
Mathematics Performance Task Block Achievement Category	Mathematics Performance Task Block Achievement Category	Mathematics Performance Task Block Achievement Category	Mathematics Performance Task Block Achievement Category	Mathematics Performance Task Block Achievement Category	