Summary of Comparisons Involving SBAC Interim Results (SY 2018 – 2019)

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

Overview

Montana uses the Smarter Balanced assessment that is aligned to Montana's Content Standards for mathematics and English Language Arts (ELA). These content standards were adopted in fall of 2011 and implemented July 1, 2013. The test began its rollout during the 2012–2013 school year with a large-scale volunteer pilot. This is the general population assessment given online as a computer adaptive test for Grades 3–8. In 2013–2014, there was a census field test administration where the state participated in the U.S. Department of Education "double-testing" waiver for states. In 2014–2015, the first operational assessment was given in all schools.

This report has been prepared to provide an overview of 2018–2019 Montana data from the Smarter Balanced SBAC Interim assessments. Comparisons of student performance on the Interim assessments can be made that show individual student growth and group performance trends. In addition, comparisons can be made between Interim assessment test takers and the remaining non-test takers by grade level.

The first analysis examines whether Interim test takers show growth on the Summative assessment in comparison to their Interim assessment scores. There are many possible reasons for growth including student and teacher familiarity with the assessment platform, teacher and student familiarity with grade level content, teacher familiarity with grade-level standards, as well as data-driven instruction. The Interim assessment provides 'just-in-time' feedback to the teacher that allows for data-driven instruction. The null hypothesis for this analysis was, "Student performance on the Interim assessment shows no relationship to Summative test scores". This hypothesis aimed to test if there were differences between the Interim and Summative scores and what was the magnitude and direction of these differences.

The second analysis examined the degree in which Interim test takers score higher on the Summative Assessment than their non-participating peers. This analysis was important to share with the Montana schools since Interim assessments are optional for all accredited schools. (Note during the 2019–2020 school year Interim assessments were required for schools designated as 'comprehensive'). Despite being coincidental, some important demographic differences exist between those who did and did not take an Interim assessment(s). Higher percentages of Interim test takers were eligible for free and reduced lunch than students in the non-testing comparison group. Higher numbers of American Indian students took the Interim assessments as well. The focus of the analysis was to see whether these differences had any impact on the comparison (individual student and group). The null hypothesis for this analysis was, "There is no difference between comparison group performance and Interim test taker performance." This hypothesis aimed to test whether there was a relationship between the Interim test

takers score and Summative test takers score and what was the magnitude and direction of the differences.

There are two kinds of Interim assessments that schools may choose to administer to their students. The Interim Assessment Blocks (IAB) focus on specific targets and can be administered within one class period. Multiple IABs are available for each grade (3–8 and high school) for math and English Language Arts. For this analysis, the number of IAB assessments taken was not a factor. Instead we considered any student taking at least one IAB to be an IAB test taker.

The Interim Comprehensive Assessment (ICA) has the same blueprint and achievement-level indicators as the Summative assessment. Both assessments provide data in the form of a scale score that is comparable to the scores on the Summative assessment. Because scaled scores varied by grade level, all analyses shown in this report were completed by grade level.

There are several notable trends about districts that chose to participate in the Interim assessment. Kalispell had the most students who took the Interim assessments in both SY2018 and SY2019. Students took the Interim assessments in all tested grades. Billings Elementary also had strong participation during both school years with students taking the Interim assessments predominantly in the lower grades. Belgrade also had a high numbers of test takers in SY2018 and SY2019. Table 1 highlights 15 Districts with the largest numbers of students who took at least one IAB and the percent of the total population that took the Interim assessments; for example, Kalispell represents 14.52% of the population that took at least one ELA IAB and 11.39% of the math population. When compared with SY2018, Browning Elementary, Livingston Elementary, Harlem Elementary, and Whitefish Elementary participated either for the first time or at a greater degree of frequency.

Table 1: District-Level Participation Frequencies (Top 15)

	ELA IAB			**Math IAB**	
	Frequency	**Percent of Total**		**Frequency**	**Percent of Total**
Kalispell Elem	1520	14.52	**Kalispell Elem**	1348	11.39
Billings Elem	921	8.80	**Billings Elem**	1299	10.97
Belgrade Elem	817	7.81	**Belgrade Elem**	1111	9.38
Ronan Elem	660	6.31	**Missoula Elem**	743	6.28
Missoula Elem	550	5.25	**Butte Elem**	723	6.11
Butte Elem	521	4.98	**Browning Elem**	689	5.82
Browning Elem	476	4.55	**Ronan Elem**	661	5.58
Evergreen Elem	463	4.42	**Evergreen Elem**	458	3.87
Hellgate Elem	455	4.35	**Libby K-12 Schools**	432	3.65
Helena Elem	433	4.14	**Helena Elem**	419	3.54
Libby K-12 Schools	427	4.08	**Hellgate Elem**	331	2.80
West Valley Elem	338	3.23	**Columbia Falls Elem**	312	2.64
Dillon Elem	249	2.38	**Whitefish Elem**	251	2.12
Livingston Elem	216	2.06	**Harlem Elem**	235	1.98

	ELA IAB			Math IAB	
	Frequency	**Percent of Total**		**Frequency**	**Percent of Total**
Whitefish Elem	208	1.99	**Bozeman Elem**	232	1.96

Districts participated for a variety of reasons. The Interim assessments are entirely optional, and schools participated at different rates, and in different grade levels. It is believed the motivation to participation in the Interim assessments likely driven by past participation or recommendations from others on the benefits of the Interims. Many school districts have expressed a lack of awareness of the option to participation at no cost in the Interim assessments or have chosen not to acknowledge the benefits. More importantly, among those Districts that did participate, participation at certain grade levels is a local decision (or, if all students in a grade level participate). The aggregate of these local decisions leads to the conclusion that school districts reacted to the Interim assessment in highly different and variable ways.

One way to gauge the intensity of a school district's use of the IAB was to look at the number of IABs taken by students (Table 2). Belgrade had more students take IABs in math than any other school in state. Kalispell also had high frequency of usage of the IAB assessments.

Table 2: Number of IAB Tests Taken by District (Top 15)

ELA	Count	Math	Frequency
Kalispell Elem	4141	**Belgrade Elem**	3488
Belgrade Elem	2789	**Kalispell Elem**	3410
Ronan Elem	1692	**Billings Elem**	2945
Libby K-12 Schools	1603	**Butte Elem**	2856
Billings Elem	1564	**Browning Elem**	2725
Missoula Elem	1547	**Ronan Elem**	2156
Evergreen Elem	1341	**Evergreen Elem**	1878
Colstrip Elem	1083	**Libby K-12 Schools**	1824
Butte Elem	1055	**Missoula Elem**	1605
Browning Elem	977	**Hellgate Elem**	885
Helena Elem	883	**Whitefish Elem**	878
Hellgate Elem	780	**Helena Elem**	816
Fort Benton Elem	663	**Colstrip Elem**	793
West Valley Elem	568	**Columbia Falls Elem**	736
Troy Elem	529	**Troy Elem**	656

Fewer students participated in the ICAs. Stevensville Elementary and Lame Deer Elementary represent the majority of the population that took the ICAs. Table 3 contains a list of schools that most frequently took the ICAs.

Table 3: ICA Participation by District

	ELA LEA			Math LEA	
	Frequency	**Percent of Total**		**Frequency**	**Percent of Total**
Stevensville Elem	505	63.05	**Stevensville Elem**	497	50.00
Lame Deer Elem	148	18.48	**Lame Deer Elem**	153	15.39
DeSmet Elem	62	7.74	**Monforton Elem**	131	13.18
Fort Benton Elem	54	6.74	**Fort Benton Elem**	76	7.65
Pendroy Elem	11	1.37	**DeSmet Elem**	67	6.74

Interesting trends were found when looking at the grades involved in the IAB for ELA, that is, Grade 3 (2355), Grade 4 (2385), and Grade 5 (2150) accounted for the largest number of students who took the IAB for ELA as shown in the Table 4. Also found in this table is the distribution of non-participating students, the comparison group.

Table 4: ELA IAB Grade Level Distribution

	Interim ELA IAB Test Takers		**Did Not Participate ELA IAB**	
Grade	**Frequency**	**Percent of Grade Level Population**	**Frequency**	**Percent of Grade Level Population**
3	2355	21.24%	8733	78.76%
4	2385	20.68%	9150	79.32%
5	2150	18.14%	9700	81.86%
6	1603	13.94%	9893	86.06%
7	1111	9.74%	10297	90.26%
8	817	7.63%	9888	92.37%
Total	10419	15.30%	57661	84.70%

Presented in Table 4 is the percentages of total test takers out of Montana's tested grade student population that took or did not take an IAB for ELA or an ICA for ELA. The IAB for ELA test takers yielded 15.30% of the total grades 3–8. This percentage was more than 20% of the tested grade student population took the assessments in grades 3 and 4.

The total number of students that took the IAB for Math (11,828) and the total population of the comparison group (55,988) are different than from what has been reported above. The same trend remains with more students taking the IAB in the lower grades than in the upper grades. There are approximately the same number of students in each tested grade in Montana, that is, grades 3–8. The distribution between tested grades in the treatment group was roughly the same as the comparison group. The tested grade numbers for the ICA in ELA and ICA in Math were also evenly distributed. The total numbers of students taking the ICA in ELA was 801 and for the ICA in Math was 994.

Demographics

In examining the demographic makeup of Interim test takers, two trends stand out. First, using participation in the National School Lunch Program (NSLP) as a proxy for economically disadvantaged, we find that students who qualify for free and reduced lunch services took the Interim assessment more frequently than their peers in the comparison group. For example, 53.96% of students who took the IAB in ELA received free or reduced lunch, compared with 44.32% of their grade-level peers that did not take the IAB (Table 5).

Table 5: Economic Disadvantage

	ELA IAB Test Takers		**Did Not Participate**	
	Frequency	**Percent of Total**	**Frequency**	**Percent of Total**
Free	5121	48.93	22120	38.36
Reduced	527	5.03	3437	5.96
Not Eligible	4819	46.04	32104	55.68
Total	10467	100	57661	100

A similar trend was found for students taking the IAB in math and their comparison group. Among the Math IAB tested students, 53.91% qualified for free or reduced lunch, compared with 44.25% of their grade-level peers that did not take the IAB. Even greater percentages of ICA test takers qualified as economically disadvantaged. Among the ICA in ELA test takers, 60.92% qualified for free or reduced lunch. Among the ICA in Math test takers, 55.73% qualified for free or reduced lunch. These findings highlight schools with greater students qualifying for free and reduced lunch were administered Interim assessments.

Second, there are two noteworthy observations among the students who took the IAB in ELA for the racial demographic variable. First, there was a large population (1600) of American Indian/Alaskan Native students representing 15.29% of the population of total test takers. When compared with students that did not participate in the Interim assessments, 10.23% of the population is American Indian. Smaller percentages of the other race/ethnicity subgroups participated in the Interim assessments. For example, 75.07% of the Interim assessment population was White, compared to 78.52% of the non-participating population (Table 6).

Table 6: Race and Ethnicity Trends

	ELA IAB Test Takers		**Did Not Participate**	
	Frequency	**Percent of Total**	**Frequency**	**Percent of Total**
Hispanic	466	4.45	2979	5.17
American Indian/ Alaskan Native	1600	15.29	5899	10.23
Asian	67	0.64	377	0.65
Black	44	0.42	460	0.80
Native Hawaiian/ Pacific Islander	26	0.25	122	0.21
White	7858	75.07	45273	78.52

	ELA IAB Test Takers		Did Not Participate	
	Frequency	**Percent of Total**	**Frequency**	**Percent of Total**
Multi-Racial	406	3.88	2551	4.42
Total	10467	100	57661	100

With Math IAB test takers we see a more pronounced trend. American Indian students comprised 16.99% of the Math Interim population, compared with 9.71% of the tested grade students that did not participate in the Interim assessments. Of the students that took the ICA in ELA, 20.60% compared to 18.21% of the Math ICA test takers were American Indian.

The distribution for gender is roughly the same for the ELA IAB, Math IAB, and the ICAs. Of the ELA IAB test-takers 50.78% were male compared with 51.29% of the comparison group. The ICA does show a slightly greater divide between the genders with female students comprising 45.69% of students who took the ELA ICA.

Results

English Language Arts (ELA)

There are two indicators of a student's score – the scale score and the achievement level. These scale scores fall on a continuous scale (ranging from approximately 2000 to 3000) that increases across grades. Students must fall into one of four achievement level reporting categories to describe student performance as determined by cut scores set at certain points on the scale to describe student proficiency (see Smarter Balanced Achievement Level Reporting Scores).

Table 7: Proficiency Levels for ELA and Math

Mathematics

Grade	Level 1	Level 2	Level 3	Level 4
	Novice	Nearing Proficient	Proficient	Advanced
3	<2381	2381–2435	2436–2500	>2500
4	<2411	2411–2484	2485–2548	>2548
5	<2455	2455–2527	2528–2578	>2578
6	<2473	2473–2551	2552–2609	>2609
7	<2484	2484–2566	2567–2634	>2634
8	<2504	2504–2585	2586–2652	>2652
9	<2517	2517–2600	2601–2675	>2675
10	<2533	2533–2613	2614–2696	>2696
11	<2543	2543–2627	2628–2717	>2717

English Language Arts/Literacy

Grade	Level 1	Level 2	Level 3	Level 4
	Novice	Nearing Proficient	Proficient	Advanced
3	<2367	2367–2431	2432–2489	>2489
4	<2416	2416–2472	2473–2532	>2532
5	<2442	2442–2501	2502–2581	>2581
6	<2457	2457–2530	2531–2617	>2617
7	<2479	2479–2551	2552–2648	>2648
8	<2487	2487–2566	2567–2667	>2667
9	<2489	2489–2570	2571–2671	>2671
10	<2491	2491–2576	2577–2677	>2677
11	<2493	2493–2582	2583–2681	>2681

This study analyzed scale scores since they are more nuanced. The scale score is framed so that it is comparable across assessments (Interim & Summative). These results are disaggregated by grade. The research questions for the IAB assessments were, "Do students who take at least one Interim

assessment and the Summative assessment experience growth in test scores over time?" and "Do students who take at least one Interim score higher than their peers who did not take an Interim assessment?"

A t-test is a statistical test that compares the significance of the difference between two means (averages). An important statistic to look at is the mean difference. A positive score indicates that from the time that a student took a certain Interim assessment to the Summative window there was growth in the average normalized scale score. When comparing groups on the same assessment, an Independent Sample t-test can provide knowledge of the direction and magnitude of any differences. The next step was to look at the significance of the t-test (Sig 2-tailed). Highlighted in yellow are those analyses whose significance level meets the $p<.05$ standard for measuring statistical significance. This significance level indicates that there is less than a 5% chance that the mean difference occurred by chance. Most of the findings addressed in this document are significant at the $p < .001$ level, indicating that there is a 0.1% chance that the mean difference occurred by chance.

To answer the first question, this analysis incorporated a paired sample t-test which can be used when a student has taken the same or two comparable assessments (normed in a similar manner). In this case, the comparison was between each student's Interim(s) and Summative assessment scaled score. For example, if a student took more than one Interim, each Interim score was compared to the Summative score. The Interims are normed to the framework of the Summative assessment.

All grade-level analyses for the paired sample t-tests for the ELA IAB population were significant. This means that the growth in test scores from the Interim to the Summative assessment by the same student were statistically significant. This growth found was meaningful, for example, Grade 4 (26.36), Grade 5 (25.09), and Grade 7 (23.36) show large gains (Table 8).

Table 8: ELA IAB Paired Sample T-tests

Grade	Test	Mean	N	Std. Deviation	Std. Error Mean	Mean Difference	T	df	Sig. (2-tailed)
3	Interim Scale Score	2414.67	5944	111.76	1.45	16.25	11.87	5943	0.000
3	Summative Scale Score	2430.92	5944	100.70	1.31				
4	Interim Scale Score	2451.78	5326	123.05	1.69	26.36	18.51	5325	0.000
4	Summative Scale Score	2478.14	5326	92.24	1.26				
5	Interim Scale Score	2488.16	5125	125.84	1.76	25.09	16.88	5124	0.000
5	Summative Scale Score	2513.25	5125	96.28	1.34				
6	Interim Scale Score	2530.77	4532	115.79	1.72	14.47	9.01	4531	0.000
6	Summative Scale Score	2545.24	4532	103.06	1.53				
7	Interim Scale Score	2536.52	2646	125.23	2.43	23.26	12.50	2645	0.000
7	Summative Scale Score	2559.78	2646	96.81	1.88				
8	Interim Scale Score	2566.91	1655	114.52	2.81	12.29	5.61	1654	0.000
8	Summative Scale Score	2579.20	1655	92.21	2.27				

It is important to also pay attention to the mean scores (not just the mean differences). From the mean scores, we can identify the achievement level in which the average score falls. By the Summative

assessment at each grade level, the average ELA IAB score was mostly in the proficient category, Grade 3 (Nearing Proficient), Grade 4 (Proficient), Grade 5 (Proficient), Grade 6 (Proficient), Grade 7 (Proficient), and Grade 8 (Proficient). The average score on the ELA IAB was in the 'Nearing Proficient' category before progressing into the 'Proficient' category on the Summative assessment.

Another way to look at the significance between two mean scores was to look at the correlations. A correlation measures a relationship between two or more things, in this case the mean scores on the ELA IAB and the paired Summative assessment. All paired grade-level samples see moderate correlations between test scores (the standard for a strong correlation is ≥0.7 and the standard for a moderate correlation is 0.5–0.69).

Table 9: ELA IAB Paired Sample Correlations

Grade	N	Correlation	Sig.
3	5944	0.51	0.000
4	5326	0.57	0.000
5	5125	0.57	0.000
6	4532	0.52	0.000
7	2646	0.66	0.000
8	1655	0.65	0.000

To answer the second question, "Do students who take at least one Interim score higher than their peers who did take an Interim assessment?" This study used an independent sample t-test, which is used when two different populations take the same assessment, in this case the Summative Assessment. The comparison was between the population participating in the ELA IAB or Math IAB assessments compared with the population not participating in the IAB assessments. The common metric is the Summative assessment scale score.

When answering the second research question, we analyzed whether students who took an ELA IAB scored significantly higher on the Summative assessment when compared with non-participating peers. If a student took more than one ELA IAB, they were counted only once. At all tested grades, the ELA IAB students scored significantly better. For example, grade 4 scored 30.03 points higher, Grade 7 33.10 points higher, and Grade 8 35.79 points higher (Table 10). To put this in context, that is nearly half a grade-level achievement-level category.

Table 10: Independent Sample Test (State Comparisons)

Grade	Population	N	Mean	Std. Deviation	Std. Error Mean	Mean Difference	t	Df	Sig. (2-tailed)
3	Interim ELA	2326	2427.42	106.34	2.20	21.55	6.08	11057	0.000
3	Did Not Participate	8733	2405.87	162.03	1.73				
4	Interim ELA	2359	2473.39	93.80	1.93	30.03	8.52	11507	0.000
4	Did Not Participate	9150	2443.35	164.36	1.72				
5	Interim ELA	2130	2512.96	100.97	2.19	23.93	6.59	11828	0.000
5	Did Not Participate	9700	2489.04	160.63	1.63				
6	Interim ELA	1575	2539.44	105.06	2.65	28.69	6.84	11466	0.000

Grade	Population	N	Mean	Std. Deviation	Std. Error Mean	Mean Difference	t	Df	Sig. (2-tailed)
6	Did Not Participate	9893	2510.74	161.21	1.62				
7	Interim ELA	1087	2567.07	94.27	2.86	33.10	6.55	11382	0.000
7	Did Not Participate	10297	2533.97	163.69	1.61				
8	Interim ELA	803	2577.18	91.69	3.24	35.79	6.21	10689	0.000
8	Did Not Participate	9888	2541.39	161.31	1.62				

On the Summative assessment, students who took the Interim assessment had mean test scores with the following achievement levels: Grade 3 (Nearing Proficient), Grade 4 (Proficient), Grade 5 (Proficient), Grade 6 (Proficient), Grade 7 (Proficient), Grade 8 (Proficient). Non-participating students had the following achievement levels based on average scores: Grade 3 (Nearing Proficient), Grade 4 (Nearing Proficient), Grade 5 (Nearing Proficient), Grade 6 (Nearing Proficient), Grade 7 (Nearing Proficient), and Grade 8 (Nearing Proficient).

Math

Modest gains in the mean difference between the Math IAB scale scores and the Summative assessment (Math) scaled scores are shown in Table 11. Five grade levels had significant increases in their average scale score. Two of those differences were substantial: Grade 5 (14.59) and Grade 6 (16.32). The grade 7 mean difference was not significant.

Table 11: Math IAB Paired Sample T-tests

Grade	Assessment	Mean	N	Std. Deviation	Std. Error Mean	Mean Difference	T	Df	Sig. (2-tailed)
3	Interim Scale Score	2431.58	9081	101.53	1.07	3.81	-3.55	9080	0.000
3	Summative Scale Score	2435.40	9081	92.89	0.97				
4	Interim Scale Score	2482.39	6988	102.90	1.23	5.75	-5.85	6987	0.000
4	Summative Scale Score	2488.14	6988	84.01	1.01				
5	Interim Scale Score	2494.36	6930	115.03	1.38	14.59	-13.55	6929	0.000
5	Summative Scale Score	2508.95	6930	96.16	1.16				
6	Interim Scale Score	2505.78	5121	115.62	1.62	16.32	-12.92	5120	0.000
6	Summative Scale Score	2522.10	5121	101.01	1.41				
7	Interim Scale Score	2561.14	3827	112.67	1.82	0.53	0.35	3826	0.727
7	Summative Scale Score	2560.61	3827	93.02	1.50				
8	Interim Scale Score	2561.08	2529	132.57	2.64	6.07	2.57	2528	0.010
8	Summative Scale Score	2555.00	2529	99.42	1.98				

Another way to look at the fit of a relationship was to look at the correlation (Table 12). For all grades except Grade 3, the correlation coefficients were moderate in size. This means the trends in Interim scale scores and Summative scale scores were moderately different.

Table 12: Math IAB Paired Sample Correlations

Grade	N	Correlation	Sig.
3	9081	0.45	0.000
4	6988	0.63	0.000
5	6930	0.65	0.000
6	5121	0.66	0.000
7	3827	0.60	0.000
8	2529	0.51	0.000

Even though Math IAB test takers did not show as substantial gains in their score on the Summative assessment as ELA IAB test takers, there were sizeable differences between those who took the Interim and those that did not. As shown in Table 13, there were large differences between the grade-level populations, for example, Grade 3 (26.80), Grade 4 (29.79), Grade 5 (29.85), and Grade 7 (39.68), indicating statistically significant findings on the Summative assessment for students who took at least one Math IAB.

Table 13: Independent Sample Tests (State Comparisons)

Grade	Math Assessment	N	Mean	Std. Deviation	Std. Error Mean	Mean Difference	T	Df	Sig. (2-tailed)
3	Interim Math	2691	2437.10	95.82	1.85	26.80	8.24	11012	0.000
3	Did Not Participate	8323	2410.30	159.66	1.75				
4	Interim Math	2510	2483.26	87.25	1.74	29.79	9.03	11430	0.000
4	Did Not Participate	8922	2453.47	158.60	1.68				
5	Interim Math	2389	2512.80	100.29	2.05	29.85	8.81	11819	0.000
5	Did Not Participate	9432	2482.95	157.72	1.62				
6	Interim Math	1725	2523.58	99.53	2.40	21.40	5.22	11420	0.000
6	Did Not Participate	9697	2502.18	164.85	1.67				
7	Interim Math	1305	2560.40	94.41	2.61	39.68	8.42	11336	0.000
7	Did Not Participate	10033	2520.72	166.72	1.66				
8	Interim Math	1023	2551.75	100.29	3.14	21.55	4.01	10601	0.000
8	Did Not Participate	9580	2530.20	168.92	1.73				

ICA

English Language Arts

There were no significant differences in ELA performance on the ELA ICA and their Summative assessment performance. Trends in the mean differences were modest and mixed (Table 14). This indicated that students scored similarly on the ELA ICA as on the Summative assessment. This finding was not surprising as the ICA are the full-length assessments most like the Summative assessment with the same blueprint as the Summative and full range of claim-, target-, and standard-level expectations.

The Summative assessment achievement levels for the mean scale score were Grade 3 (Novice), Grade 4 (Novice), Grade 5 (Nearing Proficiency), Grade 6 (Nearing Proficient), Grade 7 (Nearing Proficient), and Grade 8 (Nearing Proficient).

Table 14: ICA ELA Paired Sample T-tests

Grade	Assessment	Mean	N	Std. Deviation	Std. Error Mean	Mean Difference	t	Df	Sig. (2-tailed)
3	Interim Scale Score	2362.144	139	81.25	6.89	3.92	-0.78	138	0.438
3	Summative Scale Score	2366.065	139	93.23	7.91				
4	Interim Scale Score	2409.206	126	95.78	8.53	-0.53	0.11	125	0.913
4	Summative Scale Score	2408.675	126	90.99	8.11				
5	Interim Scale Score	2442.467	135	93.86	8.08	-0.25	0.05	134	0.959
5	Summative Scale Score	2442.215	135	103.34	8.89				
6	Interim Scale Score	2484.649	114	105.43	9.87	6.93	-1.45	113	0.150
6	Summative Scale Score	2491.579	114	95.79	8.97				
7	Interim Scale Score	2539.774	124	81.98	7.36	3.08	-0.63	123	0.531
7	Summative Scale Score	2542.855	124	86.95	7.81				
8	Interim Scale Score	2546.645	138	94.43	8.04	-3.70	0.68	137	0.499
8	Summative Scale Score	2542.942	138	101.79	8.66				

ICA Math

Trends were positive for the mean difference of ICA Math test takers. Increases in scores (ICA to Summative) range from 17.16 to 43.71. All analyses were significant at the p<.001 level, indicating a high degree of certainty that test scores dramatically improved. This was reflected in the achievement levels by the Summative assessment where, in some categories there was movement from Novice to Nearing Proficiency: Grade 3 (Nearing Proficiency), Grade 4 (Nearing Proficiency), Grade 5 (Nearing Proficiency), Grade 6 (Nearing Proficient), Grade 7 (Nearing Proficiency), and Grade 8 (Nearing Proficiency).

For example, Lame Deer School District experienced strong gains in terms of the mean difference between assessments. Grade 3 (53.15), Grade 4 (60.89), and Grade 6 (31.35) experienced increases in scale scores that were significant at the p < .001 level. Stated another way, Lame Deer students in Grades 3-4 and 6, progressed at least half a proficiency level over the short two-month period between the Interim assessment and the Summative assessment.

Table 15: ICA Math Paired Sample T-tests

Grade	Assessment	Mean	N	Std. Deviation	Std. Error Mean	Mean Difference	t	Df	Sig. (2-tailed)
3	Interim Scale Score	2340.691	136	66.00	5.66	43.71	-10.64	135	0.000
3	Summative Scale Score	2384.397	136	72.55	6.22				
4	Interim Scale Score	2412.08	125	91.95	8.22	30.03	-6.74	124	0.000
4	Summative Scale Score	2442.112	125	79.94	7.15				

Grade	Assessment	Mean	N	Std. Deviation	Std. Error Mean	Mean Difference	t	Df	Sig. (2-tailed)
5	Interim Scale Score	2440.69	142	79.53	6.67	17.17	-3.76	141	0.000
5	Summative Scale Score	2457.859	142	91.54	7.68				
6	Interim Scale Score	2511.205	161	109.65	8.64	17.16	-4.38	160	0.000
6	Summative Scale Score	2528.36	161	114.62	9.03				
7	Interim Scale Score	2548.016	184	78.14	5.76	21.27	-5.58	183	0.000
7	Summative Scale Score	2569.288	184	90.13	6.64				
8	Interim Scale Score	2523.028	211	101.67	7.00	27.46	-6.71	210	0.000
8	Summative Scale Score	2550.493	211	100.78	6.94				

Conclusion

These findings indicate that overall, Interim assessment participation has a positive impact on student performance on the Summative assessment. Interim assessment use has grown for both ELA and math for all tested grades over the past four years. During the 2018–2019 school year, more Interim assessments were taken by students qualifying for free and reduced lunch than their non-participating peers. Since we did not control for economic disadvantage in these analyses, the findings regarding this subgroup are even more remarkable.

The most significant impact can be seen by students who took the ELA IAB assessments. Students who took at least one Interim assessment outperformed on the Summative assessment students that did not take Interim assessments for all tested grades (Table 8). In addition, students who took Interim assessments showed statistically significant personal growth from their score on the Interim assessment(s) and their summative score. This finding may indicate that teachers are making data-driven decisions to support their classroom instructional practices and may be adjusting instruction to align to the rigorous expectations of our Montana Content Standards as measured by both the Interim and Summative assessments.

In a sense, the Interim assessments is both a measure and a treatment showing that students who took Interim assessments outperformed their non-participating peers on nearly every measure as well as showing individual student growth in achievement from the Interims to the Summative. Further research may investigate any impact on Summative assessment scores resulting from or related to the number of Interim assessments a student takes. As more school districts implement Interim assessments, further analysis is recommended to see if these trends remain consistent across larger and more diverse student populations.