



THE SCHOOL DISTRICT OF
PHILADELPHIA

Assessing Student Performance Before and During Virtual Learning: A Cohort Comparison of Student Performance on 2019-20 Winter and 2020-21 Fall aimswebPlus and Star Assessments

This report compares performance in reading and math assessments for the same set of students over time. The main metric used in this report is the National Percentile Rank (NPR). NPR is a norm-referenced measure that compares student performance to a national sample of students.

Key Findings

- On average, decline in student National Percentile Rank (NPR) performance from 2019-20 Winter (in-person instruction) to 2020-21 Fall (virtual instruction) was small. The overall average decline was highest (6 percentiles) for aimswebPlus Reading (grades 1-5).
- Students who were in kindergarten in 2019-20 and in first grade in 2020-21 had the largest decline in Winter to next-Fall performance.
- Overall, the decline in performance from Winter to next-Fall represents a mostly equal decline across students of different demographic subgroups.
- Although all demographic groups experienced Winter to next-Fall declines equally, year-to-year comparisons show that existing overall performance differences across demographic groups persisted.

Ebru Erdem,
Policy Research Analyst

Joy Lesnick,
Deputy Chief

Office of Evaluation,
Research, and
Accountability

January 2021

Contents

Overview: Assessments at the School District of Philadelphia.....	3
Required Formative Assessments in SDP.....	3
Assessing Student Learning During the 100% Online Education Period.....	3
Research Questions.....	4
Methods.....	5
Metrics.....	5
Determining Comparable Tests	6
Determining the Cohort	7
Findings.....	8
AimswestPlus Reading	8
Analysis of Winter to Next-Fall Changes in Average National Percentile Rank: 2019-20 to 2020-21 and 2018-19 to 2019-20	8
Star.....	15
Star Reading: Analysis of 2019-20 Winter to 2020-21 Fall Changes in National Percentile Rank	15
Star Math: Analysis of 2019-20 Winter to 2020-21 Fall Changes in National Percentile Rank..	17
Star Reading and Math Test Duration Analysis.....	20
Conclusions.....	22
Appendix.....	23
More about AimswestPlus Tests at the District.....	23
More about Star Tests at the District.....	23
Comparable Tests and Cohorts.....	24

Keeping Track of School Years and Testing Windows

Looking across different testing windows in different school years can be confusing to understand. In this report, we always refer to the school year first, followed by the testing window (e.g. 2020-21 Fall).

Figure 1. A Visual Map of Multi-Year Assessment Performance Analyses



The two diagonal arrows (labeled A and B) help illustrate the analyses presented in this report. First, we analyze student performance from 2019-20 Winter to 2020-21 Fall (arrow B) and then compare the change in performance to the previous year for aimswebPlus Reading (arrow A). The vertical arrow represents a Fall-to-Fall comparison of student performance published in a previous report (see footnote 6).

Overview: Assessments at the School District of Philadelphia

Required Formative Assessments in SDP

Beginning in the 2020-21 school year, the School District of Philadelphia (SDP) required that formative assessments in reading and math be administered in all schools. Formative assessments administered throughout the school year allow for monitoring student progress and growth as well as developing and targeting interventions where necessary in a timely manner. The aimswebPlus assessment is required for grades K-5 and the Star assessment is required for grades 6-12.^{1, 2}

Assessing Student Learning During the 100% Online Education Period

Students in grades K-5 were scheduled to take the aimswebPlus assessments for their grade level three times during the 2019-20 school year (Fall, Winter, and Spring). Students in grades 6-12 were scheduled to take the Star assessment four times during the 2019-20 school year (Term 1, Term 2, Term 3, and Term 4).³ However, the schedule of planned formative assessments during the 2019-20 school year was disrupted by the switch to online learning in Spring 2020. In line with Covid-19 containment efforts in Philadelphia, all SDP school buildings closed on March 13, 2020. By that point, Fall and Winter aimswebPlus assessments and Terms 1, 2, and 3 Star assessments were completed. Although SDP implemented online instruction for the remainder of the school year, no formative assessments were administered for Spring/Term 4, and all state-required standardized tests were also cancelled.⁴

SDP started the 2020-21 school year fully online and has been continuing online instruction as of January 2020. During this period, 2020-21 Fall assessments were completed between September 14 and October 23, 2020. The winter assessment window dates were December 14, 2020, to January 29, 2021, and the window includes an 11-day winter break.⁵ Fall and Winter 2020-21 assessment windows for both the aimswebPlus and Star assessments were the same.

The pivot to 100% online learning was unprecedented, sudden, and unexpected. During a short period of time, SDP worked to provide students with computers and internet access, and to provide teachers with professional development about remote instruction. Students and families also had to

¹ As of 2020-21 Fall special education students who are not exempted based on their IEPs take aimswebPlus Reading and Math in grades K-5, both aimswebPlus and Star tests in reading and math in grades 6-8, and Star Reading and Math in grades 9-12.

² SDP had been using AIMSweb reading as a required early literacy assessment for K-3 since 2014-15 Spring and switched to aimswebPlus in the 2018-19 school year. Starting in 2019-20 Fall, aimswebPlus Reading became required for K-5 and AimswebPlus Math became available but was not a required assessment for grades K-5. Star was introduced in grades 6-10 during the 2019-20 school year and became required for grades 6-12 starting 2020-21 Fall. See Appendix for more details on aimswebPlus and Star assessments.

³ Throughout the report, in analyses that refer to 2019-20 Winter, the testing windows are as follows: aimswebPlus (K-5)-January 2-31, 2020; Star (6-8)-Term 3, March 2-12, 2020; Star (9-12)-Term 2, February 5-21, 2020.

⁴ See [Evaluation of the Spring 2020 Covid-19 Continuity of Education Plan](#) for details on the Continuity of Education plan for 2019-20 Spring and [Assessment Calendar for 2020-21](#) for details on the assessment windows for 2020-21 school year.

⁵ Before Fall 2020, SDP provided additional weeks on either end of the official testing window to capture testing data from students who may have had barriers to completing their assessment(s) during the District-specified testing window. For example, for Fall 2019, the expanded window was from September 3 to October 31, 2019. The data in this report reflect performance during the expanded testing windows for assessments before the 2020-21 school year. Beginning in the 2020-21 school year, all testing windows are strict and all testing has to be completed within the window announced in the assessment calendar.

work through obstacles beyond technology and access to establish remote learning environments for students. All these factors may have had effects on student learning experiences and test performance.

Challenges experienced by families and students throughout online learning also affected the administration of assessments. In addition, there were challenges specific to remote administration of the tests as well. In in-person/at-school settings, both aimswebPlus and Star assessments would normally be administered at school, under the guidance of a teacher. In addition, some aimswebPlus tests would typically have been administered individually by the teacher. Under online learning conditions, all assessments were administered online, with different levels of parent or guardian guidance and/or intervention. Tests that were typically administered by the teacher in-person were administered over individual online meetings with each student.

Research Questions

An Office of Evaluation, Research, and Accountability report examined the differences in student participation and tier performance between 2019-20 Fall and 2020-21 Fall to study how participation and performance in testing under online learning conditions differed from in-person, in-school testing at the same window the year before.⁶ The results presented in this report provide more nuance to that analysis by examining student performance on their last in-person/at-school test (2019-20 Winter) compared to their first online/at-home test (2020-21 Fall).

Two primary research questions guided this analysis:

1. What were the differences in student performance, as measured by National Percentile Rank, between 2019-20 Winter and 2020-21 Fall for aimswebPlus Reading (2020-21 grade 1-5 students) and Star Reading and Math (2020-21 grade 7-11 students)?
 - 1.a. Were there identifiable patterns in aimswebPlus Reading and Star Reading and Math performance differentials between 2019-20 Winter and 2020-21 Fall with respect to grade level, race/ethnicity, English learner status, 2020-21 Fall performance tier on the test, and Economic Disadvantage?
2. How did the change in aimswebPlus Reading performance between 2019-20 Winter and 2020-21 Fall compare to the change in student performance on the same assessments the previous school year (between 2018-19 Winter and 2019-20 Fall)?

A third exploratory question also provides information about the time students take to complete the assessments:

3. For Star Reading and Star Math, did the time it took students to complete the test differ between 2019-20 Winter, which was administered in-person/at-school and 2020-21 Fall, which was administered online/at-home?

By comparing the performance of the same students in 2019-20 Winter and 2020-21 Fall, we can examine the effects of multiple factors. First, some students included in this analysis have

⁶ See: [K-12 Student Participation and Performance on aimswebPlus and Star Assessments, Fall 2019 and Fall 2020](#), December 2020.

experienced “summer learning loss,” a decline in academic skills, especially early literacy skills, over the summer months.⁷ Existing literature suggests that some decline from the previous school year is expected for Fall assessments. The analysis here captures potential summer learning loss as well as the possible effects of both online learning since March 13, 2020, and online/at-home test administration in 2020-21 Fall.

Although separating out the effects of these factors is not possible for all tests, the cohort analysis of aimswebPlus and Star performance in this report provides valuable insights about changes in student performance before and during online instruction.

Methods

This report compares performance in reading and math assessments for the same set of students over time. AimswebPlus Reading test results are used for elementary grades (K-5) and Star Reading and Math test results are used for middle and high school grades (6-12).

Metrics

National Percentile Rank (NPR)

The main metric used in this report is the National Percentile Rank (NPR). NPR is a norm-referenced measure that compares students’ raw score (number correct) to a national sample of students.⁸ Students who perform better score a higher percentile rank, with the 100th percentile being the best performance. In the analyses in this report, the NPR for individual students are averaged for the student groups analyzed. The analyses below report the average NPR⁹ that the identified groups of students scored in the comparable tests as well as the change in average NPR for the same students between Winter and next-Fall.^{10, 11}

The previous ERA report (see fn. 6) compared student tier performance in Fall windows of 2019-20 and 2020-21 school years. Performance tiers are an alternative way of analyzing student performance and are valuable in determining necessary instructional interventions. NPR is related to the aimswebPlus performance tiers SDP uses to target interventions because both are based on the number correct on the tests. For aimswebPlus, students are placed into three performance tiers: Tier 1 (At Target), Tier 2 (Strategic Intervention), or Tier 3 (Intensive Intervention), based on

⁷ See: [A Four-Year Summary of “Summer” Learning Loss: Changes in K-2 Independent Reading Levels from June to November, 2016-2019](#), June 2020.

⁸ NPR is provided by the vendor and each vendor calculates it based on the national sample of test takers.

⁹ NPR only shows performance of individual students relative to all other test takers. When averaged, it shows how a group of students performed compared to all other test takers. NPR does not measure student growth over time; measures such as Student Growth Percentile (SGP) are utilized for understanding student growth.

¹⁰ Sometimes a student transfers from one District school to another within the testing window and they are tested in each school. There may also be situations where a student is tested more than once within the same school. In cases where a student is tested multiple times within the same window for the same test, the latest best score is included in the analyses and any other test results are disregarded.

¹¹ Starting in Fall 2020-21, aimswebPlus Reading and Math tests were also administered in Spanish, especially in dual language programs. Spanish test takers are required to take the English language test as well and the English test is prioritized. Spanish tests are not included in the analyses below; throughout the report, only English language tests were included. Star is currently only available in English.

number correct.^{12, 13} Similarly, Star has four tiers, based on NPR: At/Above Benchmark for 40th percentile and above, On Watch for 25th to 29th percentiles, Strategic Intervention for 10th to 24th percentiles, and Intensive Intervention for 9th percentile and below. This report uses NPR because (1) NPR provides a more detailed look at student performance at different points of the year relative to a national sample, and (2) is on a 100-point scale so we can see patterns of change in performance across different student groups.

Time to Completion (TTC)

Another metric used in this report is average Time to Completion (TTC), and it is currently available only for Star assessments. This metric shows the average duration of tests in minutes; that is, how long it took students, on average, to complete the assessment once they start the session. Time to Completion is not a measure of performance; it is used here as a metric to study differences in test administration in in-person/at-school and online/at-home settings.

Determining Comparable Tests

The report does not include an analysis of student performance on the aimswebPlus Math assessments, because these tests became required only in 2020-21 Fall. Although they were available before 2020-21 Fall, participation was low because they were not required. Low and selective participation in aimswebPlus Math before 2020-21 Fall prohibits any meaningful comparative analysis.

AimswebPlus Reading (formerly referred to as AIMSweb/aimswebPlus literacy) has a longer history in the District, allowing for comparisons including earlier years. This report primarily compares, as with Star, aimswebPlus Reading performance in 2019-20 Winter and 2020-21 Fall. However, a comparison of 2018-19 Winter and 2019-20 Fall aimswebPlus Reading performance is also presented. This comparison gives us a baseline about how aimswebPlus Reading performance is expected to change from Winter of a year to the Fall of the next year under normal circumstances, without Covid-19 induced conditions such as online instruction and test administration. Comparing the results from this 2018-19 Winter to 2019-20 Fall comparison with the 2019-20 Winter to 2020-21 Fall comparison allows us to isolate the “summer learning loss” effect and understand the impact of online instruction on reading assessment performance.

Although aimswebPlus Reading data is available for earlier years, testing requirements for this test have changed over time. In addition to the changes in overall District testing requirements, the specific aimswebPlus Reading tests required for different grade levels changed as well. Because reading skills change dramatically from kindergarten to second grade, administered tests are not standard across grade levels. This presents an additional challenge in determining comparable tests across cohorts for aimswebPlus Reading.

¹² For aimswebPlus, tiers are criterion-referenced and cut-points in each window are determined according to the likelihood of meeting Spring/end-of-year performance targets, measured in terms of NPR.

¹³ Based on performance on Star, students within Tier 1 receive the core curriculum, while Tier 2 students receive additional small group interventions, and Tier 3 students receive more intensive interventions.

For the analyses below, NPR based on Nonsense Word Fluency (NWF) performance was used for the comparisons of reading performance of students who were in kindergarten in Winter and first grade in the next Fall (Table 1 and Table A.1). All comparisons for students who were first to fourth grade in Winter and second to fifth grade in the next Fall are based on Oral Reading Fluency (ORF) NPR performance.¹⁴

Table 1. Comparable Tests

Test	Possible Comparisons
aimswebPlus Reading	NPR based on NWF for grades K-1 and ORF for grades 1-5
aimswebPlus Math	No comparisons possible
Star Reading	2020-21 Grades 7-11 (2019-20 Grades 6-10), NPR based on composite score
Star Math	2020-21 Grades 7-11 (2019-20 Grades 6-10), NPR based on composite score

Note: See Appendix Table A.1. for more details about the cohorts and comparable aimswebPlus Reading tests.

For Star Reading and Math, all test takers receive a composite score, which has been consistent between 2019-20 and 2020-21 school years (Table 1). The only difference in Star administration between the two years is that it became a required assessment for grades 11 and 12 only in 2020-21 Fall. In this report, we compare the National Percentile Rank based on the composite Star Reading and Math scores only for students who were in sixth to tenth grade in 2019-20 and exclude the grade levels for which the assessment was not required.

Determining the Cohort

Students who received a score during the 2019-20 Winter testing window on the specific tests identified in the previous section were determined to be the base cohort for this analysis (see Appendix A for more details). These students were most likely enrolled in the next higher grade level in 2020-21 Fall.¹⁵ Students who completed the test in both 2019-20 Winter and 2020-21 Fall testing windows are the cohort of students used for this analysis.

This report analyzes test results for students who took the tests only.¹⁶ As of the 2020-21 school year, only students who have completed all of the component tests required to get a composite score are considered participants. By contrast, in this report, students who have scores in the comparable tests identified above are included in the analyses, even if they may not have a composite score.¹⁷

The 2019-20 Winter testing window (see fn. 3) was the last window in which assessments were administered before District school buildings were closed on March 13, 2020, due to Covid-19

¹⁴ See the Appendix for more information about the comparable aimswebPlus Reading tests for the grade levels.

¹⁵ The numbers of students included in the analyses and basic descriptives about the cohort are included at the beginning of aimswebPlus Reading and Star Reading and Math sections below.

¹⁶ At each testing window, there are students who did not take the assessments although they were enrolled at a District school and thus were eligible to take the test. They are not included in the report because there were no performance outcomes for them.

¹⁷ Due to this discrepancy, the students who are included in these analyses do not overlap perfectly with students who are counted as participants in other District reports and QlikBAM dashboards where participation is based on composite scores.

containment efforts in Philadelphia. Among the 2019-20 Winter test-takers, those who took a comparable test in the 2020-21 Fall testing window were identified. For each of the three tests analyzed in this report, the number of identified students (that is, students with results during both windows) varies.¹⁸

Findings

The results from the analyses for student performance on aimswebPlus and Star are presented separately below. First, the comparison of 2020-21 Fall and 2019-20 Winter aimswebPlus Reading performance is presented, including the performance of student subgroups. In the same section we address Research Question 2, comparing Winter to next-Fall performance differentials in two consecutive years for the same base cohort of 2019-20 Winter aimswebPlus test takers. In the next section, the 2020-21 Fall and 2019-20 Winter comparison and subgroup analyses are repeated for both Star Reading and Star Math. In addition, the section on Star also presents a test duration analysis for 2020-21 Fall and 2019-20 Winter assessment windows.

AimswebPlus Reading

Analysis of Winter to Next-Fall Changes in Average National Percentile Rank: 2019-20 to 2020-21 and 2018-19 to 2019-20

There were 41,145 grade 1-5 students in 2020-21 who took the aimswebPlus reading test in 2019-20 Winter and 2020-21 Fall. Of these students, 20,160 were female and 20,985 were male students. A great majority (85%) of them were non-English Learners (35,080) and 6,065 were English Learners. Black/African American students constituted the highest number in this cohort with 18,775 students (46%), followed by 10,108 Hispanic/Latino, 6,586 White, 3,909 Asian, 1,653 Multiracial, 59 Native Hawaiian/Pacific Islander, and 55 American Indian/Alaskan Native students. Consistent with District averages, 30,562 of the students were Economically Disadvantaged and 10,568 were not.

Winter to Next-Fall Changes by Grade Level

Table 2 shows the average NPR for students in grades 1 to 5 in 2020-21, who were enrolled in grades K to 4 in 2019-20. It presents average NPR for 2019-20 Winter, 2020-21 Fall and the difference between the two. The overall average decline in NPR between the two testing periods was 6.37 percentiles. While the average NPR declined about 4-5 percentiles for grades 2 to 5, students who went from kindergarten to first grade experienced a larger drop in average NPR (15 percentiles).

¹⁸ These figures are reported in the Totals row of the Number of Students Columns in each table in the aimswebPlus Reading, Star Reading, and Star Math sections.

Table 2. aimswebPlus Reading National Percentile Rank by Grade Level, 2019-20 Winter to 2020-21 Fall

Grade Level in 2020-21	Number of students	Average 2019-20 Winter NPR*	Average 2020-21 Fall NPR**	Average NPR Differential
1	7439	42.35	27.86	-14.49
2	8601	31.68	28.01	-3.67
3	8690	32.02	27.05	-4.97
4	8487	31.32	26.06	-5.25
5	7928	31.96	27.55	-4.41
Total	41145	33.66	27.29	-6.37

* Last assessment window before Covid-19 school closure

** First assessment administered during Covid-19 school closure

How to read this table: The table shows performance in 2019-20 Winter and 2020-21 Fall for each grade level and the change from Winter to next-Fall. For example, there were 7,439 2020-21 first graders, who took an aimswebPlus Reading test who also took the comparable test in the 2019-20 Winter window, when they were kindergarteners. These students' average NPR was 42nd percentile in 2019-20 Winter and 28th percentile in 2020-21 Fall. The change from Winter to next-Fall for 2020-21 first graders was 15 percentiles.

In order to examine how much of the decline in NPR across grade levels observed in Table 2 is due to the change to digital learning as a result of Covid-19, Table 3 presents a retrospective analysis of the past performance of the 19-20 Winter test takers included in the analysis in Table 2. Studying the patterns for the changes in average NPR between 2018-19 Winter and 2019-20 Fall, where instruction and testing happened completely in person and at school, can show us whether the decline in Table 2 is expected or due to online learning conditions.

Comparing the overall average NPR for Fall of the two years, we see that students scored lower in 2020-21 Fall overall (Tables 2 and 3). While the corresponding overall average NPR for Winter assessments held steady at the 33rd percentile, the overall average NPR was the 32nd percentile in 2019-20 Fall and the 27th in 2020-21 Fall, a difference of 5 percentiles.¹⁹

Some Winter to next-Fall decline in NPR is not uncommon: overall average NPR declined 2 percentiles between 2018-19 Winter and 2019-20 Fall, when K-5 students were attending school in-person (Table 3). Table 3 also shows that students who were kindergarteners in Winter and first-graders in the next-Fall experience a larger decline compared to other grades. However, the Winter to next-Fall decline for first-graders in Fall 2019-20 was 7 percentiles compared to 15 percentiles for 2020-21 Fall first-graders. In addition, for higher grade levels, the Winter to next-Fall decline ranged from 0 (no decline) to 2 percentiles between 2018-19 Winter and 2019-20 Fall, a much smaller decline than the 4-5 percentiles from 2019-20 Winter to 2020-21 Fall.

To compare cohort performance across years, we compare the results of Tables 2 and 3, and see Winter to next-Fall NPR declines are largely limited to the kindergarten to first grade transition. In

¹⁹ [Our previous report](#) (see fn. 6) found that 24.6% of 2020-21 first graders placed in Tier 1 based on the composite aimswebPlus score in the fall, whereas 33.9% of 2019-20 first graders placed in Tier 1 in the previous fall (p.14). Comparing the Fall average NPR for first-graders in Table 1 and 2, 2020-21 first graders scored at 28th percentile on average while 2019-20 first graders scored at 36th percentile on average. This is a decline of 8 percentiles. By comparison, slightly higher percentage of fourth- and fifth-grade students placed at Tier 1 in 2020-21 Fall than 2019-20 Fall. Fall to Fall comparison of the average NPR for these grade levels (not the longitudinal cohort) show only a 4-percentile decline.

these two consecutive school years, the cohort that experienced a Winter to next-Fall NPR decline from kindergarten to first grade did not experience a decline as large when they moved from first to second grade. More specifically, the first graders in 2019-20 Fall, who showed a 7-percentile decline in average NPR from their kindergarten year performance in 2018-19 Winter, experienced only a 4-percentile decline as second graders in 2020-21 Fall from 2019-20 Winter. Despite spending Spring of their first-grade year and through the Fall testing window in their second-grade year in online learning, their Winter to next-Fall average NPR decline was smaller than the previous year when they went from kindergarten to first grade, all under in-person instruction.

Table 3. aimswebPlus Reading National Percentile Rank by Grade level 2018-19 Winter to 2019-20 Fall

Grade Level in 2019-20	Number of students	Average 2018-19 Winter NPR	Average 2019-20 Fall NPR	Average NPR Differential
1	7604	42.76	36.25	-6.51
2	8695	30.85	30.87	0.02
3	8706	31.87	30.15	-1.72
4	8666	30.84	29.45	-1.39
5	7721	31.86	31.76	-0.10
Total	41392	33.44	31.57	-1.87

Winter to Next-Fall Changes by Race/Ethnicity

In the section above, tables 2 and 3 showed how the overall 6 percentile decline from 2019-20 Winter to 2020-21 Fall and the overall 2 percentile decline from 2018-19 Winter to 2019-20 Fall were distributed across grade levels. Examining the 2019-20 Winter to 2020-21 Fall average NPR changes across racial/ethnic groups shows that the overall decline in average NPR was similar across all groups, except for Native Hawaiian/Pacific Islander students. There were only 59 Native Hawaiian/Pacific Islanders in the sample and these students had a 9 percentile decline, compared to other groups whose decline was in the 5-7 percentile range (Table 4).

Table 4. aimswebPlus Reading National Percentile Rank by Race/Ethnicity, 2019-20 Winter to 2020-21 Fall

Race/Ethnicity	Number of students	Average 2019-20 Winter NPR	Average 2020-21 Fall NPR	Average NPR Differential
American Indian/Alaskan Native	55	30.36	23.58	-6.78
Asian	3909	47.22	40.88	-6.34
Black/African American	18775	30.11	23.17	-6.94
Hispanic/Latino	10108	25.69	20.50	-5.19
Multiracial/Other	1653	40.67	34.06	-6.61
Native Hawaiian/Pacific Islander	59	39.71	30.49	-9.22
White	6586	46.17	39.69	-6.48
Total	41145	33.66	27.29	-6.37

Racial/ethnic groups contribute about equally to the overall 2-percentile decline from 2018-19 Winter to 2019-20 Fall; NPR declines ranged from 0.5 to 3 percentiles for all groups except Native Hawaiian/Pacific Islander students. The decline for the 70 Native Hawaiian/Pacific Islander students was higher from 2018-19 Winter to 2019-20 Fall (4 percentiles) (Table 5).

Table 5. aimswebPlus Reading National Percentile Rank by Race/Ethnicity, 18-19 to 19-20

Race/Ethnicity	Number of Students	Average 2018-19 Winter NPR	Average 2019-20 Fall NPR	Average NPR Differential
American Indian/Alaskan Native	53	26.92	26.38	-0.55
Asian	3620	46.38	45.28	-1.10
Black/African American	19201	30.31	27.85	-2.45
Hispanic/Latino	10164	25.66	25.16	-0.50
Multiracial/Other	1891	38.61	36.49	-2.12
Native Hawaiian/Pacific Islander	70	38.51	34.10	-4.41
White	6393	46.36	43.75	-2.61
Total	41392	33.44	31.57	-1.87

The comparison of Winter to next-Fall average NPR performance shows mostly equal decline across students of different racial/ethnic groups, and a higher decline under online learning for all groups (Tables 4 and 5). However, *the year-to-year comparison also shows the persistence of differences across racial/ethnic groups in overall performance as measured by NPR*. Asian and White students in this sample consistently scored at or above 40th percentile. Black/African American students in this sample consistently scored within the 20-30th average NPR range. Hispanic/Latino students in this sample scored lower, at the 20-26th average NPR rate.

Winter to Next-Fall Changes by English Learner Status:

AimswebPlus Reading average NPR performance from 2019-20 Winter to 2020-21 Fall by English Learner status is examined in Table 6. Interestingly, the Winter to next-Fall decline during this period of online learning was smaller for English Learners (3 percentiles) compared to non-English Learners (7 percentiles). However, the average NPR for English Learner students was at least 10 percentiles lower than non-English Learners, despite higher decline for non-English Learners.

Table 6. aimswebPlus Reading National Percentile Rank by English Learner Status, 2019-20 Winter to 2020-21 Fall

English Learner Status	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Fall NPR	Average NPR Differential
Non-English Learner	35080	35.77	28.89	-6.88
English Learner	6065	21.46	18.04	-3.42
Total	41145	33.66	27.29	-6.37

Note: English Learner status as of Fall 2020-21

The pattern of overall lower performance by English Learners but a smaller decline in performance from Winter to next-Fall when compared to non-English Learners holds for the in-person instruction and testing period results of 2018-19 Winter to 2019-20 Fall (Table 7). In fact, English Learners did not show a decline from Winter to next-Fall in this period, while non-English Learners showed a 2 percentile decline. However, the overall reading performance of English Learners (21st percentile on average) was lower than non-English learners' (33-35th percentiles on average).

Table 7. aimswebPlus Reading National Percentile Rank by EL Status, 2018-19 Winter 2018-19 to 2019-20 Fall

English Learner Status	Number of Students	Average 2018-19 Winter NPR	Average 2019-20 Fall NPR	Average NPR Differential
Non-English Learner	35681	35.42	33.24	-2.18
English Learner	5711	21.04	21.17	0.13
Total	41392	33.44	31.57	-1.87

Note: English Learner status as of Fall 2019-20.

Winter to Next-Fall Changes by Fall Performance Tiers

As mentioned above, SDP uses performance tiers based on the number of correct answers on aimswebPlus for targeting necessary interventions. In order to understand whether students who were at different performance levels experienced Winter to next-Fall performance changes differently, Tables 8 and 9 compare the Winter to next-Fall change in average NPR by the performance tiers the students were placed in based on their performance in the respective Fall testing. For example, Table 8 shows the change from 2019-20 Winter to 2020-21 Fall for students who placed in the three performance tiers based on their aimswebPlus Reading scores in the 2020-21 Fall window for NWF for first grade and ORF in other grades. Because performance tiers are indirectly associated with NPR, it is expected that students in Tier 1 have a higher average NPR and those in Tier 3 have a lower average NPR.

The change in average NPR from Winter to next-Fall varies a lot by tier (Tables 8 and 9). Students who placed in the top tier (Tier 1) in 2020-21 Fall experienced the largest decline, on average, from 66th percentile in 2019-20 Winter to 52nd percentile in 2020-21 Fall. This 14-percentile decline is higher than the decline students who placed in Tier 2 showed (5 percentiles), while Tier 3 students held steady (Table 8).

Table 8. aimswebPlus Reading National Percentile Rank by 2020-21 Fall Performance Tiers, 2019-20 Winter to 2020-21 Fall

2020-21 Fall Performance Tier	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Fall NPR	Average NPR Differential
Tier 1	16064	66.05	51.84	-14.21
Tier 2	7105	28.67	23.42	-5.24
Tier 3	17976	6.69	6.88	0.19
Total	41145	33.66	27.29	-6.37

How to read this table: This table includes 41,145 students who were in grades 1-5 in 2020-21 Fall, in grades K-4 in 2019-20 Winter and took comparable tests in both testing windows. These students are grouped according to the performance tiers they placed in based on their 2020-21 Fall performance and average NPR were calculated for the three groups for the two testing windows. According to the table, 16,064 students placed in Tier 1 in 2020-21 Fall and these students, on average, scored at 66th percentile in 2019-20 Winter and at 52nd percentile in 2020-21 Fall. These represent, on average, a 14-percentile decline between the two windows.

Under typical circumstances, some Winter to next-Fall decline in average NPR is expected for Tier 1 students. Students who placed in Tier 1 based on their 2019-20 Fall aimswebPlus Reading scores showed an 8-percentile decline from their average NPR performance from 2018-19 Winter (Table 9). In 2020-21 Fall, Tier 2 students showed a 5-percentile Winter to next-Fall decline under online learning. In comparison, students who placed in Tier 2 in 2019-20 Fall did not show a Winter to next-Fall change. During in-person instruction, Tier 3 students in the sample improved on NPR, on average, by 2 percentiles.

Table 9. aimswebPlus Reading National Percentile Rank by 2019-20 Fall Performance Tiers, 2018-19 Winter to 2019-20 Fall

2019-20 Fall Performance Tier	Number of Students	Average 2018-19 Winter NPR	Average 2019-20 Fall NPR	Average NPR Differential
Tier 1	16022	65.68	57.95	-7.72
Tier 2	7200	28.61	29.00	0.39
Tier 3	18170	6.93	9.33	2.41
Total	41392	33.44	31.57	-1.87

How to read this table: This table includes 41,392 students who were in grades 1-5 in 2019-20 Fall, in grades K-4 in 2018-19 Winter and took comparable tests in both testing windows. These students are grouped according to the performance tiers they placed in based on their 2019-20 Fall performance and average NPR were calculated for the three groups for the two testing windows. According to the table, 16,022 students placed in Tier 1 in 2020-21 Fall and these students, on average, scored at 66th percentile in 2018-19 Winter and at 58th percentile in 2019-20 Fall. These represent, on average, a 8 percentile decline between the two windows.

Winter to Next-Fall Changes by Economic Disadvantage:

The same analysis is repeated with respect to economic disadvantage status in Tables 10 and 11. Results are similar to the results with other demographic breakouts: some decline is expected and the decline is similar between the subgroups but higher under online instruction compared to the decline under in-person instruction. Looking across the results presented in Tables 10 and 11 shows the persistence of differences in average NPR in Reading scores between economically disadvantaged and advantaged students. In all testing windows studied, students who are not economically disadvantaged perform, on average, about 13 percentiles higher than the

economically disadvantaged students. One relatively positive finding might be that the performance difference between the two groups has not increased during the period of online instruction.

Table 10. aimswebPlus Reading National Percentile Rank by Economic Disadvantage, 2019-20 Winter to 2020-21 Fall

Economic Disadvantage Status	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Fall NPR	Average NPR Differential
Economically Disadvantaged	30562	30.09	23.74	-6.36
Not Economically Disadvantaged	10568	43.99	37.57	-6.41
Total	41145	33.66	27.29	-6.37

Note: There were 15 students whose economic disadvantage status was not known. They are included in the counts; however, data on their performance are not reported.

Table 11. aimswebPlus Reading National Percentile Rank by Economic Disadvantage, 2018-19 Winter to 2019-20 Fall

Economic Disadvantage Status	Number of Students	Average 2018-19 Winter NPR	Average 2019-20 Fall NPR	Average NPR Differential
Economically Disadvantaged	31299	29.93	28.34	-1.59
Not Economically Disadvantaged	9939	44.57	41.83	-2.74
Status Unknown	154	29.02	27.93	-1.09
Total	41392	33.44	31.57	-1.87

Student performance over two school years

All of the analyses above were repeated for students who were in Grades 2-5 in Fall 2020-21 who also had comparable Winter and Fall scores for the previous two school years (2018-19 Winter, 2019-20 Fall, 2019-20 Winter, and 2020-21 Fall). 41,392 students had test scores for all four testing windows. The performance and change in performance for this strict cohort of students did not differ significantly from the results reported above when accounted for the fact that first-graders in 2020-21 Fall were not included in this cohort because they were not yet school-aged in 2018-19. These students experienced the largest Winter to next-Fall decline summarized in our analyses above; their inclusion and exclusion makes a difference in overall averages.

Star

Star Reading: Analysis of 2019-20 Winter to 2020-21 Fall Changes in National Percentile Rank

Students who were required to take Star Reading and Math assessments in both 2019-20 Winter and 2020-21 Fall are in seventh to eleventh grades in the current (2020-21) school year. For these students, the overall average NPR for Star Reading was the 25th percentile in both 2019-20 Winter and 2020-21 Fall windows (a 0.09 percentile decline).

The number of students in grades 7-11 in 2020-21, who were in grades 6-10 in 2019-20 and who took Star Reading in both 2019-20 Winter and 2020-21 Fall, is 25,263. Black/African American students constituted 47% of the cohort, in addition to 21% Hispanic/Latino students, 15% White students, and 11% Asian students. Of these students, 12% were English Learners and 67% were Economically Disadvantaged. Forty percent of the students in this cohort were placed in the Strategic Intervention tier based on their 2020-21 Fall Star Reading performance.

Winter to Next-Fall Changes by Grade Level

Examining the average NPR differences by grade level shows that grades 7 to 9 experienced 0.5- to 2-percentile decline, while grades 10 and 11 experienced a 2-percentile improvement in average NPR (Table 12). Average NPR did not vary too much between grade levels, ranging from 22 to 27 percentiles.

Table 12. Star Reading National Percentile Rank by Grade Level, 2019-20 Winter to 2020-21 Fall²⁰

Grade Level in 2020-21	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Fall NPR	Average NPR Differential
7	6093	26.84	24.94	-1.90
8	5533	26.17	24.81	-1.36
9	4354	25.78	25.22	-0.57
10	5067	21.53	23.29	1.76
11	4216	23.10	25.54	2.44
Total	25263	24.82	24.73	-0.09

Winter to Next-Fall Changes by Race/Ethnicity

The lack of difference in overall average NPR between 2019-20 Winter and 2020-21 Fall holds true for racial/ethnic groups (Table 13). As with aimswebPlus Reading results above, the subgroup differences in the average NPR itself gives us a more nuanced picture than the subgroup differences in Winter to next-Fall changes. Asian and White students scored around the 35th percentile, while Black/African American and Hispanic/Latino students scored around 15 percentiles lower, at 20-21st percentiles.

²⁰ Average NPR for the 38,217 students who took Star Reading in 2020-21 Fall was 26, 1 percentile higher. All students who took Star Reading test tenth and eleventh grades respectively scored 3 and 4 percentiles higher than the tenth- and eleventh-graders in the cohort.

Table 13. Star Reading National Percentile Rank by Race/Ethnicity, 2019-20 Winter to 2020-21 Fall

Race/Ethnicity	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Fall NPR	Average NPR Differential
American Indian/Alaskan Native	65	27.85	26.83	-1.02
Asian	2810	34.65	35.31	0.66
Black/African American	11969	20.76	20.77	0.01
Hispanic/Latino	5206	19.84	19.80	-0.04
Multiracial/Other	1330	28.36	27.86	-0.50
Native Hawaiian/Pacific Islander	32	28.34	26.75	-1.59
White	3851	35.70	34.84	-0.86
Total	25263	24.82	24.73	-0.09

Winter to Next-Fall Changes by English Learner Status

English Learners, in fact, improved in Star Reading average NPR by 2 percentiles from 2019-20 Winter to 2020-21 Fall (Table 14). However, the average NPR for English Learners was around 9th percentile compared to 27th percentile for non-English Learners. This is a persistent difference of more than 15 percentiles between the two groups.

Table 14. Star Reading National Percentile Rank by EL Status, 2019-20 Winter to 2020-21 Fall

English Learner Status	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Fall NPR	Average NPR Differential
Non-English Learner	22203	27.21	26.85	-0.36
English Learner	3060	7.49	9.32	1.83
Total	25263	24.82	24.73	-0.09

Winter to Next-Fall Changes by Fall Performance Tier:

Students who placed at the At/Above Benchmark tier based on their 2020-21 Fall Star Reading performance have improved from their NPR performance from 2019-20 Winter by 6 percentiles (Table 15). Students who placed at the On Watch tier also improved, but only marginally. Students who placed in either intervention tiers, on average, showed a 3-percentile decline from their average NPR in 2019-20 Winter.

Table 15. Star Reading National Percentile Rank by 2020-21 Fall Performance Tiers

2020-21 Fall Performance Tier	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Fall NPR	Average NPR Differential
At/Above Benchmark (40-100%)	6574	53.54	59.88	6.34
On Watch (25-39%)	3494	30.79	31.47	0.68
Strategic Intervention (10-24%)	5155	19.52	16.49	-3.03
Intensive Intervention (<10%)	10040	6.66	3.59	-3.07
Total	25263	24.82	24.73	-0.09

How to read this table: This table includes 25,263 students who took the Star Reading assessment when they were in grades 6-10 in 2019-20 Winter and in grades 7-11 in 2020-21 Fall. These students are grouped into four performance tiers based on their 2020-21 Fall performance. Unlike aimswebPlus, Star performance tiers are based on NPR: At/Above Benchmark students have scored between 40th to 100th percentiles. Students who have scored at the At/Above Benchmark tier in 2020-21 scored, on average, at 60th percentile whereas they had scored, on average, at the 54th percentile. The high performing students have shown a 6-percentile improvement on average.

Winter to Next-Fall Changes by Economic Disadvantage

Similar to patterns with respect to other demographic categories, the change in average NPR in Star Reading from 2019-20 Winter to 2020-21 Fall was minimal for both Economically Disadvantaged and non-Economically Disadvantaged students (Table 16). However, again similarly, the two groups differed in average NPR consistently across testing windows. Economically Disadvantaged students scored at 22nd percentile and non-Economically Disadvantaged students scored at 31st percentile, a 9-percentile difference.

Table 16. Star Reading National Percentile Rank by Economic Disadvantage, 2019-20 Winter to 2020-21 Fall

Economic Disadvantage Status	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Fall NPR	Average NPR Differential
Economically Disadvantaged	16851	21.86	21.64	-0.22
Not Economically Disadvantaged	8390	30.78	30.95	0.17
Total	25263	24.82	24.73	-0.09

Note: There were 22 students whose economic disadvantage status was not known. They are included in the counts; however, data on their performance are not reported.

Star Math: Analysis of 2019-20 Winter to 2020-21 Fall Changes in National Percentile Rank

For students who were in grades 7-11 in Fall 2020-21, the overall average NPR for Star Math was the 39th percentile in 2019-20 Winter and the 38th percentile in 2020-21 Fall (a 0.63 percentile decline).

The number of students who took Star Math in both 2019-20 Winter and 2020-21 Fall is 23,642. Students who took Star Math in both windows are similarly distributed across racial/ethnic groups: 47% of students were Black/African American, 21% were Hispanic/Latino, 15% were White, and 12% were Asian. Of the cohort, 13% were English Learners and 67% were Economically

Disadvantaged. In terms of overall 2020-21 Fall Star Math performance, 23% of the students in this cohort were placed in the Strategic Intervention tier.

Winter to Next-Fall Changes by Grade Level:

Although there was not a considerable change in average NPR from 2019-20 Winter to 2020-21 Fall overall, the change was not even for all grade levels. Students who were 9th graders in 2019-20 and 10th graders in 2020-21 improved by 6 percentiles (Table 17). On the other hand, 2020-21 seventh and eighth-graders each experienced a decline of around 5 percentiles. The levels of average NPR in both testing windows vary in a narrow 35 to 41st percentile range across the grade levels.

Table 17. Star Math National Percentile Rank by Grade Level, 2019-20 Winter to 2020-21 Fall²¹

Grade Level in 2020-21	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Fall NPR	Average NPR Differential
7	5418	40.78	35.43	-5.35
8	5448	41.09	36.33	-4.76
9	4162	39.55	40.38	0.83
10	4818	34.72	40.88	6.16
11	3796	37.82	39.63	1.81
Total	23642	38.93	38.29	-0.63

Winter to Next-Fall Changes by Race/Ethnicity:

Compared to other groups, White students had a larger decline in average NPR from 2019-20 Winter to 2020-21 Fall: 3 percentiles, from 51st to 48th (Table 18). In comparison, Hispanic/Latino students showed a small decline, from 33rd to 32nd percentile, and Black/African American students showed a small improvement, from 31st to 32nd percentile. Asian students had a higher average NPR, in the low 60th percentiles, and showed a 1-percentile decline between the two testing windows.

Table 18. Star Math National Percentile Rank by Race/Ethnicity, 2019-20 Winter to 2020-21 Fall

Race/Ethnicity	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Fall NPR	Average NPR Differential
American Indian/Alaskan Native	57	41.47	43.00	1.53
Asian	2781	63.69	62.53	-1.16
Black/African American	11089	31.08	31.58	0.50
Hispanic/Latino	4862	33.35	32.31	-1.04
Multiracial/Other	1240	40.14	38.23	-1.91
Native Hawaiian/Pacific Islander	29	55.79	56.21	0.41
White	3584	50.94	48.18	-2.76
Total	23642	38.93	38.29	-0.63

²¹ Average NPR for the 37,231 students who took Star Math in 2020-21 Fall was 39.6, 1 percentile higher. All students who took Star Reading test tenth and eleventh grades respectively scored 4 and 5 percentiles higher than the tenth- and eleventh-graders in the cohort.

Winter to Next-Fall Changes by English Learner Status

Changes in Average NPR for Star Math with respect to English Learner status from 2019-20 Winter to 2020-21 Fall were minor (Table 19). As with aimswebPlus Reading and Star Reading above, the more interesting finding in the average NPR comparison between the two testing windows is the persistence of performance differential between the two groups. English Learners scored on average at the 28th percentile, about 12 percentiles lower than non-English Learners, who scored at the 40th percentile on average.

Table 19. Star Math National Percentile Rank by English Learner Status, 2019-20 Winter to 2020-21 Fall

English Learner Status	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Fall NPR	Average NPR Differential
Non-English Learner	20647	40.48	39.72	-0.75
English Learner	2995	28.24	28.44	0.21
Total	23642	38.93	38.29	-0.63

Winter to Next-Fall Changes by Fall Performance Tier

Students who performed best in the 2020-21 Fall Star Math assessment improved on their performance in 2019-20 Winter Star Math assessment by 5 percentiles, on average (Table 20). Students who performed at other performance tiers showed a decline from 4 to 6 percentiles.

Table 20. Star Math National Percentile Rank by 2020-21 Fall Performance Tiers

2020-21 Fall Performance Tier	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Fall NPR	Average NPR Differential
At/Above Benchmark (40-100%)	10607	62.12	67.16	5.04
On Watch (25-39%)	3054	35.42	31.88	-3.54
Strategic Intervention (10-24%)	4484	22.75	16.44	-6.31
Intensive Intervention (<10%)	5497	9.31	3.98	-5.33
Total	23642	38.93	38.29	-0.63

How to read this table: This table includes 23,642 students who took the Star Math assessment when they were in grades 6-10 in 2019-20 Winter and in grades 7-11 in 2020-21 Fall. These students are grouped into four performance tiers based on their 2020-21 Fall performance. Star performance tiers are based on NPR: At/Above Benchmark students have scored between 40th to 100th percentiles. Students who scored at the At/Above Benchmark tier in 2019-20 Winter scored on average, at the 62nd percentile, and in 2020-21 Fall they scored, on average, at the 67th percentile, an improvement of 5 percentiles. Students who scored in the Intensive Intervention Tier in 2020-21 Fall decreased an average of 5 percentiles in their performance from 2019-20 Winter to 2020-21 Fall.

Winter to Next-Fall Changes by Economic Disadvantage:

Economically disadvantaged students in the analytic sample consistently performed at about the 35th percentile on average, compared to non-economically disadvantaged students who performed at about the 45th percentile on average (Table 21). Neither group of students showed a decline larger than 1 percentile in the average NPR between 2019-20 Winter to 2020-21 Fall.

Table 21. Star Math National Percentile Rank by Economic Disadvantage, 2018-19 Winter to 2019-20 Fall

Economic Disadvantage Status	Number of students	Average 2019-20 Winter NPR	Average 2020-21 Fall NPR	Average NPR Differential
Economically Disadvantaged	15937	35.68	34.87	-0.81
Not Economically Disadvantaged	7686	45.68	45.42	-0.26
Total	23642	38.93	38.29	-0.63

Note: There were 19 students whose economic disadvantage status was not known. They are included in the counts; however, data on their performance are not reported.

Star Reading and Math Test Duration Analysis

According to the Renaissance, the vendor that produces the Star assessment, the Star Reading test should take about 15 minutes and the Star Math test should take about 20 minutes to complete.²² Tables 22 and 23 present the average time it took students to complete Star Reading and Math tests, respectively. The tables compare the average test durations in 2019-20 Winter and 2020-21 Fall by grade level. These comparisons are not meant to be a measure of performance. They are intended as indicators of effect of in-person/at-school vs. at-home test administration.

For students in grades 7 to 11, the average reading test duration overall was 21 minutes in 2019-20 Winter and 23 minutes in 2020-21 Fall (Table 22). Note that these durations are 6-8 minutes longer than the average duration suggested by the vendor. The 2020-21 Fall average duration was longer by 1 to 2 minutes than 2019-20 Winter and the difference in average duration between the two windows increases with grade level.

Table 22. Star Reading Time to Completion (TTC) by Grade Level, 2019-20 Winter to 2020-21 Fall

Grade Level in 2020-21	Number of Students	Average 2019-20 Winter TTC (minutes)	Average 2020-21 Fall TTC (minutes)	Average TTC Differential (minutes)
7	6093	21.44	22.39	0.95
8	5533	21.76	22.92	1.16
9	4354	22.07	23.13	1.05
10	5067	20.86	23.04	2.18
11	4216	20.70	22.87	2.17
Total	25263	21.38	22.84	1.46

Note: The standard deviation for 2019-20 Winter was 7.3 and for 2020-21 Fall it was 8.8.

For the same grade span, the average math test duration overall was 25 minutes in 2019-20 Winter and 28 minutes in 2020-21 Fall (Table 23). The overall average math test duration was 5-8 minutes longer than the average duration suggested by the vendor. The 2020-21 Fall average duration was 2 to 4 minutes longer than average duration in 2019-20 Winter. Seventh graders had the lowest

²² See: [Parent’s Guide to STAR Assessments](#). According to [Star Assessments for Reading Technical Manual](#) (2019), in 2015-16 school year, median duration for Star Reading was 19 minutes and grade level standard deviation ranged from 4.6 to 7.2 (p.8). According to [Star Assessments for Math Technical Manual](#) (2020), in 2015-16 school year, median duration for Star Math was 25 minutes and standard deviation for different grade levels ranged from 5.8 to 8.25 (p. 7-8).

difference in average duration between the two windows (2.32 minutes) and eleventh-grade students had the highest difference (3.49 minutes).

Table 23. Star Math Time to Completion by Grade Level, 2019-20 Winter to 2020-21 Fall

Grade Level in 2020-21	Number of Students	Average 2019-20 Winter TTC (minutes)	Average 2020-21 Fall TTC (minutes)	Average TTC Differential (minutes)
7	5418	26.31	28.63	2.32
8	5448	26.21	28.63	2.42
9	4162	25.27	28.00	2.72
10	4818	24.74	27.02	2.28
11	3796	23.71	27.20	3.49
Total	23642	25.37	27.96	2.59

Note: The standard deviation for 2019-20 Winter was 10.8 and for 2020-21 Fall it was 13.

Test durations under at-home test administration were longer than in-person/at-school administration; however the difference, on average, was not markedly large.

Conclusions

As a result of the Covid-19 school closures in spring 2020, SDP did not administer the planned District-wide assessments for reading and math (aimswebPlus for grades K-5 and Star for grades 6-12) nor the state standardized assessments (PSSA). To examine whether student performance declined as a result of Covid-19 school closures and the resulting online learning approach, we examined student performance on the last available test in 2019-20 (Winter testing window) and compared it to their performance on the first available test in 2020-21 (Fall testing window). We then examined whether the change in Winter to next-Fall performance for students was different during the Covid-19 closure period compared to the previous Winter to next-Fall testing period, using National Percentile Rank (NPR) as a way to measure whether students were maintaining growth relative to a normed national sample.

We found that aimswebPlus Reading performance for students in grades 1-5 in 2020-21 (K-4 in the 2019-20 school year) declined an average of 6 percentiles from 2019-20 Winter to 2020-21 Fall. Some of this decline is not unexpected. Previous research has identified “summer learning loss” as a problem nationwide, and in Philadelphia, a comparable group of students showed a 2-percentile decline in the same reading tests in the prior year (2018-19 Winter to 2019-20 Fall). What is striking about the decline is how large it is for students who were 2019-20 kindergarteners and 2020-21 first-grade students. While comparable kindergarten to first-grade students showed a decline of 7 percentiles from Winter to next-Fall in the school-year before Covid-19 school closures (Table 3), during online learning, the decline from 2019-20 Winter to 2020-21 Fall was 15 percentiles for the 2020-21 first-graders (Table 2).

For grades 7 to 11 (6-10 in the 2019-20 school year), performance on Star Reading and Star Math, as measured by NPR, and test duration did not change significantly from 2019-20 Winter to 2020-21 Fall. Analysis of performance details across demographic groups highlight no difference in the way performance changed before and after the transition to online learning, but the persistence of differences in performance that stay consistent during this period of change.

For both aimswebPlus and Star results, the changes from 2019-20 Winter to 2020-21 Fall are not too different between different demographic groups; online learning did not impact different groups disproportionately. However, the pre-online learning performance differences between demographic groups continue to persist.

Appendix

More about AimswebPlus Tests at the District

In 2014-15 Spring, the School District of Philadelphia (SDP) adopted Pearson’s AIMSweb (now aimswebPlus²³), a universal early literacy screening, benchmarking, and progress-monitoring tool, as its major instrument to assess and progress-monitor literacy proficiency for all students in kindergarten through third grade. In 2019-20 Fall, SDP extended the requirement of administering aimswebPlus literacy assessments to include fourth- and fifth-grade students. In Fall 2020, aimswebPlus Reading and aimswebPlus Math assessment administration became required for all K-5 students and K-8 Special Education students.

The aimswebPlus Reading assessments students take vary by grade level and begin in kindergarten by assessing pre-literacy skills such as letter recognition, letter sounds, and blending. Later, these assessments progress to include identification of common words, oral and silent reading fluency, and reading comprehension. All of these skills contribute to the development of proficient readers and allow teachers to pinpoint skills that need additional work in order to improve overall reading achievement.

AimswebPlus assessments are administered during set testing windows in the fall, winter, and spring. For literacy, prior to 2020-21 Fall, each grade level had one “core” assessment that was used for District-level accountability and “required” assessments used for additional monitoring purposes. Participation was reported for core and required assessments separately. As of 2020-21 Fall, there are a set of required tests that vary by grade level. Some of these tests contribute to the composite score and as of 2020-21 Fall, participation and performance metrics are based on students receiving a composite score. Tests that were administered in 2020-21 Fall but are not included in the composite score are used for instructional and screening purposes.

More about Star Tests at the District

Star Assessments, provided by Renaissance, are a suite of computer-adaptive assessments that measure students’ reading and math skills, monitor achievement and growth, and track understanding of skills aligned to state standards and Common Core State Standards. Renaissance offers six different tests as a part of the Star Assessment suite, including Star Reading, Star Math, Star Early Literacy, Star CBM (curriculum-based measures), Star Custom, and Star in Spanish. Star Reading, Star Math, Star Early Literacy, and Star in Spanish are adaptive tests, meaning the questions students receive adjust depending upon their answers. Therefore, two students taking the same assessment can receive different question sets depending on their responses.

For students in grades 6-10, Star is an approved universal screener tool, and for students in grades 6-8, Star Reading is also used to inform instructional reading levels. Finally, Star Reading may be used for MTSS for ELA in grades 6-12. Starting in 2019-20, SDP launched Star Reading and Star

²³ SDP transitioned from AIMSweb to aimswebPlus for grades K-3 starting Fall 2018-19 and aimswebPlus Reading became required for K-5 a year later. As part of the aimswebPlus assessment package, math assessments were made available to K-5 teachers for the purposes of progress monitoring students but teachers were not required to administer the math assessments. From Fall 2020 on, aimswebPlus math assessments are required for K-5.

Math to inform District-wide instruction and progress monitoring. Initially, Star Reading and Star Math were required for students in grades 6-10 and were optional for students in grades 11-12. Starting in Fall 2020, Star Reading and Star Math assessments are required for all students in grades 6-12, including students receiving special education services.

Comparable Tests and Cohorts

Table A.1: AimswebPlus Reading Comparable Tests for 2018-19 Winter-2019-20 Fall and 2019-20 Winter-2020-21 Fall Performance Comparisons

2018-19 Winter Grade level	Comparison test	2019-20 Fall Grade level	2019-20 Winter (base cohort)	Comparison test	2020-21 Fall Grade level
		K	K	NWF	1
K	NWF	1	1	ORF	2
1	ORF	2	2	ORF	3
2	ORF	3	3	ORF	4
3	ORF	4	4	ORF	5
4	ORF	5	5	ORF	

Availability of comparable aimswebPlus Reading test results from 2018-19 school year allows us to do the additional retrospective comparison for the students who were tested during the 2019-20 Winter window. For this retrospective analysis in the aimswebPlus reading section above, the list of students who took comparable tests in Winter 2019-20 is again the base cohort. The only difference in comparable tests in the retrospective analysis is, only base cohort first-grade students who have taken NWF are included in order to be able to examine their progress over their NWF results from their kindergarten year, 2018-29. As a result, the cohort used in this analysis includes first-grade students who have taken ORF in 2019-20 Winter, NWF in 2019-20 Fall, and NWF in 2018-29 Winter, when they were in kindergarten. Similarly, it includes second- to fifth-grade students in 2019-20 who have taken ORF in all three windows -2019-20 Winter, 2019-20 Fall, and 2018-29 Winter, while they were in grades 1-4 (Table A.1.).