



Assessing Student Performance Through a Year of Virtual Learning: A Cohort Comparison of Student Performance on 2019-20 Winter and 2020-21 Spring Star Assessments and End-of-Year Review

This report, along with the accompanying report on aimswebPlus performance, concludes the analyses [in our January 2021 report](#) that focused on student reading and math test performance from 2019-2020 Winter (prior to closing schools due to the Covid-19 pandemic) to 2020-21 Fall (during virtual learning for all students) and [in our April 2021 addendum](#) that compared the performance of the same set of 2019-20 Winter test takers to their 2020-21 Winter reading and math performance. We compare Star reading and math performance of the 2019-20 Winter test takers to their performance in 2020-21 Spring and we also present our summary of this cohort's assessment performance progress throughout the 2020-2021.

The main metric used in this report, as in the previous analyses, is National Percentile Rank (NPR). NPR is a norm-referenced measure that compares student performance to a national sample of students.

Key Findings

- On average, student performance during the last testing window during the virtual/hybrid period (2020-21 Spring) was similar to student performance during the last in-person/at-school administration of Star Reading and Star Math in 2019-20 Winter.
- As the previous analyses in this cohort study series showed, performance of different demographic groups did not show large fluctuations throughout the year; however, performance differences between these groups persisted.

Ebru Erdem, Ph.D.,
Director

Joy Lesnick, Ph.D.,
Deputy Chief

Office of Evaluation,
Research,
and Accountability

September 2021

Contents

District-Wide Testing and Research Analyses During the 2019-20 and 2020-21 School Years	2
Keeping Track of School Years, Testing Windows, and Research Reports	3
Research Questions for the Final Cohort Analyses	4
The National Percentile Rank (NPR) Metric	4
Comparable Tests and the Cohort Sample	5
Findings	6
Star Reading	6
Star Reading: Analysis of 2019-20 Winter to 2020-21 Spring Changes in National Percentile Rank	6
Star Math: Analysis of 2019-20 Winter to 2020-21 Spring Changes in National Percentile Rank	9
Trends in National Percentile Rank throughout 2020-21	12
Star Reading: End-of-Year Trends	12
Star Math: End-of-Year Trends	14
Conclusions	16
Appendix A. Demographic Details for the Cohort	17
Appendix B: National Percentile Rank Analyses Compared to Normal Curve Equivalent-Based National Percentile Rank	19
Overall Cohort Performance in Star Reading and Star Math Throughout 2020-21 with NCE-based NPR Comparisons	21
Winter-to-Spring Star Reading Tables with NCE-based NPR Comparison	22
Winter-to-Spring Star Math Tables with NCE-based NPR Comparison	25

District-Wide Testing and Research Analyses During the 2019-20 and 2020-21 School Years

Throughout the 2020-21 school year, we have been examining student performance on interim assessments to better understand how student performance has been affected by the Covid-19 pandemic and the District's shift to 100% virtual learning in March 2020. Beginning in 2020-21, the aimswebPlus assessment was required for grades K-5 and the Star assessment was required for grades 6-12, and these tests were administered online throughout the 2020-21 school year. The dates of the 2019-20 and 2020-21 administration windows were as follows:

- 2019-20 Fall (in-person)
 - aimswebPlus (grades K-5*): September 3-October 31, 2019;
 - Star (grades 6-12): Term 1, September 19-October 4, 2019.
- 2019-20 Winter (in-person)
 - aimswebPlus (grades K-5*): January 2-31, 2020;
 - Star (grades 6-8): Term 3, March 2-12, 2020;
 - Star (grades 9-12): Term 2, February 5-21, 2020.
- 2019-20 Spring (not administered due to Covid-19)
- 2020-21 Fall (online)
 - September 14-October 23, 2020 for both tests, all grades.
- 2020-21 Winter (online)
 - December 14, 2020-January 29, 2021 for both tests, all grades.¹
- 2020-21 Spring (online/in-person)
 - aimswebPlus (grades K-5*): April 14-June 11, 2021;
 - Star (grades 6-12): May 3-June 11, 2021.

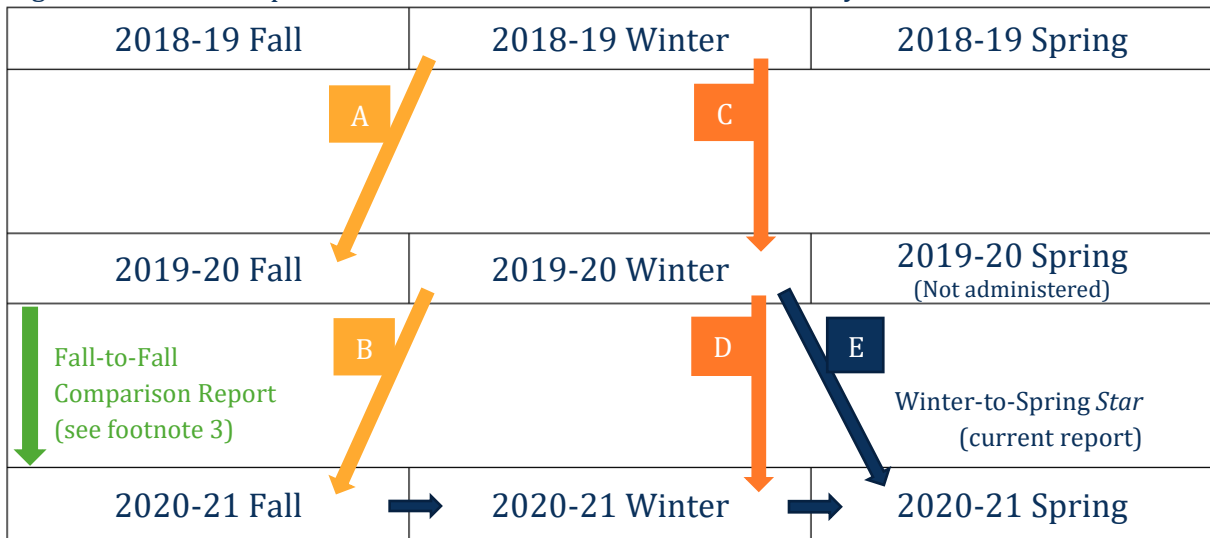
*AimswebPlus was required for special education students in grades 6-8 and is administered during the aimswebPlus windows for Grades K-5.

¹ Results from schools that started aimswebPlus 2020-21 Winter testing on December 1, 2021, were included in the April 2021 addendum.

Keeping Track of School Years, Testing Windows, and Research Reports

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 green vertical arrow represents a Fall-to-Fall comparison of average student performance published in a November 2020 report.² This was our first examination of how students performed on Fall interim assessments as compared to students who were in that same grade the previous year.

We then looked more closely at the same students' performance over time. The two diagonal arrows labeled A and B illustrate the analyses presented in the report dated January 2021.³ In that report, we analyzed changes in performance between 2019-20 Winter, the last in-person/at-school assessments administered, and 2020-21 Fall, the first online/at-home assessments administered (arrow B). The January 2021 report also compared the 2019-20 to 2020-21 change in performance to the previous year (2018-19 to 2019-20) for aimswebPlus Reading (arrow A).

An addendum published in April 2021 repeated the same analyses, comparing the 2019-20 Winter results to the performance of *the same cohort of students* in the 2020-21 Winter assessment window (arrow D). The April 2021 addendum also compared the change in performance to the Winter-to-Winter changes in the previous year for aimswebPlus Reading (Arrow C).⁴

² See: [K-12 Student Participation and Performance on aimswebPlus and Star Assessments, Fall 2019 and Fall 2020](#), November 2020.

³ See [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](#), January 2021.

⁴ See [Assessing Student Performance Before and During Virtual Learning: A cohort Comparison of Student Performance on 2019-20 Winter and 2020-21 Winter aimswebPlus and Star Assessments](#), April 2021.

This report, which is one of the two final reports in this series, examines student performance *on Star assessments* from 2019-20 Winter to 2020-21 Spring (Arrow E) as well as analyzing the trends over the course of the three assessment windows during the 2020-21 school year. An accompanying report does the same for *aimswebPlus*.

Research Questions for the Final Cohort Analyses

The analyses presented in this report complement our previous reports by comparing students' performance on their last in-person/at-school Star assessments (2019-20 Winter) to their latest virtual/hybrid learning period assessment (2020-21 Spring) at the time of writing.⁵

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 Spring Star Reading and Math (2020-21 grade 7-11 students)?
 - 1.a. Were there identifiable patterns in Star Reading and Math performance differentials between 2019-20 Winter and 2020-21 Spring with respect to grade level, race/ethnicity, English learner status, 2020-21 Spring performance tier on the test, and Economic Disadvantage status?
2. What were the trends in Star Reading and Math performance, as measured by average NPR, throughout the 2020-21 school year?

As in the previous analyses in this series, the primary metric we examine in this addendum is the average National Percentile Rank (NPR) for Star.

The National Percentile Rank (NPR) Metric

NPR is a norm-referenced score that measures a student's performance in reference to the performance of a national sample of students (the norm sample) at the same grade level. It shows what percentage of students in the norm sample scored lower than the particular student. The vendor for Star Reading and Math has established norm samples which are representative of the national demographic distributions of K-12 students. Star Reading and Math conducted a norming process using samples of test takers between August 15, 2014, and June 30, 2015. It is important to note that the norm samples used to identify NPR were established prior to school closures as part of Covid-19 containment efforts. Therefore, NPR norms are not affected by any Covid/virtual learning related performance changes.⁶

⁵ Starting on March 8, 2021, SDP started hybrid-learning (students attending in person two days a week) in Phases. For more about hybrid-learning at SDP see: <https://www.philasd.org/coronavirus/schoolstart2020/#highlights>.

⁶ More details about the norm samples can be found in the technical reports for these tests: [Star Reading](#), p. 89; [Star Math](#), p. 74.

Comparable Tests and the Cohort Sample

Our January 2021 report includes detailed descriptions of the assessments administered at the School District of Philadelphia (SDP) in 2019-20, before the pivot to virtual instruction in response to COVID-19, as well as in 2020-21 Fall, during virtual instruction. In this report, as in the previous cohort analyses in this series, 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 we exclude the grades for which the assessment was not required prior to 2020-21 (Table 1).

Table 1. Comparable Tests

Test	Possible Comparisons
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

Students who completed the Star Reading and Math tests in both 2019-20 Winter and 2020-21 Spring testing windows are the cohort of students used for this analysis.⁷ The 2019-20 Winter testing window was the last window in which assessments were administered before District school buildings were closed on March 13, 2020, due to Covid-19 containment efforts in Philadelphia. Among the 2019-20 Winter test-takers, those who took a Star tests in the 2020-21 Spring testing window were identified. For each of the tests analyzed in this report, the number of identified students (that is, students with results during both windows pairs) varies.⁸

Note that the students whose performance is analyzed in this report are not necessarily the same set of students whose performance was analyzed in the January 2021 report or the April 2021 addendum. The January 2021 report analyzed students who were tested in both 2019-20 Winter and 2020-21 Fall, and the April 2021 addendum analyzed students who were tested in both 2019-20 Winter and 2020-21 Winter. While there is a big overlap between the 2019-20 Winter test takers who were then tested in either window of 2020-21, there are students who participated in only one or two of the three 2020-21 assessment windows.

The 2020-21 Spring testing window coincided with the administration of Pennsylvania System of School Assessment (PSSA) for grades 6-8 in this cohort. The District prioritized the administration of the state tests and this might have impacted the number of students in the study cohort who took 2019-20 Winter who also took 2020-21 Spring Star tests, which is lower than the other two windows in 2020-21. See Table A.1 in the appendix for a comparison of the cohort in this analysis to all 2020-21 Spring test takers. A demographic breakdown of the 2020-21 Spring cohort whose Star Reading and Math performance is analyzed in this report is presented in Table A.2 in the appendix.

⁷ The 2020-21 Spring analysis studies the outcomes for students who took comparable tests in both 2019-20 Winter and 2020-21 Spring. The January 2021 report studied the outcomes for students who took the same tests in both 2019-20 Winter and 2020-21 Fall and the April 2021 addendum studied the outcomes for students who took the same tests in both 2019-20 Winter and 2020-21 Winter. We did not restrict the analysis to students who took the tests in all four testing windows; thus, although the sets of students analyzed for each 2020-21 testing window highly overlap, they are not exactly the same.

⁸ These figures are reported in the Totals row of the Number of Students Columns in each table in the Star Reading and Star Math sections.

Findings

The results from the analyses of student performance on Star Reading and Star Math are presented separately below. First, the comparison of 2020-21 Winter and 2019-20 Spring Star Reading and Star Math performance is presented, including the performance of student subgroups. In the next section, we summarize the progression of 2019-20 Winter test takers' Star Reading and Math performance through the three testing windows of 2020-21 school year.

Star Reading

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

Students who were required to take Star Reading and Math assessments in both 2019-20 Winter and 2020-21 Spring were in grades 7-11 in the 2020-21 school year. For these students, the overall average NPR for Star Reading was the 25th percentile in all windows analyzed: 2019-20 Winter, 2020-21 Fall, 2020-21 Winter, and 2020-21 Spring.

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 Spring, was 19,894. Black/African American students constituted 46% of the cohort, in addition to 20% Hispanic/Latino students, 16% white students, and 13% Asian students. Of these students, 13% were English Learners and 69% were economically disadvantaged. Demographic distributions are similar to the profiles of test takers in the previous 2020-21 testing windows. Forty percent of the students in this cohort were placed in the Intensive Intervention tier based on their 2020-21 Spring Star Reading performance (compared to 41% in the previous windows).

Winter-to-Spring Changes by Grade

Examining the average NPR differences by grade shows that grades 7 to 9 experienced a 0.8- to 1.5-percentile decline, while grades 10 and 11 experienced a 2.4 and 1.3-percentile respective improvement in average NPR (Table 2). Average NPR did not vary much between grades, ranging from 24 to 27th percentiles.

Table 2. Star Reading National Percentile Rank by Grade, 2019-20 Winter to 2020-21 Spring⁹

Grade in 2020-21	Number of Students	Average 2019-20 Winter NPR*	Average 2020-21 Spring NPR	Average NPR Differential
7	5147	27.30	26.52	-0.78
8	4633	26.80	25.57	-1.24
9	3284	27.41	25.89	-1.52
10	3744	21.33	23.71	2.38
11	3086	23.49	24.79	1.30
All Students	19894	25.49	25.40	-0.09

* The students included in this analysis are students who have been tested in both 2019-20 Winter and 2020-21 Spring. This is only a subset of the 30,128 grade 7-11 students who took Star Reading in 2020-21 Spring (see Table A.2 in Appendix A).

** The students whose results are shown here are those who have been tested both in 2019-20 Winter and 2020-21 Spring. This is not necessarily the same set of students who have been tested both in 2019-20 and 2020-21 Fall or Winter. For this reason, results for 2019-20 Winter may differ from the January 2021 report and the April 2021 addendum.

Winter-to-Spring Changes by Race/Ethnicity

Among racial/ethnic groups with large enough sample sizes, the only notable differences between 2019-20 Winter and 2020-21 Spring Star Reading average NPR for 2020-21 grade 7-11 students who took the test at both time periods were for Asian students, who performed 3 percentiles higher, and white students, who performed 2 percentiles lower (Table 3). The subgroup differences in the average NPR itself continue to give us a more nuanced picture than the subgroup differences in Winter-to-Spring changes. **Asian and white students scored around the 35th percentile, while Black/African American and Hispanic/Latino students scored about 15 percentiles lower, around the 20th percentile.**

Table 3. Star Reading National Percentile Rank by Race/Ethnicity, 2019-20 Winter to 2020-21 Spring

Race/Ethnicity	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Spring NPR	Average NPR Differential
American Indian/Alaskan Native	52	29.73	27.27	-2.46
Asian	2488	35.22	37.72	2.50
Black/African American	9124	20.78	20.66	-0.12
Hispanic/Latino	3995	20.32	20.33	0.00
Multiracial/Other	1035	28.35	27.68	-0.67
Native Hawaiian/Pacific Islander	26	33.92	37.58	3.65
White	3174	36.82	34.86	-1.96
All Students	19894	25.49	25.40	-0.09

⁹ Average NPR for the 30,128 students who took Star Reading in 2020-21 Spring was 26. Comparing the 2020-21 Spring average NPR for all test takers at grades 7-11 vs. study cohort, 7th and 8th grade cohort students performed about 1 percentile better than all test takers, 9th grade outcomes were the same, and 10th and 11th grade cohort students performed 3-4 percentiles worse than the whole sample of test takers.

Winter-to-Spring Changes by English Learner Status

English Learners improved their Star Reading average NPR by 3 percentiles from 2019-20 Winter to 2020-21 Spring (Table 4). However, the average NPR for English Learners was around the 11th percentile, compared to the 28th percentile for non-English Learners. This is a persistent difference of more than 15 percentiles between the two groups.

Table 4. Star Reading National Percentile Rank by EL Status, 2019-20 Winter to 2020-21 Spring

English Learner Status	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Spring NPR	Average NPR Differential
Non-English Learner	17354	28.09	27.58	-0.51
English Learner	2540	7.71	10.45	2.74
All Students	19894	25.49	25.40	-0.09

Note: English Learner status as of Spring 2020-21.

Winter-to-Spring Changes by Spring Performance Tiers

Students who placed at the At/Above Benchmark tier based on their 2020-21 Spring Star Reading performance improved from their 2019-20 Winter NPR performance by 9 percentiles (Table 5). Students who placed at the On Watch tier also improved, but only by 1 percentile. Students who placed in either intervention tier, on average, showed a 4-5-percentile decline from their average NPR in 2019-20 Spring.

Table 5. Star Reading National Percentile Rank by 2020-21 Spring Performance Tier

2020-21 Spring Performance Tier	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Spring NPR	Average NPR Differential
At/Above Benchmark (40-100%)	5414	52.67	61.28	8.61
On Watch (25-39%)	2698	30.81	31.69	0.88
Strategic Intervention (10-24%)	3802	20.01	16.36	-3.66
Intensive Intervention (<10%)	7980	7.86	3.23	-4.63
All Students	19894	25.49	25.40	-0.09

How to read this table: This table includes 19,894 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 Spring. These students are grouped into four performance tiers based on their 2020-21 Spring performance. Star performance tiers are based on NPR: At/Above Benchmark students have scored between the 40th and 100th percentiles. Students who scored at the At/Above Benchmark tier in 2020-21 Spring scored, on average, at the 61st percentile, whereas they had scored, on average, at the 53rd percentile the previous winter. Thus, high-performing students showed a 9-percentile improvement, on average.

Winter-to-Spring Changes by Economic Disadvantage

Similar to patterns observed in other demographic categories, the average change in average NPR in Star Reading from 2019-20 Winter to 2020-21 Spring was minimal for both economically disadvantaged and non-economically disadvantaged students (Table 6). Also similar to previous findings, the two groups differed in average NPR consistently across testing windows. Economically Disadvantaged students scored at the 23rd percentile, and non-economically disadvantaged students scored at the 31st percentile, an 8-percentile difference.

Table 6. Star Reading National Percentile Rank by Economic Disadvantage, 2019-20 Winter to 2020-21 Spring

Economic Disadvantage Status	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Spring NPR	Average NPR Differential
Economically Disadvantaged	13736	22.64	22.75	0.11
Not Economically Disadvantaged	6158	31.84	31.30	-0.54
All Students	19894	25.49	25.40	-0.09

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

For students who were in grades 7-11 in 2020-21 Spring and took Star Math in both 2020-21 Spring and 2019-20 Winter testing windows, the overall average NPR for Star Math was the 40th percentile in 2019-20 Winter and the 39th percentile in 2020-21 Spring (a 1-percentile decline).

The number of students who took Star Math in both 2019-20 Winter and 2020-21 Spring was 18,492. Students who took Star Math in both windows were distributed across racial/ethnic groups similarly to the previous cohort analyses: 46% of students were Black/African American, 20% were Hispanic/Latino, 16% were white, and 13% were Asian. Of the cohort, 13% were English Learners and 70% were economically disadvantaged. In terms of overall 2020-21 Spring Star Math performance, 27% of the students in this cohort were placed in the Intensive Intervention tier.

Winter-to-Spring Changes by Grade

Change in average NPR from 2019-20 Winter to 2020-21 Spring indicates only a 1 percentile decline overall; however, the change was not even across grades. Students who were ninth-graders in 2019-20 and tenth-graders in 2020-21 improved by 3 percentiles (Table 7). According to our previous analysis, 2020-21 seventh- and eighth-graders each experienced a Winter-to-next-Fall decline of around 5 percentiles, and these students' average NPR improved 2-4 percentiles Winter to Winter. Performance of seventh- and eighth-graders in the 2020-21 Spring sample declined 2-4 percentiles Winter-to-Spring. The average NPR in both 2019-20 Winter and 2002-21 Spring testing windows varied in a narrow 36 to 42nd percentile range across grades.

Table 7. Star Math National Percentile Rank by Grade, 2019-20 Winter to 2020-21 Spring¹⁰

Grade in 2020-21	Number of Students*	Average 2019-20 Winter NPR**	Average 2020-21 Spring NPR	Average NPR Differential
7	4629	41.97	38.33	-3.64
8	4508	41.99	39.93	-2.06
9	3157	40.77	39.14	-1.63
10	3481	36.23	39.66	3.43
11	2717	39.68	40.02	0.33
All Students	18492	40.35	39.36	-0.99

* The students included in this analysis are those who have been tested in both 2019-20 Winter and 2020-21 Spring. This is only a subset of the 29,099 grade 7-11 students who took Star Math in 2020-21 Spring (See Table A.2 in Appendix A).

** The students whose results are shown here are those who have been tested in both 2019-20 Winter and 2020-21 Spring. This is not necessarily the same set of students who have been tested in both 2019-20 and 2020-21 Fall or 2020-21 Winter. For this reason, 2019-20 Winter results may differ from the January 2021 report or the April 2021 addendum.

Winter-to-Spring Changes by Race/Ethnicity

Except for the Asian students, students of all racial/ethnic groups experienced a decline in average NPR from 2019-20 Winter to 2020-21 Spring. The decline was highest for the small group of Native Hawaiian/Pacific Islander students and white students (4 percentiles) and it was in the 0.5-2-percentile range for other groups (Table 8).

Table 8. Star Math National Percentile Rank by Race/Ethnicity, 2019-20 Winter to 2020-21 Spring

Race/Ethnicity	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Spring NPR	Average NPR Differential
American Indian/Alaskan Native	46	44.15	42.85	-1.30
Asian	2406	64.91	65.67	0.76
Black/African American	8483	31.69	31.22	-0.47
Hispanic/Latino	3685	34.24	33.44	-0.80
Multiracial/Other	951	40.49	38.64	-1.85
Native Hawaiian/Pacific Islander	24	57.92	53.88	-4.04
White	2897	52.83	48.91	-3.92
All Students	18492	40.35	39.36	-0.99

Winter-to-Spring Changes by English Learner Status

Winter-to-Spring changes in average NPR for Star Math show a 3-percentile improvement for English Learners and a 2-percentile decline for non-English Learners (Table 9). As with our cohort analysis reports on the previous testing windows, the more interesting finding in the average NPR comparisons is the persistence of performance differentials between the two groups. Despite showing more improvement, English Learners scored, on average, about 10 percentiles lower than non-English Learners (31st and 41st percentiles, respectively).

¹⁰ Average NPR for the 29,099 students who took Star Math in 2020-21 Spring was 39, same as the study cohort students. All tenth and eleventh grade Star Math test takers respectively scored 3 and 4 percentiles higher than the tenth and eleventh graders in the cohort. For other grades, cohort students performed up to 3 percentiles higher than the whole sample of test takers.

Table 9. Star Math National Percentile Rank by English Learner Status, 2019-20 Winter to 2020-21 Spring

English Learner Status	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Spring NPR	Average NPR Differential
Non-English Learner	16023	42.14	40.63	-1.51
English Learner	2469	28.73	31.09	2.37
All Students	18492	40.35	39.36	-0.99

Winter-to-Spring Changes by Spring Performance Tiers

Students who performed best in the 2020-21 Spring Star Math assessment improved their performance on the 2019-20 Winter Star Math assessment by 7 percentiles, on average (Table 10). Students who performed at the two intervention tiers showed a decline from 8 to 9 percentiles.

Table 10. Star Math National Percentile Rank by 2020-21 Spring Performance Tiers

2020-21 Winter Performance Tier	Number of Students	Average 2019-20 Winter NPR	Average 2020-21 Spring NPR	Average NPR Differential
At/Above Benchmark (40-100%)	8629	62.67	69.25	6.58
On Watch (25-39%)	2179	36.12	32.10	-4.02
Strategic Intervention (10-24%)	2629	24.57	16.14	-8.44
Intensive Intervention (<10%)	5055	12.28	3.54	-8.75
All Students	18492	40.35	39.36	-0.99

How to read this table: This table includes 18,492 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 Spring. These students are grouped into four performance tiers based on their 2020-21 Winter 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 2020-21 Spring had scored, on average, at the 63rd percentile in 2019-20 Winter. In 2020-21 Winter, these students scored, on average, at the 69th percentile, an improvement of 9 percentiles. Students who scored in the Intensive Intervention Tier in 2020-21 Spring decreased an average of 9 percentiles in their performance from 2019-20 Winter to 2020-21 Spring.

Winter-to-Spring Changes by Economic Disadvantage

Economically disadvantaged students in the analytic sample consistently performed at the 34-36th percentile range, on average, compared to non-economically disadvantaged students who performed at the 44-47th percentile range, on average (Table 11). Both groups of students demonstrated a Winter-to-Spring decrease of about 1 percentile in average NPR.

Table 11. Star Math National Percentile Rank by Economic Disadvantage, 2018-19 Winter to 2019-20 Spring

Economic Disadvantage Status	Number of students	Average 2019-20 Winter NPR	Average 2020-21 Spring NPR	Average NPR Differential
Economically Disadvantaged	12877	37.35	36.38	-0.97
Not Economically Disadvantaged	5615	47.24	46.18	-1.06
All Students	18492	40.35	39.36	-0.99

Trends in National Percentile Rank throughout 2020-21

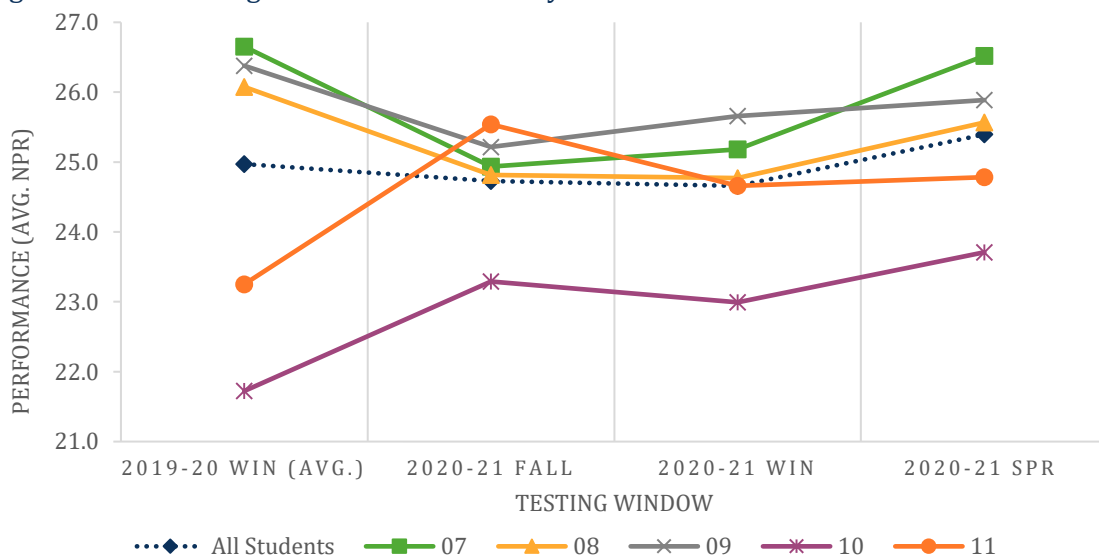
In this section we compile the average NPR for different student groups represented in the cohort across all testing windows studied: 2019-20 Winter, 2020-21 Fall, 2020-21 Winter, and 2020-21 Spring. As Table A.1 in Appendix A shows, different numbers of students among 2019-20 Winter test takers took the Star assessments in the fall, winter, and spring testing windows of the 2020-21 school year.

The demographic make-up of these three subsets of 2019-20 Winter base cohort was consistent across the 2020-21 testing windows; however, these were not exactly the same set of students. The trends shown in this section are the results of these overlapping but non-identical subsets of 2019-20 Winter test takers. Each subset had a different 2019-20 Winter average NPR, although they were very close (See Table A.1); we use the average of these three values as the 2019-20 Winter average NPR below.

Star Reading: End-of-Year Trends

Charting the cohort group performance by grade shows that students who were in grades 7-9 in 2020-21 show a different pattern than those who were in grades 10 and 11 (Figure 1). For grades 7-9, the 2020-21 Fall performance shows a decline from 2019-20 Winter and they recover from this decline -albeit at different levels- in 2020-21 Winter and Spring. To the contrary, students who were in grades 10 and 11 in 2020-21 school year, show a significant improvement in 2020-21 Fall compared to 2019-20 Winter. These students show some decline in the 2020-21 Winter testing window but end the year at a higher performance level than 2019-20 Winter. The performance for all grades varied within a narrow band between the 21st and 27th percentiles.

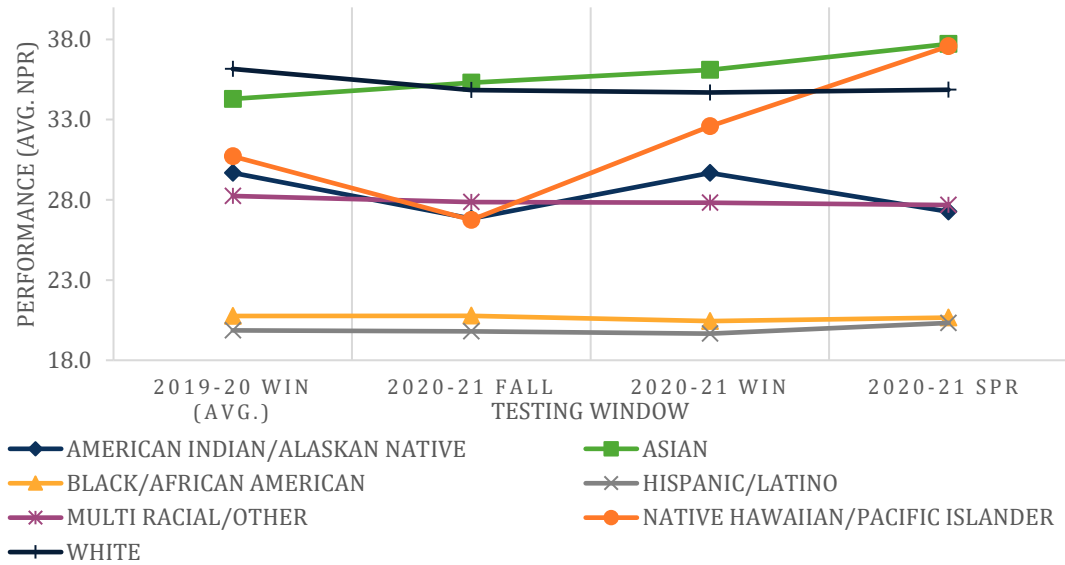
Figure 1. Star Reading Performance Trends by 2020-21 Grade



Trends for racial/ethnic groups display a different pattern, where the groups are clustered around three different levels of average NPR. Although their trends are divergent, Asian and white students performed above the 30th percentile consistently. Students who are multiracial, American

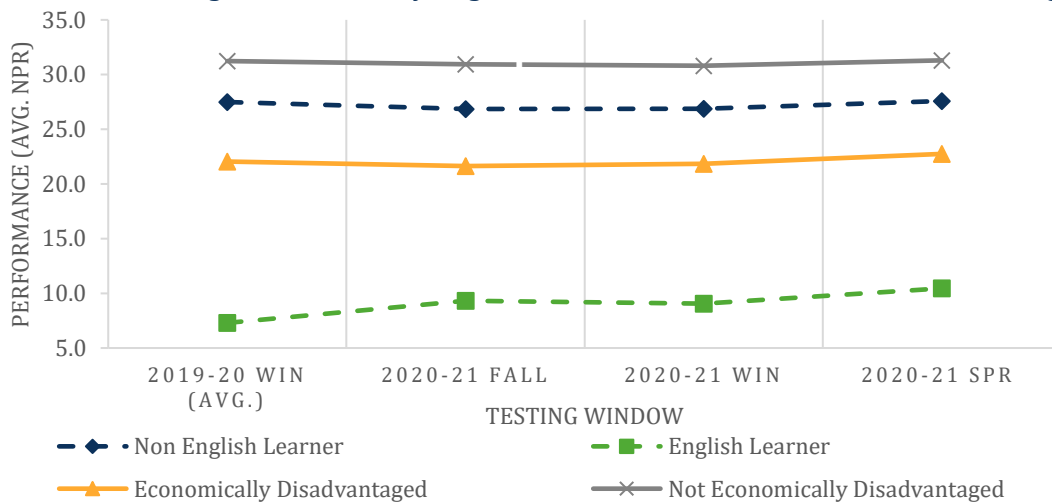
Indian/Alaskan Native, and Native Hawaiian/Pacific Islander performed in the upper 20th percentiles. Because the latter two of these groups had very few students in the sample, the performance level fluctuations are larger. Students who are Black/African American and Hispanic/Latino scored below the 20th percentile, and this remained stable throughout the year.

Figure 2. Star Reading Performance Trends by Race/Ethnicity



The divergent performance trends are also visible for students with different English Learner status (Figure 3). Although they made gains throughout the year, English Learners performed significantly lower than non-English learners. Similarly, students who are economically disadvantaged scored about 10 percentiles lower than their peers who are not economically disadvantaged. Both groups’ performance remained relatively stable through the testing windows.

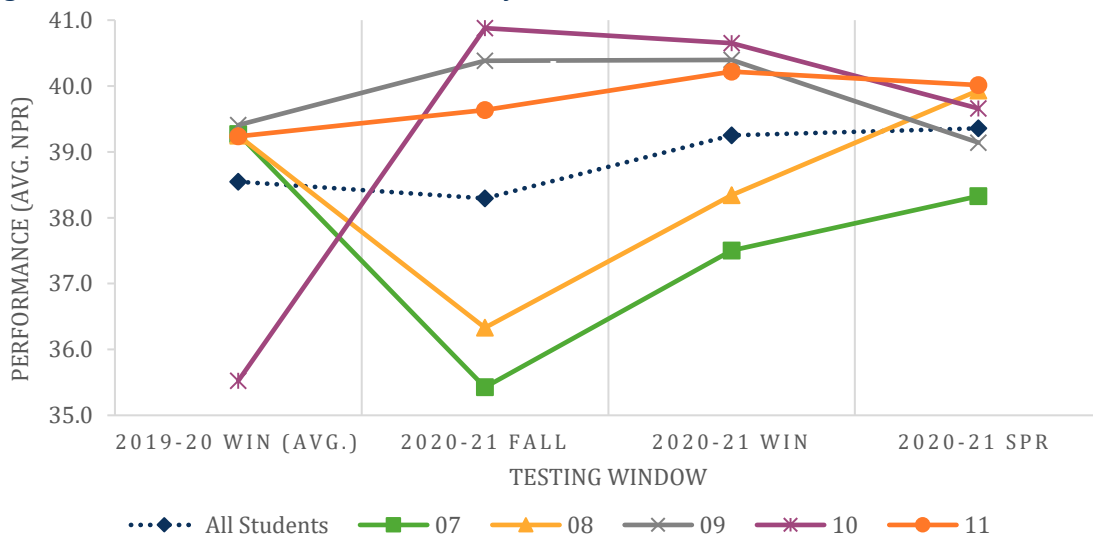
Figure 3. Star Reading Performance by English Learner Status and Economic Disadvantage Status



Star Math: End-of-Year Trends

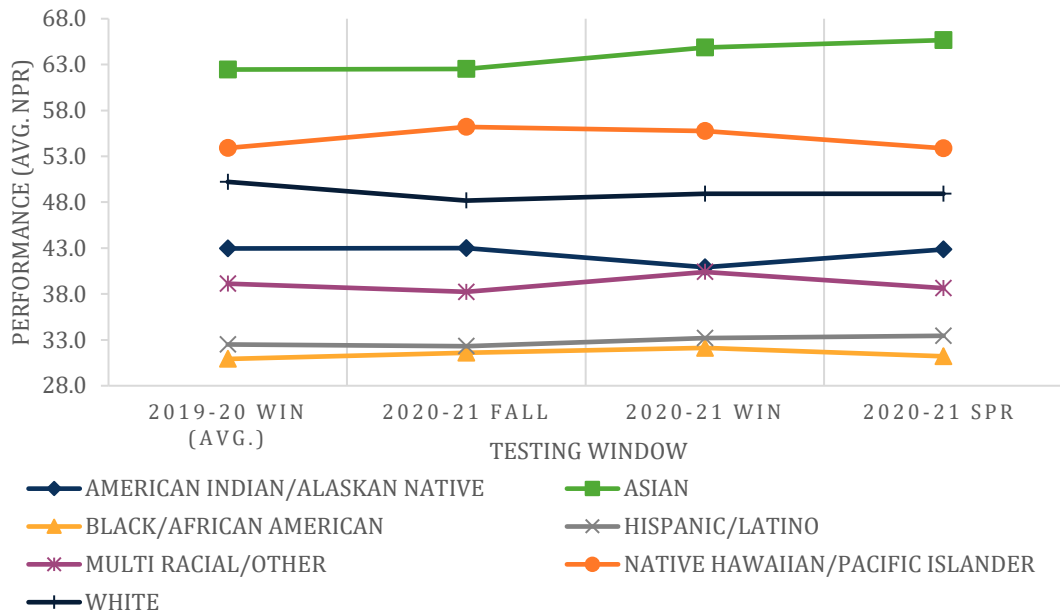
Star Math grade performance trends for the cohort analyzed show differing patterns for different grades (Figure 4). As with Star Reading, grades 7 and 8 show a pattern of Winter-to-next-Fall decline in Math, then recovery from this decline throughout the year. Ninth-graders showed Winter-to-next-Fall performance improvements but a decline in 2020-21 Spring. Tenth-graders experienced an even more striking improvement Winter-to-next-Fall; however, they showed some decline through the end of the year. These changes in average NPR by grade are confined to a narrow band between the 35th and 41st percentiles.

Figure 4. Star Math Performance Trends by Grade



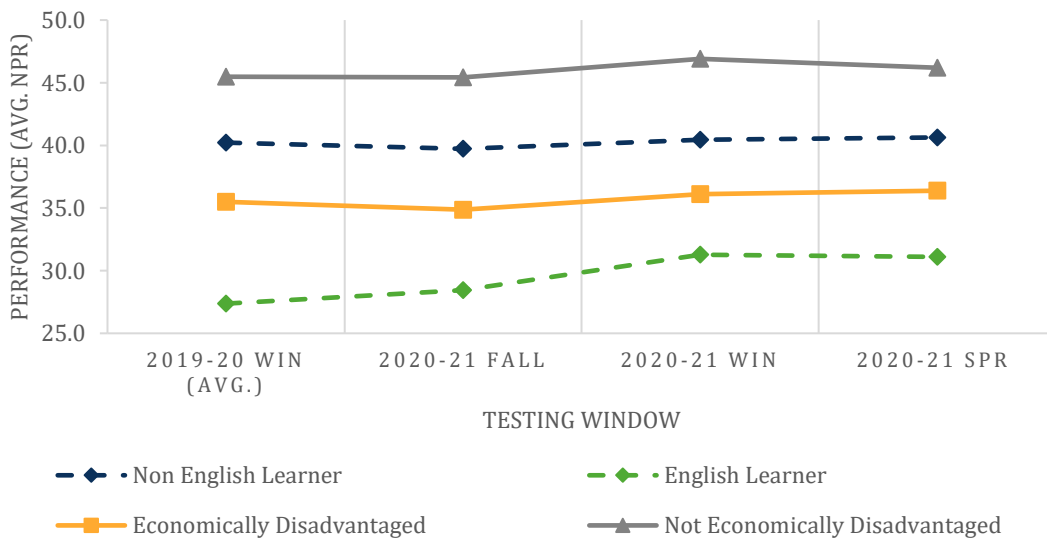
Star Math average performance levels show more variance when analyzed by race/ethnicity (Figure 5). While trends across testing windows are more or less stable with small fluctuations for each group, there are differences between racial/ethnic groups’ performance levels throughout. For example, Asian students, on average, perform higher than all other groups, above 60th percentile and show an upward trend during the year. Students who are Black/African American or Hispanic/Latino, as in Star Reading, showed lower levels of performance, below the 35th percentile.

Figure 5. Star Math Performance Trends by Race/Ethnicity



Trends in Star Math performance by English learner status and economic disadvantage status show patterns similar to Star Reading (Figure 6). English learners scored lower but showed improvements in performance throughout the year. Students who were not English learners, regardless of economically disadvantaged status, held more or less steady.

Figure 6. Star Math Performance by English Learner Status and Economic Disadvantage Status



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) or 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 (administered during the Winter testing window) and compared it to student performance on tests administered throughout the 2020-21 school year. In this last report in this series, which is accompanied by a report on aimswEBPlus performance for grades 1-5, we analyzed 2020-21 Spring assessment results from the same set of students who took comparable tests in the 2019-20 Winter testing window as well as performance trends throughout the 2020-21 school year.

For students in grades 7 to 11 (who were in grades 6-10 in 2019-20), performance on Star Reading and Star Math, as measured by NPR, did not change significantly from 2019-20 Winter to 2020-21 Spring. Analysis across demographic groups highlights no difference in the way performance changed before and after the transition to online learning, aside from persistent performance differences between subgroups that remained consistent during this period of change.

The changes from 2019-20 Winter to 2020-21 Spring are not very different between demographic groups; that is, online learning did not impact different groups disproportionately. However, consistent with our previous analyses in this series, the pre-online learning performance differences between demographic groups continued to persist.

Appendix A. Demographic Details for the Cohort

Table A.1. Comparison of the Study Cohorts to All Test Takers in the Three 2020-21 Star Testing Windows

	Assessed in 19-20 Winter and 20-21 Spring (Current Cohort)	All 7-11 Students Assessed in 20-21 Spring	Assessed in 19-20 Winter and 20-21 Winter (April Addendum Cohort)	All 7-11 Students Assessed in 20-21 Winter	Assessed in 19-20 Winter and 20-21 Fall (January Report Cohort)	All 7-11 Students Assessed in 20-21 Fall
Star Reading						
Number of Students	19,894	30,128	23,732	36,464	25,263	38,217
Average NPR, 20-21 window	25.40	26.11	24.66	26	24.73	26.1
Star Math						
Number of Students	18,492	29,099	22,335	35,773	23,642	37,231
Average NPR, 20-21 window	39.36	39.23	39.25	40.12	38.29	39.6

Table A.2. Demographic Breakdown of 2019-20 Winter to 2020-21 Spring Cohort

Subgroup	Star Reading		Star Math	
	Number of Students	% of Cohort	Number of Students	% of Cohort
American Indian/Alaskan Native	52	0.3%	46	0.2%
Asian	2488	13%	2406	13%
Black/African American	9124	46%	8483	46%
Hispanic/Latino	3995	20%	3685	20%
Multiracial/Other	1035	5%	951	5%
Native Hawaiian/Pacific Islander	26	0.1%	24	0.1%
White	3174	16%	2897	16%
Non-English Learner	17354	87%	16023	87%
English Learner	2540	13%	2469	13%
Economically Disadvantaged	13736	69%	12877	70%
Not Economically Disadvantaged	6158	31%	5615	30%
Total	19894	100%	18492	100%

Appendix B: National Percentile Rank Analyses Compared to Normal Curve Equivalent-Based National Percentile Rank

National Percentile Rank was used as the measure for student performance in this report and in the other cohort analyses in this series. However, when student assessment scores are normally distributed, which is often the case, NPR is not an interval measure. That is, the range of scores that students receive in the lower and upper ends of the NPR spectrum (around the 1st and 99th percentiles) will be much larger than the range of scores that students receive around the 50th percentile (Figure B.1). For example, scale scores for the 2020-21 Spring Star Math assessment at the 1st percentile correspond to between 0 and 567 (a 567-unit difference). In the same testing window, scaled scores for students at the 99th percentile ranged from 1042 to 1245 (a 203-unit difference).¹¹ In contrast, the scaled score range on the same test for students at the 50th percentile was 793 to 843 (a 50-unit difference). In other words, in order for a student to move from one percentile to the next, the student would have to improve their scaled score by a larger amount at the lower and higher ends of a normal distribution and by a smaller amount around the 50th percentile.

NPR is an easily accessible measure; therefore, we used average NPR for different student groups' performance in our cohort analyses without any conversions. The usual practice is to convert NPR into a "Normal Curve Equivalent (NCE),"¹² which is an interval measure, to average the NCE scores and then to convert those averages back to the percentile equivalents rather than averaging an ordinal scale (such as NPR). Other than a few notable exceptions, the average NPR improvements or declines in our samples were not big. In this report, we repeat the tables for the average NPR-based analyses discussed above with NCE conversions and report them out together. For this analysis, we first converted each student's NPR to NCE scores, then averaged the NCE scores, and converted the average NCE back to NPR ("Average 2019-20 Winter/2020-21 Spring NCE-based NPR" columns). We report the improvement/decline in performance between 2019-20 Winter and 2020-21 Spring in terms of average NCE that is back-converted to an average NPR in the "NCE-based Average NPR Differential" columns. We also calculated the difference between the average NPR differential between the two testing windows we originally reported and the post-NCE conversion back-converted average NPR ("Average NPR-Average NCE-based NPR Difference" columns). This column shows the difference between the two methods of reporting performance.

We report overall performance in terms of average NCE-based NPR in Table B.1. A comparison of improvement or decline based on average NPR and improvement or decline when NCE conversions are performed shows that the difference is always less than 1 percentile. Remember that for Star

¹¹ When scale scores are listed from the highest to the lowest, students who are in the 1st percentile represent the bottom 1% of the whole list. In the same case, 99th percentile will be the group of students who scored higher than 99 percent of the list (i.e. top 1%). The 50th percentile would be the group of students who scored higher than half of the sample (but lower than the group of students whose score places them in 51st percentile).

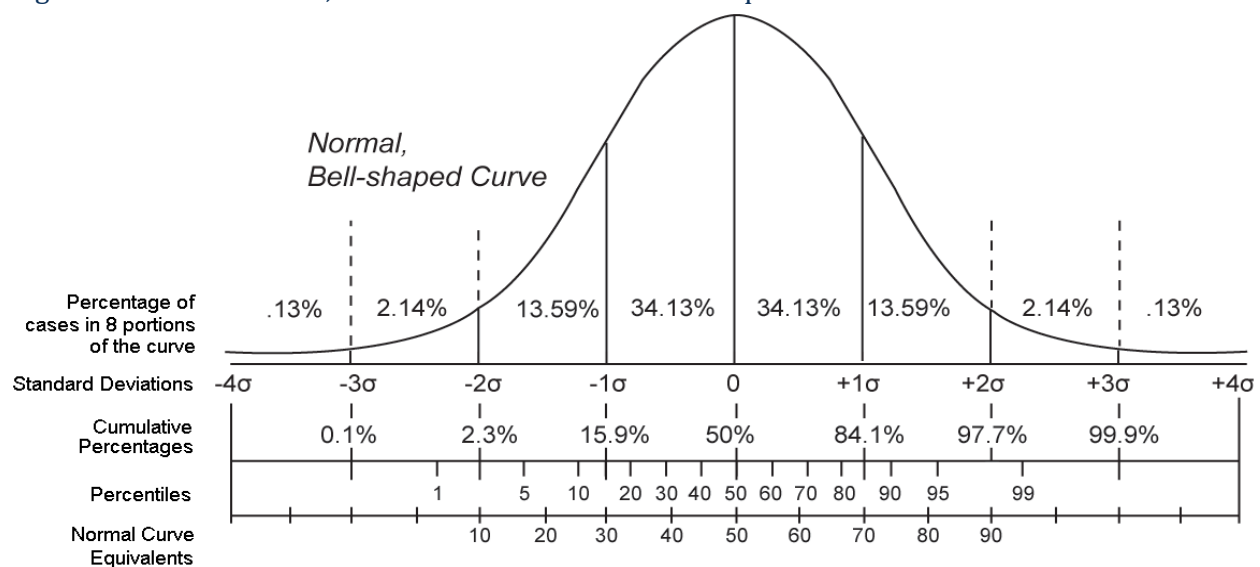
¹² For a more detailed explanation see: "Making Sense of NCEs and Standard Errors" by Pennsylvania Department of Education available at: <https://www.education.pa.gov/Documents/K-12/Assessment%20and%20Accountability/PVAAS/Methodology/MakingSenseOfNCEsAndStandardErrors.pdf>

Reading and Math, overall performance changes Winter-to-next-Fall, Winter-to-Winter, and Winter-to-Spring were usually close to 0. This is still the case when the NPR-NCE conversions are performed.

The comparison of Winter-to-Spring performance discussed in this report and the average NCE-based NPR analysis shows that the direction of change from 2019-20 Winter to 2020-21 Spring was largely consistent between the two methods (Tables B.2-B.11). Any improvement or decline observed with average NPR-based analyses was still an improvement or decline when NCE conversions were used, with a few exceptions. These exceptional cases were marginal improvements or declines close to 0 (no change).

There was variation in the magnitude of the reported change between student performance in the two assessment windows; however, the reported changes based on average NPR and average NCE-based NPR were smaller than 1 percentile. The highest NPR-NCE based NPR difference in the 2019-20 Winter to 2020-21 Spring performance change was 6 percentiles in Star Math performance among Hawaiian/Pacific Islander students, a very small group of 24 students across the cohort. For all other groups analyzed, the highest difference between average NPR and average NCE-based NPR Winter-to-Spring performance changes in Star Reading or Math was 2 percentiles.

Figure B.1. Normal Curve, Percentiles and Normal Curve Equivalents



Source: Public Domain, <https://commons.wikimedia.org/w/index.php?curid=58891646>

Overall Cohort Performance in Star Reading and Star Math Throughout 2020-21 with NCE-based NPR Comparisons

Table B.1. Overall Star Reading and Star Math NPR and NCE-Based NPR for All Assessment Windows Analyzed

Test and Cohort Analyzed	Number of Students	Average 19-20 Winter NPR	Average 20-21 Window NPR	Average NPR Differential	Average 19-20 Winter NCE-based NPR	Average 20-21 Window NCE-based NPR	Average NCE-based NPR Differential	Average NPR-Average NCE-based NPR Difference
Star Reading								
Winter-to-Fall	25263	24.82	24.73	-0.09	16.82	17.03	0.21	-0.30
Winter-to-Winter	23732	24.60	24.66	0.06	16.89	16.62	-0.27	0.33
Winter-to-Spring	19894	25.49	25.40	-0.09	17.49	17.08	-0.40	0.31
Star Math								
Winter-to-Fall	23642	38.93	38.29	-0.63	32.86	32.65	-0.21	-0.42
Winter-to-Winter	22335	36.37	39.25	2.89	30.32	33.42	3.09	-0.21
Winter-to-Spring	18492	40.35	39.36	-0.99	34.75	32.89	-1.86	0.86

How to read this table: This table shows the results for all students included in the Star analyses in this series of cohort studies without any grade or subgroup break-outs. For example, the January 2021 Winter-to-next-Fall report analyzed performance for 25,263 students who were tested for Star Reading in 2019-20 Winter and the 2020-21 Fall, the testing window analyzed in that report. These students had an average NPR of 24.82 in 2019-20 Winter and 24.73 in 2020-21 Fall, a .09 percentile decline. Then, we averaged the NCE scores for these students and calculated the NPR equivalents of the average NCE scores and found that the average NCE-based NPR was 16.82 for 2019-20 Winter and 17.03 for 2020-21 Fall, a .21 percentile improvement.

The last column shows the difference between the two ways of calculating the improvement/decline between 2019-20 Winter and the 2020-21 window noted in the row. In this example, our original analysis based on NPR showed a minor decline rather than a minor improvement and thus underestimated the Winter-to-next-Fall performance change by .30 percentiles.

Winter-to-Spring Star Reading Tables with NCE-based NPR Comparison

Table B.2. Star Reading NPR and NCE-based NPR by Grade, 2019-20 Winter to 2020-21 Spring

Grade in 20-21	Number of Students	Average 19-20 Winter NPR*	Average 20-21 Spring NPR	Average NPR Differential	Average 19-20 Winter NCE-based NPR	Average 20-21 Spring NCE-based NPR	Average NCE-based NPR Differential	Average NPR-Average NCE-based NPR Difference
7	5147	27.30	26.52	-0.78	19.53	17.92	-1.62	0.83
8	4633	26.80	25.57	-1.24	18.69	17.22	-1.48	0.24
9	3284	27.41	25.89	-1.52	19.34	17.32	-2.03	0.51
10	3744	21.33	23.71	2.38	13.69	16.02	2.33	0.05
11	3086	23.49	24.79	1.30	15.56	16.60	1.04	0.26
All Students	19894	25.49	25.40	-0.09	17.49	17.08	-0.40	0.31

Table B.3. Star Reading NPR and NCE-based NPR by Race/Ethnicity, 2019-20 Winter to 2020-21 Spring

Race/Ethnicity	Number of Students	Average 19-20 Winter NPR*	Average 20-21 Spring NPR	Average NPR Differential	Average 19-20 Winter NCE-based NPR	Average 20-21 Spring NCE-based NPR	Average NCE-based NPR Differential	Average NPR-Average NCE-based NPR Difference
American Indian/Alaskan Native	52	29.73	27.27	-2.46	23.35	19.26	-4.09	1.63
Asian	2488	35.22	37.72	2.50	28.40	32.09	3.69	-1.19
Black/African American	9124	20.78	20.66	-0.12	13.52	12.97	-0.54	0.42
Hispanic/Latino	3995	20.32	20.33	0.00	12.39	12.28	-0.11	0.12
Multiracial/Other	1035	28.35	27.68	-0.67	20.46	19.28	-1.18	0.51
Native Hawaiian/Pacific Islander	26	33.92	37.58	3.65	27.98	31.40	3.42	0.23
White	3174	36.82	34.86	-1.96	30.69	27.73	-2.96	0.99
All Students	19894	25.49	25.40	-0.09	17.49	17.08	-0.40	0.31

Table B.4. Star Reading NPR and NCE-based NPR by EL Status, 2019-20 Winter to 2020-21 Spring

English Learner Status	Number of Students	Average 19-20 Winter NPR*	Average 20-21 Spring NPR	Average NPR Differential	Average 19-20 Winter NCE-based NPR	Average 20-21 Spring NCE-based NPR	Average NCE-based NPR Differential	Average NPR-Average NCE-based NPR Difference
Non-English Learner	17354	28.09	27.58	-0.51	20.71	19.63	-1.08	0.58
English Learner	2540	7.71	10.45	2.74	4.05	5.42	1.37	1.37
All Students	19894	25.49	25.40	-0.09	17.49	17.08	-0.40	0.31

Note: English Learner status as of Spring 2020-21.

Table B.5. Star Reading NPR and NCE-based NPR by Economic Disadvantage, 2019-20 Winter to 2020-21 Spring

Economic Disadvantage Status	Number of Students	Average 19-20 Winter NPR*	Average 20-21 Spring NPR	Average NPR Differential	Average 19-20 Winter NCE-based NPR	Average 20-21 Spring NCE-based NPR	Average NCE-based NPR Differential	Average NPR-Average NCE-based NPR Difference
Economically Disadvantaged	13736	22.64	22.75	0.11	14.90	14.64	-0.26	0.36
Not Economically Disadvantaged	6158	31.84	31.30	-0.54	24.23	23.42	-0.81	0.27
All Students	19894	25.49	25.40	-0.09	17.49	17.08	-0.40	0.31

Winter-to-Spring Star Math Tables with NCE-based NPR Comparison

Table B.6. Star Math NPR and NCE-based NPR by Grade, 2019-20 Winter to 2020-21 Spring

Grade in 20-21	Number of Students	Average 19-20 Winter NPR*	Average 20-21 Spring NPR	Average NPR Differential	Average 19-20 Winter NCE-based NPR	Average 20-21 Spring NCE-based NPR	Average NCE-based NPR Differential	Average NPR-Average NCE-based NPR Difference
7	4629	41.97	38.33	-3.64	36.84	31.23	-5.61	1.97
8	4508	41.99	39.93	-2.06	37.08	33.99	-3.09	1.03
9	3157	40.77	39.14	-1.63	35.45	32.23	-3.22	1.59
10	3481	36.23	39.66	3.43	29.21	33.52	4.32	-0.89
11	2717	39.68	40.02	0.33	33.95	33.92	-0.04	0.37
All Students	18492	40.35	39.36	-0.99	34.75	32.89	-1.86	0.86

Table B.7. Star Math NPR and NCE-based NPR by Race/Ethnicity, 2019-20 Winter to 2020-21 Spring

Race/Ethnicity	Number of Students	Average 19-20 Winter NPR*	Average 20-21 Spring NPR	Average NPR Differential	Average 19-20 Winter NCE-based NPR	Average 20-21 Spring NCE-based NPR	Average NCE-based NPR Differential	Average NPR-Average NCE-based NPR Difference
American Indian/Alaskan Native	46	44.15	42.85	-1.30	40.12	37.09	-3.03	1.73
Asian	2406	64.91	65.67	0.76	68.64	69.63	0.99	-0.23
Black/African American	8483	31.69	31.22	-0.47	24.31	23.11	-1.20	0.73
Hispanic/Latino	3685	34.24	33.44	-0.80	27.05	25.50	-1.55	0.75
Multiracial/Other	951	40.49	38.64	-1.85	34.28	31.31	-2.97	1.12
Native Hawaiian/Pacific Islander	24	57.92	53.88	-4.04	60.50	50.84	-9.66	5.62
White	2897	52.83	48.91	-3.92	51.81	46.11	-5.70	1.78
All Students	18492	40.35	39.36	-0.99	34.75	32.89	-1.86	0.86

Table B.8. Star Math NPR and NCE-based NPR by English Learner Status, 2019-20 Winter to 2020-21 Spring

English Learner Status	Number of Students	Average 19-20 Winter NPR*	Average 20-21 Spring NPR	Average NPR Differential	Average 19-20 Winter NCE-based NPR	Average 20-21 Spring NCE-based NPR	Average NCE-based NPR Differential	Average NPR-Average NCE-based NPR Difference
Non-English Learner	16023	42.14	40.63	-1.51	37.19	34.61	-2.58	1.07
English Learner	2469	28.73	31.09	2.37	20.72	22.73	2.02	0.35
All Students	18492	40.35	39.36	-0.99	34.75	32.89	-1.86	0.86

Table B.9. Star Math National Percentile Rank by Economic Disadvantage, 2018-19 Winter to 2019-20 Spring

Economic Disadvantage Status	Number of Students	Average 19-20 Winter NPR*	Average 20-21 Spring NPR	Average NPR Differential	Average 19-20 Winter NCE-based NPR	Average 20-21 Spring NCE-based NPR	Average NCE-based NPR Differential	Average NPR-Average NCE-based NPR Difference
Economically Disadvantaged	12877	37.35	36.38	-0.97	30.85	29.06	-1.79	0.83
Not Economically Disadvantaged	5615	47.24	46.18	-1.06	44.26	42.34	-1.93	0.87
All Students	18492	40.35	39.36	-0.99	34.75	32.89	-1.86	0.86