RENAISSANCE

SPECIAL REPORT SERIES | SPRING 2021 EDITION

SPRING

How Kids Are Performing

Tracking the School-Year Impact of COVID-19 on Reading and Mathematics Achievement

RENAISSANCE LEARNING
P.O. Box 8036
Wisconsin Rapids, WI 54495-8036
(800) 338-4204
www.renaissance.com

7/21

Contents

Figures

Tables

Executive Summary

Renaissance's *How Kids Are Performing* reports have documented the extent to which the disruptions to teaching and learning caused by the COVID-19 pandemic have impacted US student achievement in reading and mathematics in grades 1–8. Two prior reports presented analyses of performance and growth data at Fall and Winter of the 2020–2021 school year. This Spring report summarizes performance and growth data across the entire school year and provides estimates as to what extent it would be different had the pandemic not occurred.

To estimate the impacts of the pandemic on learning, we tracked the progress of more than 3.3 million students who completed adaptive Renaissance Star Assessments in the 2019–2020 school year (prior to school shutdowns) and throughout the 2020–2021 school year. We established reasonable estimates for how each student would have been expected to grow and perform during the 2020–2021 school year had the pandemic not disrupted teaching and learning. Each student's observed performance during the 2020–2021 school year was then compared to an expected score, with results presented by subject, grade, and subgroup. Where differences were apparent, we attributed those to the impact of the pandemic on student achievement. For more details, see *Appendix A. Sample and Methods* and *Appendix B. Limitations*.

To estimate the impacts of the pandemic on learning, we tracked the progress of more than 3.3 million students who completed adaptive Renaissance Star Assessments throughout the 2020–2021 school year.

In our Fall and Winter reports, we concluded the following:

- In both fall and winter, negative impacts of the pandemic were apparent in both reading and mathematics. The shutdown of school buildings in Spring 2020 and further disruptions of learning into the 2020–2021 school year caused students to grow more slowly in both subjects than would have been expected in typical years.
- This lag in growth has caused students to be farther behind in subject-area learning progressions than in typical years.
- Negative impacts were more pronounced in math in both fall and winter.
- Students who are Black, Hispanic, or American Indian or Alaska Native experienced slightly greater negative impacts than the overall averages. Students attending schools serving low-income families (Title I) were also more negatively impacted, as were students attending schools that were public as opposed to private, and schools in rural or town areas as opposed to urban or suburban areas.

After analyzing fall-to-spring data for this Spring report, we found:

1. **Overall, while students are growing, consecutive seasons of below-typical growth have caused reading and math performance to fall farther and farther behind pre-pandemic expectations.**

   - On the Star Early Literacy and Star Reading scales, students ended the 2020–2021 school year, on average, 8 points behind expected pre-pandemic performance. This equates to a loss of 4 Percentile Rank (PR) points. We approximate it would take students about 7 weeks to make up this difference, but estimates vary by grade (ranging from 3 weeks in grades 1–3 to 14 weeks in grade 8).
   - On the Star Math scale, students ended the school year an average of 16 points behind expectations. This translates to a loss of 11 Percentile Rank points. For example, if the average student would have typically ended the year at PR 50, this year that student is at PR 39. (Note, Percentile Rank scores and growth norms cited in this study are based on pre-pandemic student data.) We estimate it would take students about 11 weeks to make up this difference but this varies considerably by grade (ranging from 5 weeks in grade 2 to 15 weeks in grade 6).

2. **The negative impacts of the pandemic vary widely by subgroup, with many subgroups experiencing staggering setbacks.**

   - We estimate that Black or African American students finished the 2020–2021 school year, on average, 19 and 11 Percentile Rank points behind where they would have been in math and reading, respectively, had the pandemic not occurred. Note, these are COVID-19 impact estimates only, which do not speak to any pre-existing achievement gaps or to the degree to which students had or did not have an equitable opportunity to learn before the pandemic.
   - Students who are Hispanic or Latino and American Indian or Alaska Native also experienced impacts that were more severe than the overall averages, as did students with disabilities, English Language Learners, and students attending urban or Title I schools.
   - It can also be instructive to examine results by combinations of subgroups to understand whether there are compounding effects. For instance, students who attended Title I schools in urban areas appear to have experienced even greater negative impacts.

With the *How Kids Are Performing* reports, our intention is to provide an overview of national achievement trends in the 2020–2021 school year and the impact of COVID on teaching and learning, as well as to support educators in interpreting and acting on their own data to identify needs and find the best possible path forward for every student. Look for additional resources in the coming months.

Introduction

Following the nearly nationwide closing of public and private school buildings in March 2020 in response to COVID-19, school leaders had to offer a shifting menu of learning modes for K–12 students during the 2020–2021 school year. After a year where most students participated in some form of remote schooling, by May 2021, the country had reached a milestone. An estimated 50 percent of elementary and middle school students were back to learning fully in-person (up from 35 percent in February), while 23 percent were hybrid (combining some in-person and some remote learning) and 26 percent were still fully remote.[1]

The novelty of widespread remote instruction, concerns about student access and engagement, temporary lack of national and state assessment data, and broader impacts of the pandemic on the lives of educators, parents, and students have generated urgent questions about achievement and growth. To estimate the impact of the Spring and Summer 2020 shutdowns, in November 2020, Renaissance published the first of a series of reports titled *How Kids Are Performing: Tracking the Impact of COVID-19 on Reading and Mathematics Achievement*. In April 2021, we released a second edition examining student performance and growth in the first half of the 2020–2021 school year. This third report covers the entire school year.

Our *How Kids Are Performing* report series has involved several million students from grades 1–8 in all 50 US states and the District of Columbia, who took Renaissance Star Assessments for early literacy, reading, and mathematics. In the studies, we estimated how each student would have been expected to perform in Fall 2020, Winter 2021, and Spring 2021, based on historical (pre-COVID) data, and compared that to observed (actual) performance at all three time points. Results were presented overall by subject and further disaggregated by grade and subgroup.

The questions addressed in this report are:

1. How much did students grow during the 2020–2021 school year compared with typical school-year growth?
2. How much has the pandemic impacted student performance by the end of the 2020–2021 school year?
3. What are the pandemic's impacts in instructional terms?

To answer these questions, we drew on a subsample of students from our prior *How Kids Are Performing* studies. Specifically, we included students who took Star tests in fall of the 2019–2020 school year, and again in both fall and spring of the 2020–2021 school year. The sample includes 3.3 million students in total from all 50 states plus DC. For details, see *Appendix A. Sample and Methods* and *Appendix B. Limitations*.

---

[1] U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2021). *Monthly school survey: Grade 4 and grade 8 combined school survey, month 5.* https://ies.ed.gov/schoolsurvey/g4g8combined_school_survey_Month5.xlsx

About Renaissance Star Assessments

Star Assessments are uniquely positioned to answer performance and growth questions arising for US educators and students as a result of the COVID-19 pandemic. At their core, Star tests are purposeful, proven, powerful, and predictive.

Star is an interim assessment that is administered periodically, usually 3–4 times, throughout the school year for screening, benchmarking, and progress monitoring. (Interim tests fall in-between daily/frequent formative activities and end-of-year state summative tests.) Star adaptive assessments inform instructional decisions about individual students and help school leaders understand how all students are performing and growing. With cancellations of state and national testing programs, and other pandemic interruptions, interim assessments like Star have taken on an even greater role for educators.

Students at tens of thousands of schools worldwide take Star Assessments to measure reading and mathematics achievement and growth. For more information, see *Research Foundation for Star Adaptive Assessments: Science of Star:* http://doc.renlearn.com/KMNet/R001480701GCFBB9.pdf. For an independent review of the reliability, validity, and other technical characteristics of Star Assessments, see the National Center on Intensive Intervention tools charts: https://intensiveintervention.org/.

Question 1

How much did students grow during the 2020–2021 school year compared with typical school-year growth?

In both reading and math, fall-to-spring student growth was slightly below typical levels, as measured by Student Growth Percentiles (SGP) in Star Assessments.

Our findings for Fall to Spring 2020–2021 indicate that progress made between the start and end of the 2020–2021 school year was lower than the progress made between the start and end of prior school years. For reading, these findings are largely consistent with the findings reported in our Fall and Winter editions of *How Kids Are Performing*. Findings for math tell a different story. In the Fall 2020 edition of *How Kids Are Performing*, we reported that fall-to-fall math growth was considerably lower than fall-to-fall growth prior to the pandemic, whereas in the Winter 2021 edition, we found that growth for the first half of the 2020–2021 school year was approaching typical pre-pandemic levels. Our findings for math for Spring 2021, which consider the full school year, indicate that growth has now fallen short of the near-typical growth achieved in winter.

Table 1 summarizes SGP results for Fall to Spring 2020–2021. Consistent with patterns noted in the Winter 2021 report, reading growth among students in grades 7 and 8 in Spring 2021 was slightly lower than other grades. Additionally, we observed that math growth among grade 3 students in spring was relatively lower than math growth observed in other grades. Note, on the SGP scale, 50 indicates typical performance, for both subjects and all grades.

Table 1. Star Student Growth Percentile results: Fall to Spring 2020–2021

	Median SGP: Fall to Spring 2020–2021	
	Reading/Early Literacy	Mathematics
Grade 1	46	n/a
Grade 2	48	47
Grade 3	46	41
Grade 4	46	44
Grade 5	45	46
Grade 6	45	44
Grade 7	44	46
Grade 8	44	46
Overall (1–8)	**45**	**45**

Interpreting the metrics

Student Growth Percentile (SGP) compares a student's growth from one period to another with academic peers nationwide, defined as students in the same grade with a similar score history. SGPs range from 1 to 99 with 50 indicating typical growth, and their interpretation is similar to Percentile Rank scores in that lower numbers indicate lower relative growth and higher numbers indicate higher relative growth. For example, an SGP of 75 means the student's growth exceeds the growth of 75 percent of students in the same grade who had a similar score history.

Other key points:

- Star SGPs are time-adjusted, meaning the growth expectations change nearly every day. Therefore, taking an assessment earlier or later than another student would not unfairly advantage or disadvantage a student.
- The data driving the Renaissance SGP model was last updated in Summer 2019, using records from the 2017–2018 school year and two prior school years, and thus it characterizes growth in pre-COVID times.

1a. Did fall-to-spring growth vary by subgroup?

We examined fall-to-spring SGPs for each subgroup to explore growth for the 2020–2021 school year (represented by "S" in figure 1), and we also considered the extent to which growth during this school year differed from each subgroup's pre-pandemic baseline ("B" in the figure). Because school shutdowns in spring of 2020 led many schools to forgo interim testing during that period, these baseline growth values were taken from the 2018–2019 school year.

As figure 1 shows, when we disaggregated results by **race and ethnicity**,[2] all groups demonstrated rates of growth that were below pre-pandemic baselines. Growth among Black, Hispanic, and American Indian or Alaska Native students was further below the overall median than all other racial and ethnic groups for both reading and math. For reading, growth was between 6 and 8 points below pre-pandemic baselines for Black, Hispanic, and American Indian or Alaska Native students; for math, growth was between 11 and 14 points below pre-pandemic baselines for these students.

Growth for **students with disabilities** and **English Language Learners** was below the overall median and below pre-pandemic baselines for both reading and math. Students with disabilities were 6 and 9 points below baselines for reading and math, respectively. English Language Learners were 5 and 10 points below baselines for reading and math, respectively. (Note, these student characteristics were available only for a small portion of our sample. See *Appendix B. Limitations*.)

Reading growth among the various **school locales** was 4 points below pre-pandemic baselines. For math, growth in rural schools was 1 point below the pre-pandemic baseline and growth in urban schools was 9 points below baseline, with the other school locales falling in between. Growth among urban students was below the overall median for both reading and math, whereas growth among other school locales was at or above the overall median for both subjects.

Growth among students attending schools categorized as **Title I Schoolwide** was 5 points below the pre-pandemic baseline for reading and 7 points below for math. Growth among students attending schools categorized as Title I Schoolwide was also below the overall median for both reading and math.

For reading, growth among students at **Catholic or other private schools** was consistent with pre-pandemic baselines, whereas growth among students at **public schools** was 4 points below the pre-pandemic baseline. For math, growth among students at Catholic or other private schools was 1 point above the baseline, whereas growth among students at public schools was 6 points below baseline. Growth for students attending Catholic or other private schools was above the overall median for both reading and math, while growth among students attending public schools was at the median for reading and slightly below the median for math.

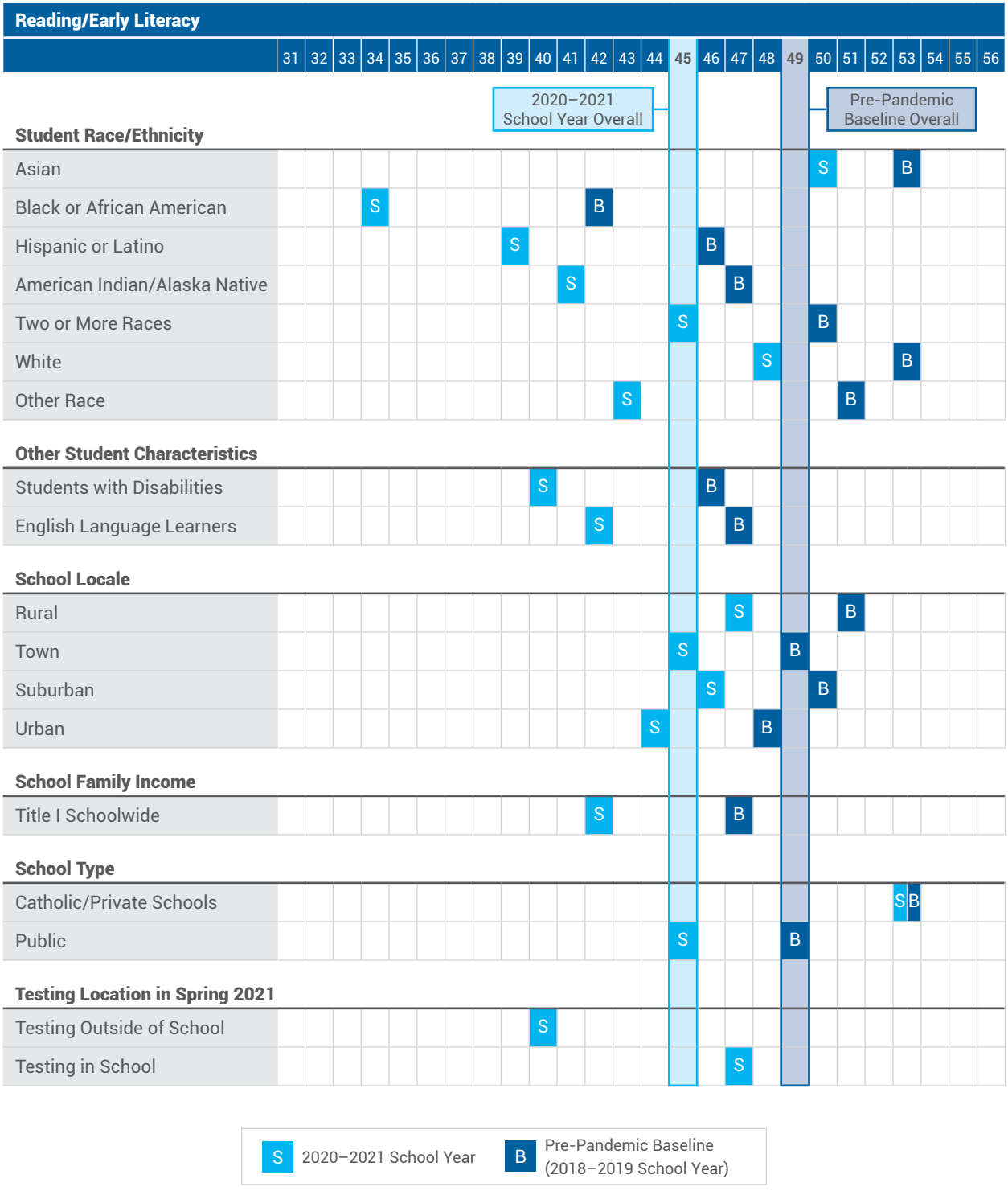
---

[2] The race and ethnicity terms used in this report follow those used by the National Center on Education Statistics, which provides standards for uniformity and comparability in how student subgroups are defined and communicated. For the sake of brevity, we often use Hispanic to represent Hispanic or Latino. Likewise, we may use Black to represent Black or African American, and American Indian to represent American Indian and Alaska Native. We recognize that the language is imprecise and often will fall short in capturing the way individuals may identify themselves. As federal agencies and educators continue to evolve in how they address questions of identity, equity, and access, we will strive to remain as accurate and inclusive as possible.

A feature in Star first introduced in early December 2020 allowed us to report results by the location the test was administered: in or outside of school. Fall-to-spring growth for students **testing outside of school** was lower than growth for students **testing in school** for both reading and math. (Note, test location designations are based on self-reported responses of where tests were taken and are not indicative of whether students were receiving primarily remote instruction, in-person instruction, or a mixture of both). Because reporting by test location was not available prior to the 2020–2021 school year, pre-pandemic baselines for test location are not available.

Overall SGPs were about 7–9 points higher for students testing in school. Examining results by grade, larger differences emerged for younger students. In both subjects, students in grades 2–4 who took tests in school achieved median growth rates that were between 11 and 17 points higher than students testing outside of school. Differences between in-school and remote testers were much smaller in higher grades.

Figure 1. Median fall-to-spring Student Growth Percentile by subgroup



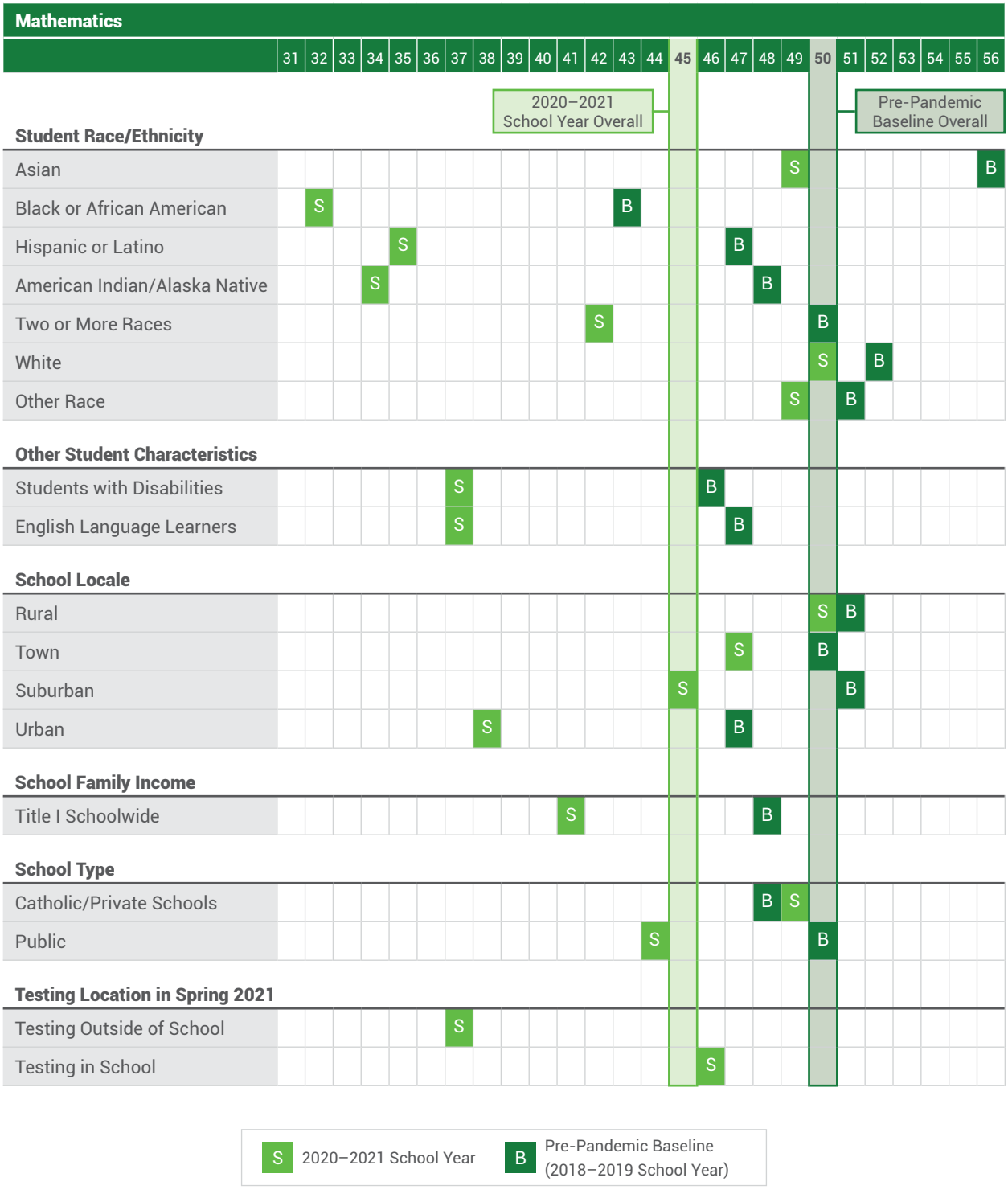
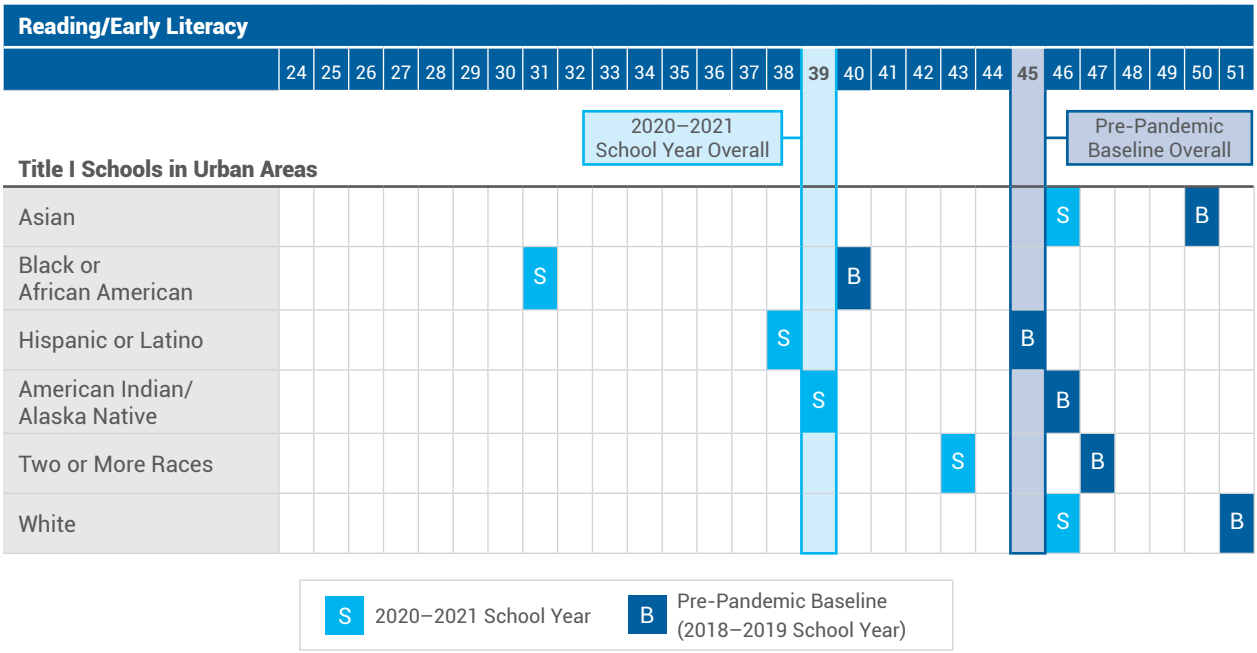
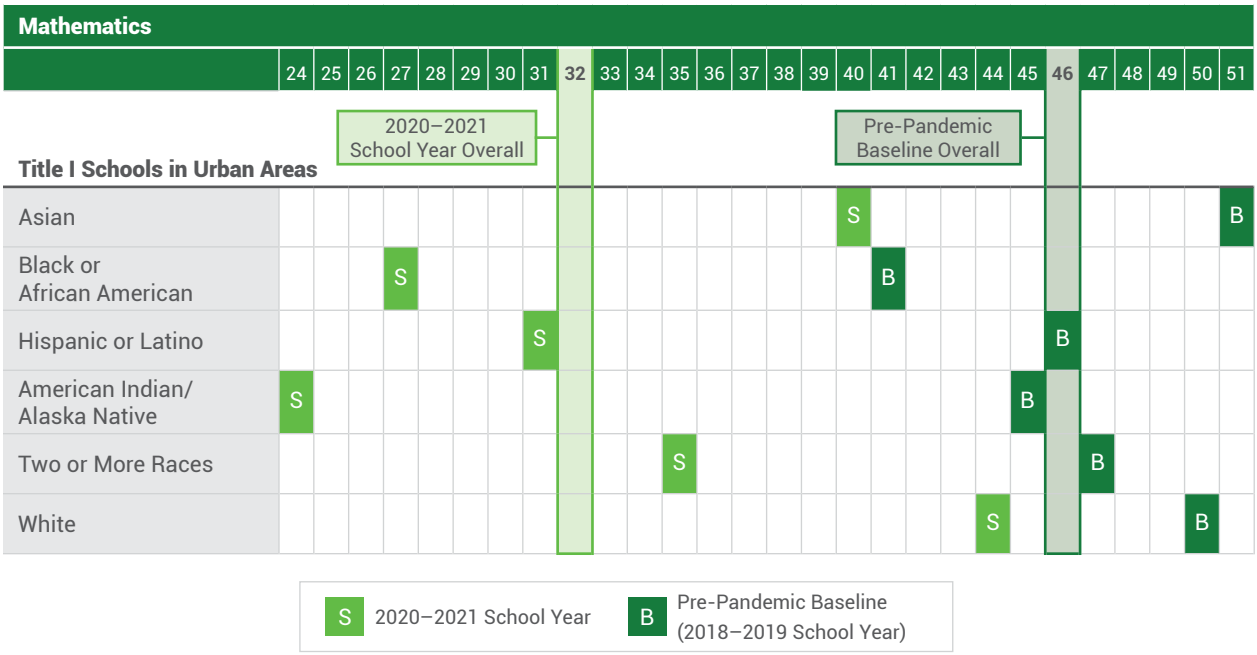


Figure 1a drills into the data for students attending **Title I schools in urban areas**, who experienced some of the greatest negative impacts, with fall-to-spring growth 6 points below their pre-pandemic baseline for reading and 14 points below for math. At those schools, Black, Hispanic, and American Indian or Alaska Native students appear to have been disproportionately impacted by the pandemic.

Figure 1a. Median fall-to-spring Student Growth Percentile by subgroup: Title I schools in urban areas





Question 2

How much has the pandemic impacted student performance by the end of the 2020–2021 school year?

By the end of the 2020–2021 school year, the average reading and math performance of students has fallen further behind pre-pandemic expectations, with math achievement more impacted than reading.

Table 2 summarizes mean expected versus observed (actual) student performance in Spring 2021. On the Star Early Literacy and Star Reading assessments, on average, students scored 8 Scaled Score (SS) points behind their expected pre-pandemic performance. On Star Math, on average, students scored 16 SS points below their expected performance. These differences translated to students being 4 Percentile Rank points lower than expected in reading and 11 PR points lower than expected in math.

Interpreting the metrics

- **Unified Scaled Scores** are calculated based on the difficulty of questions and the pattern of responses. Unified Scaled Scores are useful for comparing student performance on Star Early Literacy and Star Reading over time and across grade levels. The Star Early Literacy scale ranges from 200 to 1100 and overlaps with the Star Reading Unified Scaled Score range of 600 to 1400. Star Math Unified Scaled Scores range from 600 to 1400.
- **Percentile Rank (PR)** is a norm-referenced score that provides a measure of a student's achievement compared to other students in the same grade nationally. PRs range from 1 to 99 and indicate the percentage of other students nationally who obtained scores equal to or lower than the score of a particular student. Percentile Rank norms were last updated in pre-COVID times, in Summer 2017. PR 50 represents typical performance. Note, because PRs are not equal-interval, they should not be averaged. Our results here and in the Star software reflect a conversion of PR to an equal-interval metric (Normal Curve Equivalent/ NCE), calculation of averages in NCE, and then a conversion back to PR.

Table 2. Impact on Scaled Score and Percentile Rank performance of students: Spring

Reading/Early Literacy				
	Spring Expected Mean Unified Scaled Score	Spring Observed Mean Unified Scaled Score	Scaled Score Difference (Observed minus Expected)	Percentile Rank Difference (Observed minus Expected)
Grade 1	860	853	-7	-7
Grade 2	960	950	-10	-5
Grade 3	987	982	-5	-2
Grade 4	1022	1013	-9	-5
Grade 5	1046	1038	-8	-4
Grade 6	1067	1059	-8	-4
Grade 7	1085	1076	-9	-4
Grade 8	1100	1092	-8	-3
Overall (1–8)			**-8**	**-4**

Mathematics				
	Spring Expected Mean Unified Scaled Score	Spring Observed Mean Unified Scaled Score	Scaled Score Difference (Observed minus Expected)	Percentile Rank Difference (Observed minus Expected)
Grade 2	937	925	-12	-10
Grade 3	989	971	-18	-12
Grade 4	1032	1011	-21	-14
Grade 5	1063	1045	-18	-12
Grade 6	1083	1068	-15	-9
Grade 7	1097	1086	-11	-7
Grade 8	1109	1100	-9	-4
Overall (2–8)			**-16**	**-11**

2a. Did the impacts vary by subgroup?

As shown in figure 2, when results were disaggregated by **race and ethnicity**, students from all racial/ethnic groups were below pre-pandemic expectations for math. For reading, only Asian students showed aggregate achievement levels that were consistent with pre-pandemic expectations. In addition, students who are Black, Hispanic, and American Indian or Alaska Native have been disproportionately impacted by the pandemic, with spring performance further below expectations than the overall sample.

Students with disabilities showed greater negative impacts than the overall sample in both reading and math, with a larger impact in math. Students with disabilities testing in Spring 2021 were 5 PR points behind pre-pandemic expectations in reading and 14 PRs behind pre-pandemic expectations in math. **English Language Learners** experienced relatively greater negative pandemic impacts than the overall average in both reading and math, again with a larger impact in math. English Language Learners testing in Spring 2021 were 5 PRs behind pre-pandemic expectations in reading and 13 PRs behind pre-pandemic expectations for math. (Note, these student characteristics were available only for a small portion of the sample. See *Appendix B. Limitations*.)

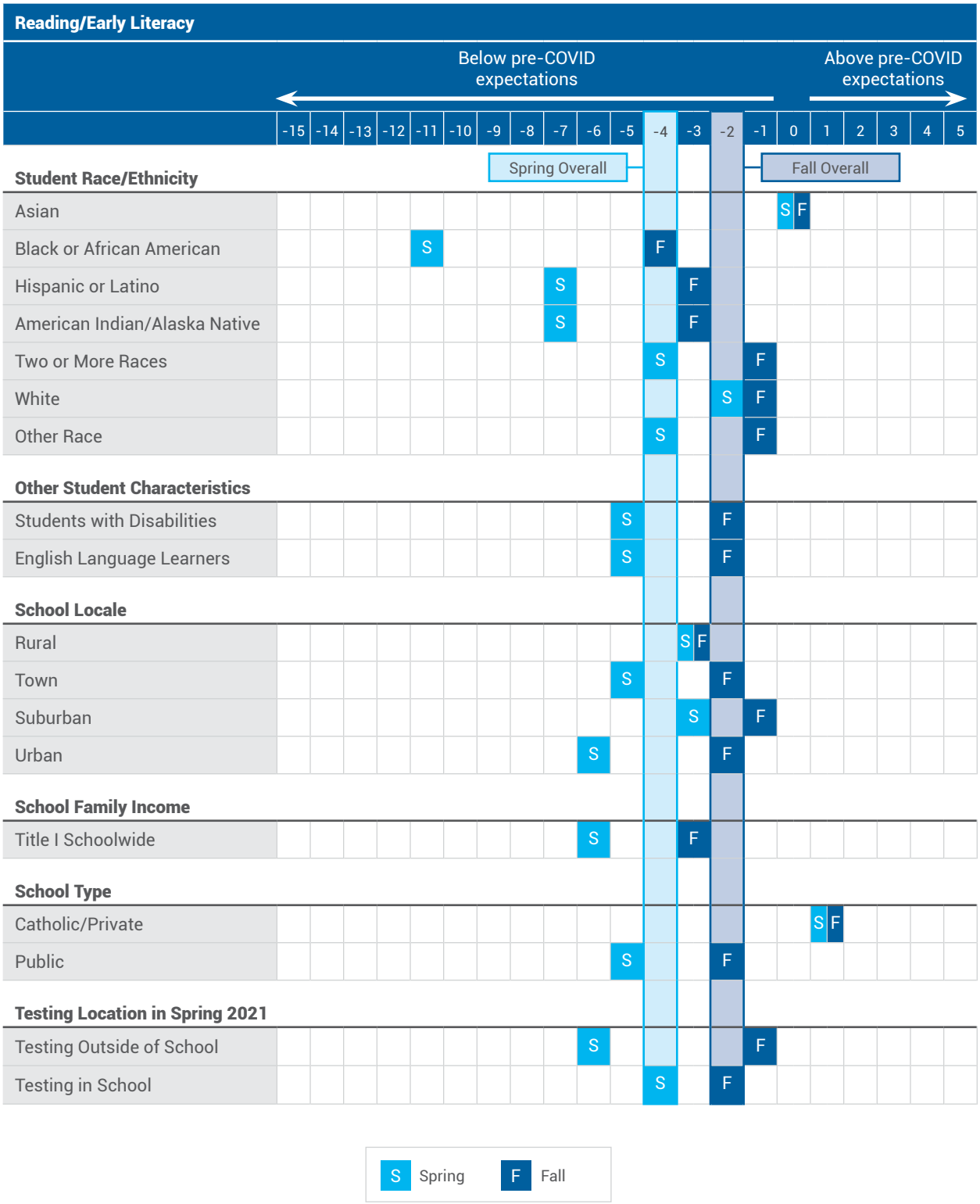
Achievement at each **school locale** was below pre-pandemic expectations. Students at **urban** schools showed the greatest COVID impacts in both reading and math, ending the school year 6 PRs behind pre-pandemic expectations for reading and 14 PRs behind pre-pandemic expectations for math.

Students attending schools categorized as **Title I Schoolwide** experienced a relatively greater negative impact of the pandemic on performance than the overall average for both reading and math. Students at these schools ended the year 6 PRs below pre-pandemic expectations for reading and 13 PRs below pre-pandemic expectations for math.

Students attending **Catholic and other private schools** performed slightly higher in reading than their pre-pandemic expectations and experienced smaller negative impacts in math than average. Student performance at **public schools** was, on average, 5 PRs below pre-pandemic expectations for reading and 11 PRs below pre-pandemic expectations for math.

For reading and math, students **testing outside of school** exhibited greater differences between typical-year performance and COVID-impacted performance than students **testing in school**.

Figure 2. Impact on Percentile Rank performance by subgroup and season



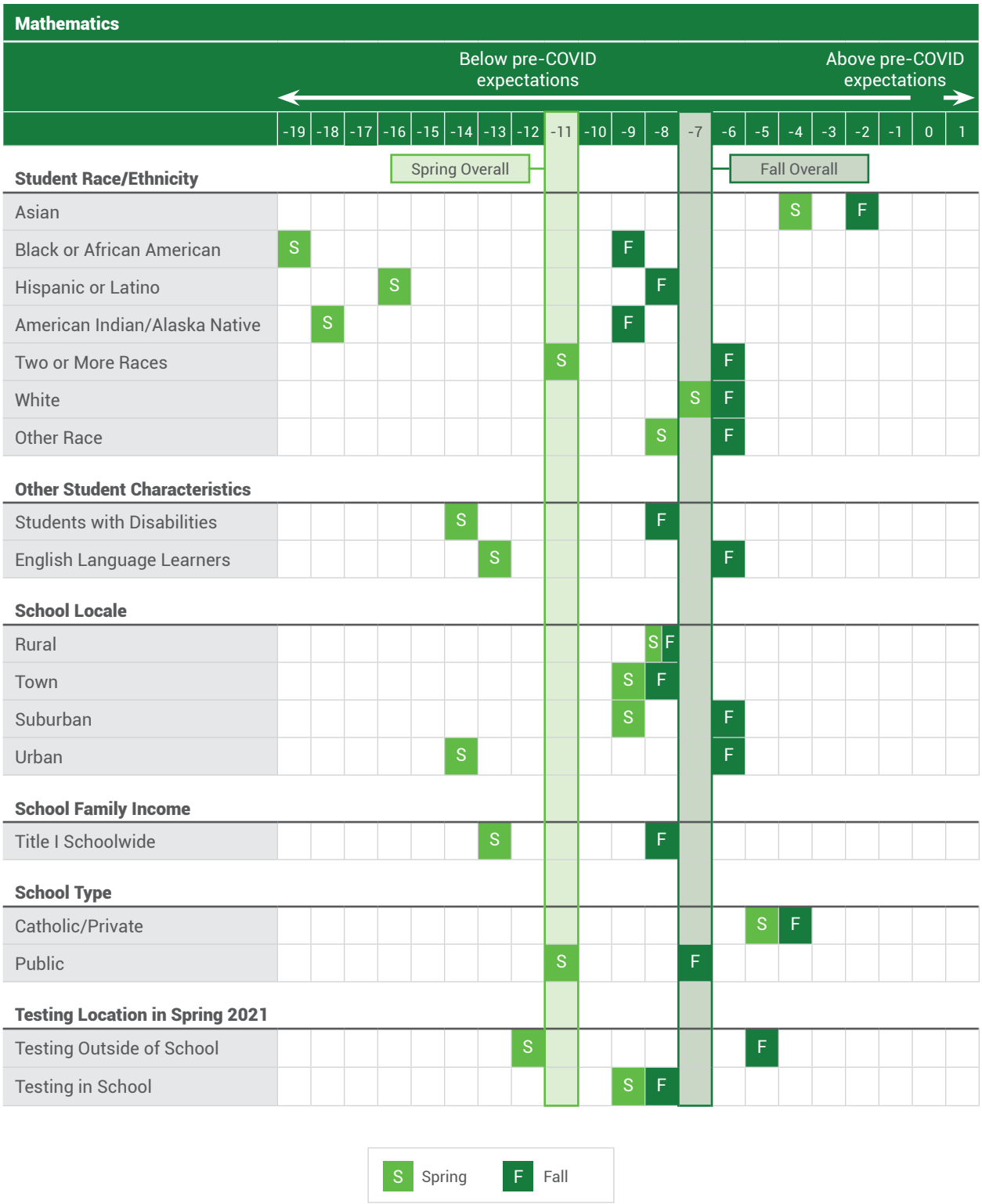
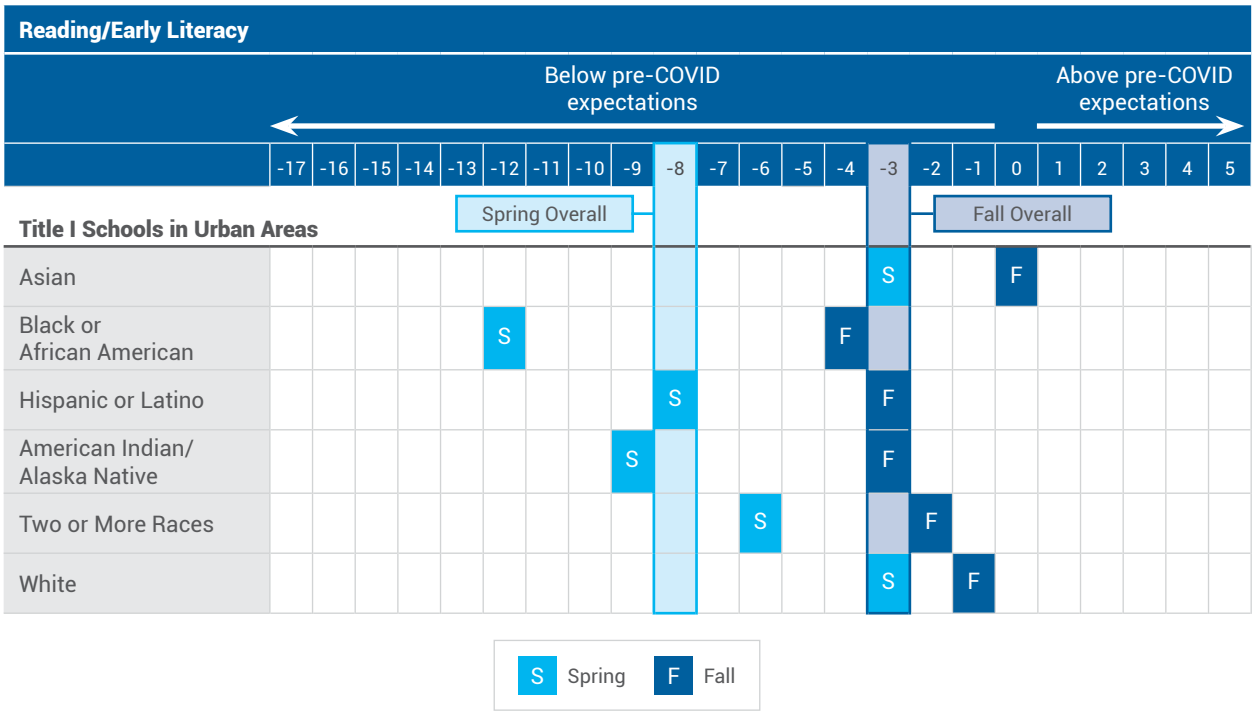
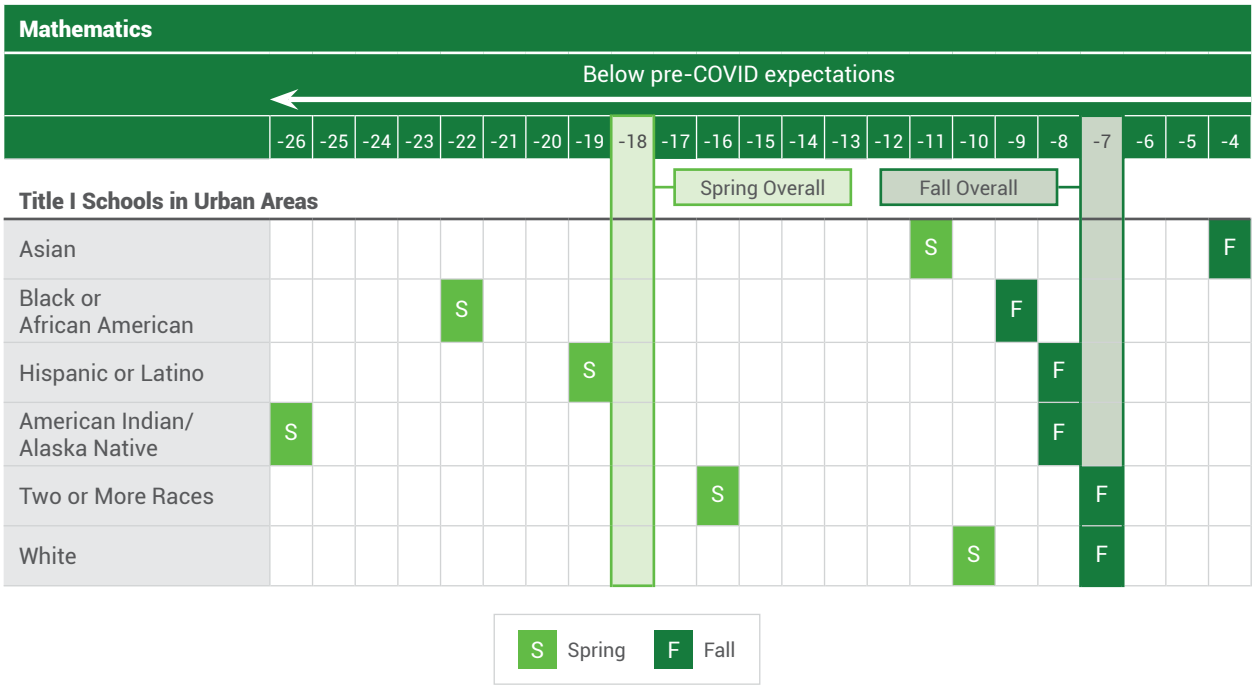


Figure 2a drills into the data for students attending **Title I schools in urban areas**, who experienced some of the greatest negative impacts, ending the year 8 PRs below pre-pandemic expectations for reading and 18 PRs below pre-pandemic expectations for math. At those schools, Black, Hispanic, and American Indian or Alaska Native students appear to have been disproportionately impacted by the pandemic.

Figure 2a. Impact on Percentile Rank performance by subgroup and season: Title I schools in urban areas



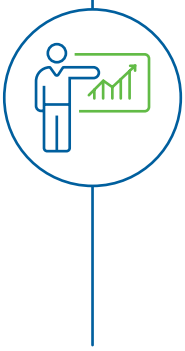


Question 3

What are the pandemic's impacts in instructional terms?

To help contextualize the degree of COVID-19 impacts on learning that have accumulated since Spring 2020, we applied pre-pandemic academic-year growth norms to convert Scaled Score differences into an estimate of how many weeks it would typically take students to close a gap between expected and observed (actual) spring performance. Weeks values from both fall and spring were categorized and are summarized in figure 3. We labeled any weeks estimate that was plus or minus 3 weeks as being approximately "close to expectations." Note, this sample includes only students who tested both in fall and spring, so weeks noted for fall ("F" in figure 3) may differ from what was presented in prior editions of *How Kids Are Performing*, because this is a different subsample of students. (See *Appendix A. Sample and Methods*.)

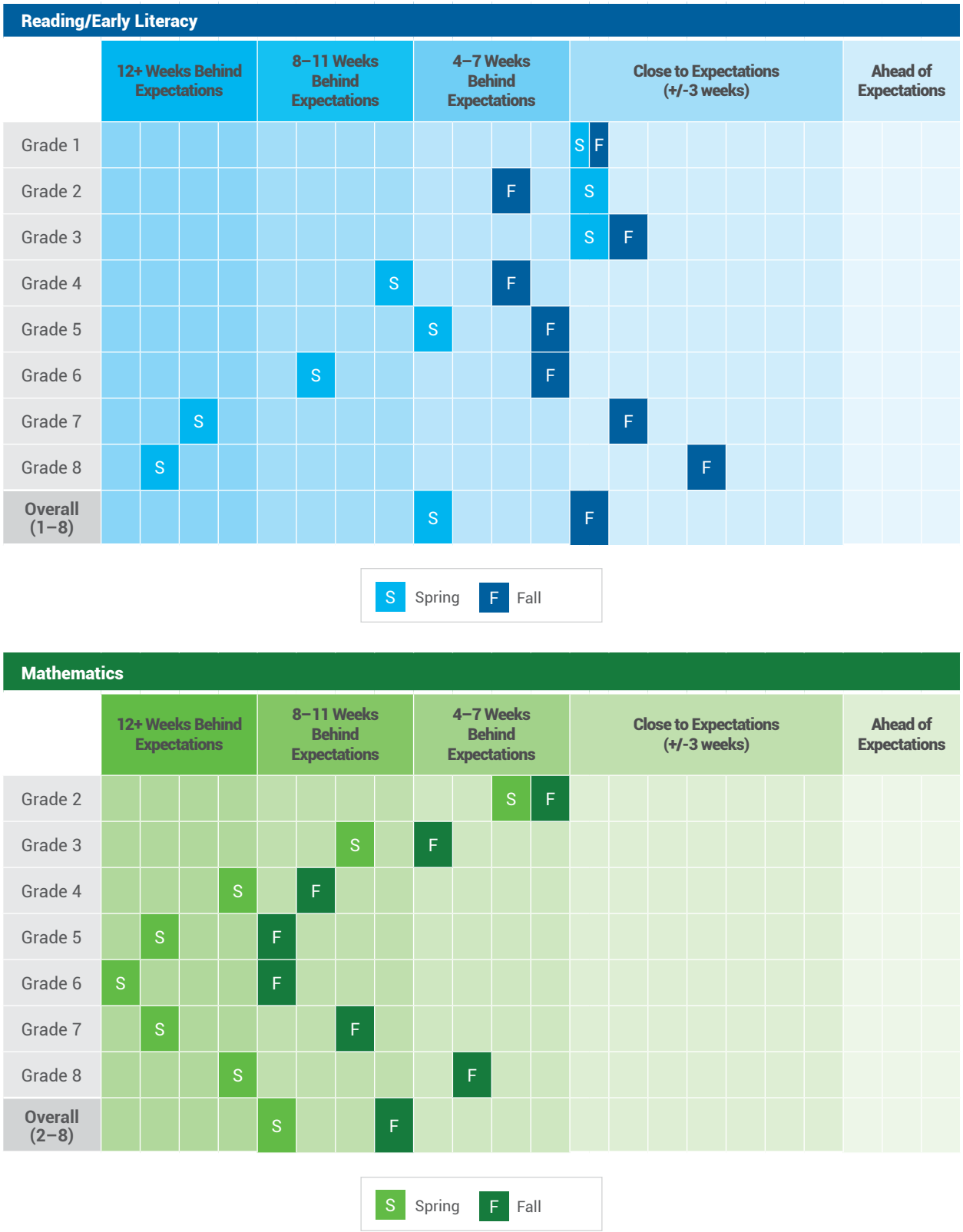
For reading, students in grades 1 through 3 have remained or moved closer to pre-pandemic expectations; however, on average, students in grades 4–8 have fallen further behind expectations. For math, students in all grades remain behind typical-year expectations to varying degrees (4–7 weeks behind expectations for grade 1, and 12+ weeks behind expectations in grades 4 through 8).



Interpreting the metrics

Weeks of instruction is a new metric created especially for this series of studies to help educators and other stakeholders interpret the magnitude of gaps caused by the pandemic. Weeks are a translation of expected versus observed Scaled Scores, divided by weekly academic-year growth rates. Growth rates are set at a student level, and are conditioned on subject, grade, and relative starting achievement level. Weeks for fall could be interpreted as the estimated amount of time for a student to reach beginning-of-year achievement levels. Weeks could be interpreted as the estimated amount of time for a student to reach year-end achievement levels.

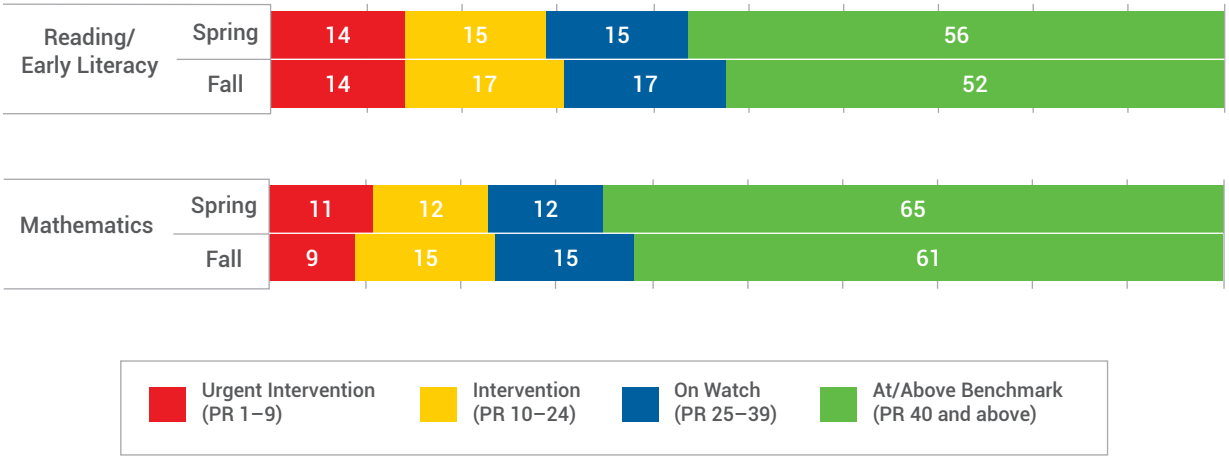
Figure 3. Fall 2020 and Spring 2021 impacts translated to weeks of academic-year instruction



To further explore COVID-19 impacts, additional analyses were conducted using a screening-and-risk assessment perspective. Many schools use a Multi-Tiered System of Support (MTSS) framework to identify and support students who are struggling before they fall too far behind, which involves screening all students with assessments such as Star a few times throughout the year. Students who perform relatively low may be deemed at-risk and receive intervention and other supports in addition to regular classroom instruction.

Figure 4 shows fall and spring MTSS category distributions from the 2020–2021 school year. In spring, we found more students were at/above benchmark compared to fall. This may appear to run counter to our other findings, which show student performance overall has suffered due to COVID-19. However, we have seen this pattern of a higher proportion of students at/above benchmark in the spring relative to fall in pre-pandemic years. In the 2018–2019 school year, there were 7 percent more students at/above benchmark in the spring when compared to fall for both reading and math, whereas this year shows only a 4 percent increase. Also, there was wide variability in COVID-19 impacts on students, so the grouping of students into broad performance categories can obscure impacts on achievement. For example, in math where average student performance was most affected, more students fell in the lowest and highest categories in spring compared to fall.

Figure 4. Percentage of students in MTSS risk categories: Spring 2021 versus Fall 2020



Interpreting the metrics

Using common and consistent risk indicators, students are placed in one of four **MTSS categories**: red (Urgent Intervention), yellow (Intervention), blue (On Watch), and green (At/Above Benchmark). Students in the red category are the lowest performing with the highest risk of experiencing academic difficulties, and thus often receive the most intensive support. Students in the yellow category are slightly higher performing, but often also receive intervention or additional support. Students in the blue category have lower-than-average performance, but typically do not warrant additional services. Students in the green category are generally considered to be at low risk for later difficulties. Note, these risk categories may be adjusted locally and are just one piece of information used by educators to make decisions about individual students. Intervention services are additionally shaped by other relevant factors such as local and state policies, other sources of information about student learning, and available capacity to serve students in need of extra support.

Discussion

In three consecutive *How Kids Are Performing* studies, we have used data from the 2020–2021 school year and historical growth norms to estimate the impact of the COVID-19 pandemic on student achievement. What have we learned?

First, interim tests took on new and important purposes. The primary role of interim assessments like Star has been to provide data to inform instructional decisions and answer questions on which students need additional support and which are responding well to interventions. While this purpose still holds, when most large-scale testing (National Assessment of Educational Progress/NAEP, state assessments) was postponed due to the pandemic, school leaders, researchers, and policy makers turned to interim assessments such as Star to determine what, if any, impact the pandemic's disruptions to education were having on student achievement by state, across the US, by subject, and by student subgroup. Star's high correlations with large-scale testing programs and flexibility of administration allowed it to serve as a reliable proxy.

Second, we learned that, on average, students ended the 2020–2021 school year substantially behind typical pre-pandemic achievement levels. Put more positively, though there were negative impacts on learning, perhaps it is a small victory and testament to the heroic efforts of educators, parents, and students that the impacts were not more severe.

As students began the 2020–2021 school year, the slowdown in learning was immediately apparent in mathematics. By the middle of the school year, those gaps between typical and observed performance appeared to be holding steady or improving slightly in some grades and subgroups.[3] However, by the end of the 2020–2021 school year, the average gap between typical pre-pandemic performance and observed performance had grown in every grade. The gap averaged 16 points on the Star Math scale, or the equivalent of 11 Percentile Rank points. In other words, we estimate that the pandemic's overall effect on student learning in math was equivalent to moving a student who would have finished the year at the 50th percentile to the 39th percentile. To interpret this impact in an instructional context, we estimate that it would take about 11 weeks to make up that ground, based on historical Star Math growth data.

Although reading performance appeared to be more resilient to the effects of the pandemic relative to math, our analysis of scores showed reading growth rates that were below typical. The consequence of several seasons of below-typical growth in reading is that, on average, students finished the 2020–2021 school year 8 Star Reading Scaled Score points and 4 Percentile Rank points below where we estimate they would have performed in any other year. Growth norms data suggest that it would take students approximately 7 weeks to close this gap.

---

[3] Renaissance Learning. (2021). *How kids are performing: Tracking the mid-year impact of COVID-19 on reading and mathematics achievement: Winter 2020–2021 edition*. https://renaissance.widen.net/s/zvq8rnrp9t/r63370

Third, as was widely predicted,[4] we documented wildly differential impacts by subgroup. While it seems that virtually no subgroup was spared some degree of learning lag, the strongest negative impacts were observed among historically marginalized populations. These include students attending Title I schools in urban areas, and students who are Black, Hispanic, American Indian or Alaska Native, as well as students with disabilities or who are classified as English Language Learners. Consecutive seasons of far-below-average rates of growth since the initial school shutdowns in 2020 have resulted in a staggering degree of negative impact on Spring 2021 performance. For example, in math, we estimate that Black or African American students are, on average, 19 Percentile Rank points behind where they would have been had the pandemic not occurred. In reading, Black or African American students are 11 Percentile Rank points behind. Note again, these are COVID-19 impact estimates only; they do not speak to any pre-existing achievement gaps or the degree to which students had or did not have an equitable opportunity to learn before the pandemic.

In addition to capturing pandemic impacts for individual subgroups, our data show that impacts were compounded among students belonging to more than one severely affected group. Since students at urban or Title I schools showed some of the greatest negative impacts, we examined the compounding impacts at Title I schools in urban areas, and the differential impacts by student ethnicity there. At urban Title I schools, Black, Hispanic, and American Indian or Alaska Native students appear to have been disproportionately impacted by the pandemic in that they exhibited the largest declines from pre-pandemic baselines for both growth and achievement.

Let's examine how educators and other stakeholders might address these impacts.

---

[4] Azevedo, J. P., Hasan, A., Goldemberg, D., Geven, K., & Iqbal, S. A. (2021). Simulating the potential impacts of COVID-19 school closures on schooling and learning outcomes: A set of global estimates. *The World Bank Research Observer*, *36*(1), 1–40. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8108634/

Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, *118*(17). https://doi.org/10.1073/pnas.2022376118

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher*, *49*(8), 549–565. https://doi.org/10.3102%2F0013189X20965918

United Nations. (2020, August). *Policy brief: Education during COVID-19 and beyond.* https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2020/08/sg_policy_brief_covid-19_and_education_august_2020.pdf

Next Steps and Instructional Implications

Equity: If you can't see equity gaps, you can't address them

The disaggregated data clearly reveal some very stark realities hiding behind the overall averages. Consider the PR changes in math performance. While the overall change of 11 points tells part of the story, it is significantly more revealing when we consider the disaggregation by race/ethnicity and see, for example, that Black and American Indian or Alaska Native students are performing 18 or more PRs lower than pre-pandemic levels. We are clearly struggling to meet the needs of these students; they have surely been disproportionately impacted. At the same time, Asian and Catholic/private school students are only 4 PRs behind, on average. In this sense, the COVID-19 related disruptions have clearly exacerbated the existing substantial achievement gaps.[5]

The disparity in performance between various demographic groups is a vivid illustration of how disaggregation by multiple key metrics is critical in revealing equity issues. Having assessment data is one thing, but finding equity issues requires being able to associate those data with other critical information and student demographics in order to get a more complete picture.

Some schools do not have the capacity to bring together information from a variety of sources (e.g., interim assessment tools, student information systems). Data warehousing and visualization platforms can be particularly useful in more fully exploring all available metrics.

Accelerated learning (vs. remediation)

Once equity issues are revealed and areas in need of attention are identified, the focus can turn to planning instruction. A significant shift is occurring around approaches to take when students are performing below grade-level expectations. While historically our approach was often remediation, many groups are now advocating for a collection of approaches under the umbrella term *accelerating learning*. This umbrella term, with several specific approaches related to it, is included in back-to-school guidance from the US Department of Education and appears in many states' guidance as well.

A primary emphasis of accelerated learning is maximizing the time students spend with grade-level content through a purposeful consideration of essential prerequisite skills and targeted "just in time" instruction and support. Within this approach, knowledge of essential grade-level skills and necessary prerequisite skills is crucial. Renaissance's Focus Skills Resource Center (www.renaissance.com/focus-skills) can provide detailed information on both essential grade-level skills and necessary prerequisites, and these resources have recently been expanded to also include Focus Skills for Spanish reading (https://www.renaissance.com/focus-skills-spanish/).

---

[5] One researcher examined NAEP data and estimated an 8-year achievement gap in typical grade 4 classrooms: Wiliam, D. (2020, August 31). COVID-19 learning loss: What we know and how to move forward. *Education Week*. https://www.edweek.org/education/opinion-covid-19-learning-loss-what-we-know-and-how-to-move-forward/2020/08

Trip Steps

Clearly, the content area where performance has been most impacted by COVID-19 related disruptions is math. While working to accelerate learning in math, having knowledge of the most difficult skills for students to master at each grade level is insightful.

Through the empirical validation process used by Renaissance in crafting its learning progressions, a subset of disproportionately difficult math skills was identified. We refer to them as *Trip Steps*. If learning is a staircase, then all steps are not created equal. In mathematics particularly, some skills, at some grade levels, are extraordinarily difficult to master. We call these skills Trip Steps because they can cause a stumble in learning, much the same way that an extraordinarily tall step in a staircase can cause an awkward or strained bit of climbing.

Our Learning Science and Content teams continually review the progression of learning in both reading and mathematics; however, we found that Trip Steps are unique to math. An example Trip Step is "Find the area of a rectangle by multiplying side lengths," which is introduced in the latter half of most grade 3 standards documents.

This is an important skill because it lays the foundation for future success in geometry and problem solving. Prior to third grade, students have been engaged in measurement (length and width) and may have been initially exposed to area using tiles and other manipulatives, but in third grade, both multiplication and a new process for finding area—multiplying side lengths rather than counting tiles—are combined.

Educators who better understand the challenge reflected by skills that are Trip Steps—something generally not visible in standards documents or curricula—will be better prepared to ready learners for meaningful engagement with, and a smoother ascent toward mastery of, these skills. For more information, see *Trip Steps for Mathematics* at www.renaissance.com/focus-skills.

Effective practices for marginalized students

Our results show that the pandemic disproportionately impacted educational outcomes for certain student groups, with the largest impacts for Black or African American and American Indian or Alaska Native students. Likewise, our instructional interventions should be disproportionately targeted for those in greatest need. Next we will examine a few research-based recommendations for effective practices for marginalized groups.

Culturally relevant pedagogy, initially introduced by Ladson-Billings[6] to support learning for African American students, is one promising approach and is based on three concurrent ideas: (1) focus on student success and learning; (2) develop cultural competence, which includes an emphasis on student identity and culture of origin; and (3) develop critical consciousness, or an awareness of and ability to analyze and evaluate real-world issues, particularly those associated with inequities. Culturally relevant teaching programs with a focus on the perspectives, histories, and experiences of marginalized communities that were implemented in San Francisco[7] and Tucson[8] showed positive academic outcomes for students across a variety of measures.

---

[6] Ladson-Billings, G. (1992). Liberatory consequences of literacy: A case of culturally relevant instruction for African American students. *The Journal of Negro Education*, *61*(3), 378–391.

[7] Dee, T. S., & Penner, E. K. (2017). The causal effects of cultural relevance: Evidence from an ethnic studies curriculum. *American Educational Research Journal*, *54*(1), 127–166.

[8] Cabrera, N. L., Milem, J. F., Ozan, J., & Marx, R. W. (2014). Missing the (student achievement) forest for all the (political) trees: Empiricism and the Mexican American studies controversy in Tucson. *American Educational Research Journal*, *51*(6), 1084–1118.

Other research suggests that providing teacher incentives that target the lowest performing students can alter teacher behavior (e.g., using more small groups, individualized instruction, and instructional time) as well as improve student outcomes.[9] In contrast, incentive models aimed at increasing average student performance may actually exacerbate existing achievement gaps.[10] Finally, there is a significant body of research to suggest that more instructional time and more time engaged in a rigorous educational program with high expectations lead to better learning outcomes. This is true for students overall, but is especially true for students most impacted by the pandemic. For example, in urban elementary and middle schools, student attendance was significantly related to academic achievement in both reading and math.[11] Another study found increased reading achievement for Black students from high-poverty schools who finished their school year by learning about specific strategies for reading at home and were subsequently provided with free books during the summer.[12]

As our educational systems strive to accelerate learning post-COVID-19, it is critical that interventions are evaluated not only for their general efficacy evidence but also the evidence specific to the relevant student or school groups for which they will be used. By continuing to monitor student progress and disaggregate the results by different subgroups, we can gain a more comprehensive picture of both educational impacts and our students' recovery.

---

[9] Lavy, V. (2009). Performance pay and teachers' effort, productivity, and grading ethics. *American Economic Review*, *99*(5), 1979–2011.

[10] Hill, A. J., & Jones, D. B. (2021). Paying for whose performance? Teacher incentive pay and the black-white achievement gap. *Educational Evaluation and Policy Analysis*, *43*(3), 445–471.

[11] Gottfried, M. A. (2010). Evaluating the relationship between student attendance and achievement in urban elementary and middle schools: An instrumental variables approach. *American Educational Research Journal*, *47*(2), 434–465.

[12] Same, M. R., Guarino, N. I, Pardo, M., Benson, D., Fagan, K., & Lindsay, J. (2018, February). *Evidence-supported interventions associated with Black students' educational outcomes: Findings from a systematic review of research*. REL Midwest. https://files.eric.ed.gov/fulltext/ED581117.pdf

Appendix A. Sample and Methods

The status of interim testing in 2020–2021 and why it matters

To determine whether COVID-19 impacted student achievement, one of the first questions we considered was, who is completing Star Assessments? If the number and composition of test takers in 2020–2021 looked identical to the population taking Star in prior years, we knew we might be able to address impact-related questions, at least in part, by simply comparing aggregated test results by season, subject, grade, and subgroup.

However, although interim assessments like Star were widely used during this pandemic year, there have been shifts in usage that make simple year-over-year comparisons unhelpful with assessing impact. In the Fall edition of *How Kids Are Performing*, we noted that overall testing volume was down in Fall 2020 compared to Fall 2019, and the same was true in winter and spring compared with pre-pandemic testing. Relatedly, the composition of students taking Star was somewhat different.

There are several plausible pandemic-related reasons for the decline in interim testing, including declining enrollments as some students switched to homeschooling or enrolled in other institutions. Also, many students have simply not shown up or participated in remote learning due to lack of access or other reasons.

Our approach

The approach used in this study is the same as in our Fall and Winter studies, which was to estimate where each student would be performing at various points in the 2020–2021 school year had the pandemic not occurred. We then compared the extent to which their observed performance met that expectation, and aggregated results by subject, grade, and student subgroup. Achievement expectations were set using a combination of historical (pre-COVID-19) Star data and a prior Star score for each student. Expectations were time-adjusted to match the date the student tested, meaning we assumed scores would generally be higher in late May than in early April, for example. The extent to which Star scores can be expected to grow and change is a function of subject, grade, and prior performance.

The process began by applying fall-to-fall growth norms to set expectations for how each student would perform in Fall 2020 based on their Fall 2019 performance. Then, for the present study, we grew each expected Fall 2020 score to the date of their spring test using within-year growth norms. All growth expectations—the fall-to-fall and within-year growth models—reflected recent historical (pre-COVID-19) data from millions of US K–12 Star users to set expectations, so that they reflected typical performance over time in pre-pandemic school years. To the extent the expected scores differ from actual observed 2020–2021 scores, we attributed those differences as representing COVID-19's impact.

We restricted the current sample to include only those students who had a history of Star testing over the past two school years. To be included, students had to have taken Star in Fall 2019, Fall 2020, and Spring (April 1 through May 30) of the 2020–2021 school year. Limiting our sample to a consistent set of students over time helped to reduce bias that might have been introduced by allowing the sample to fluctuate.

The second set of sample restrictions was implemented to ensure a reasonable degree of fidelity in Star test administrations. While this is generally not a concern when students test at school under educator supervision, it may be a greater issue with remote administrations. An examination of scores from Spring 2020, when many (if not most) students were testing remotely, revealed a slightly greater proportion of unusual scores compared with prior years. These could have been an indication of students receiving assistance from a family member, or simply not putting forth a reasonable effort. In an abundance of caution, an approach was developed for identifying and removing suspicious student scores that fell well outside what would be expected from a student given their prior performance. These extreme scores could be unusually low or high. In this case, for each student, a range of expected scores was created using the 5th and 95th percentiles of expected growth. Students scoring outside of their range of expected scores were flagged and removed.

Because the sampling strategy (a) required students to have a history of Star test-taking back to Fall 2019 and at both fall and spring of the 2020–2021 school year, and (b) involved the removal of suspicious records demonstrating extreme low or high changes relative to prior performance, the sample reduced from our fall study (5 million students) to about 3.3 million students for this Spring edition. Specific counts by product, grade, and demographic subgroup are available later in this section.

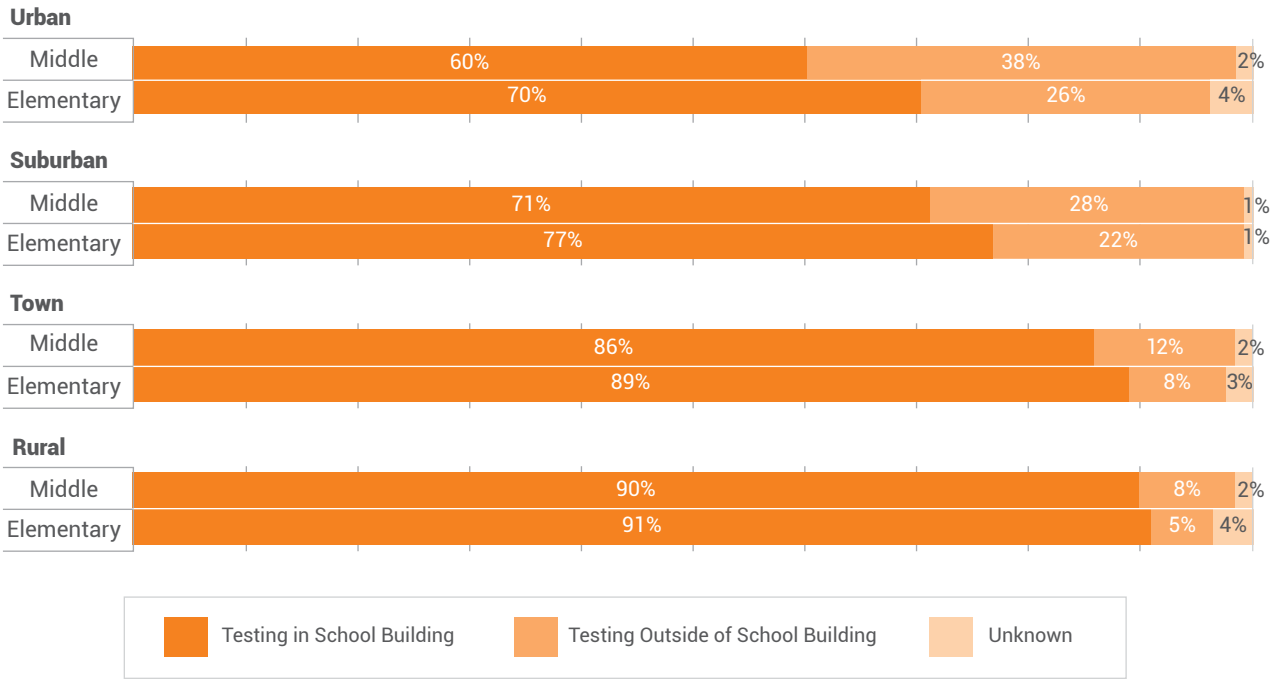
Note: Given that this study included a smaller sample, these changes also resulted in slightly different Fall 2020 results for some grades and subgroups compared to what was published in the Fall edition of *How Kids Are Performing*.

Testing location

One notable shift in testing during the pandemic has been that many schools engaged in remote testing while closed to in-person instruction or as part of hybrid models of instruction. As a result, the sample of students included in this study contained many more students testing remotely than has occurred in prior school years. Beginning in December 2020, Renaissance was able to determine the location of each test based on self-reported information. During the assessment, students indicated whether tests were taken in the student's school or outside of school, allowing us to examine results by test location.

Figure A1 presents Star testing counts for the spring sample by school locale: overall, about 20 percent of Star tests were recorded as remote tests; 78 percent were recorded as in-school; and the remaining tests were taken from an unknown location.

Figure A1. Star test location by school locale



Our sample

Our sample was limited to US schools and consisted of 3,320,662 students: 1.96 million students from 6,954 schools who took Star Early Literacy and/or Star Reading in both 2019 (when they were in grades K–7) and the 2020–2021 school year (grades 1–8), and 1.36 million students from 3,955 schools taking Star Math in both 2019 (grades 1–7) and the 2020–2021 school year (grades 2–8). The specific sample sizes by subject and grade are indicated in table A1. Note that due to smaller sample sizes in high school, this report focused on students through grade 8 only. The sample was national in scope: schools in all 50 states, plus DC, were represented.

Some of our analyses sought to understand whether the pandemic had a differential impact on student subgroups. Some of those categories were defined as individual student characteristics, while others were characteristics of the schools students attended (see table A1).

Note, entry of student-level characteristics in the Star assessment system is voluntary; slightly more than half of the students in our sample had race data, and just about 1 percent were indicated as being a student with disabilities or an English Language Learner. Based on national data, we know the actual rates for these characteristics are higher, but many schools opt to not enter these characteristics, somewhat limiting our ability to present a complete picture of their performance and growth.

Table A1. Sample sizes and school- and student-level characteristics

Student Counts by Spring 2021 Grade		
	Reading/Early Literacy	**Mathematics**
Grade 1	94,717	n/a
Grade 2	147,663	167,996
Grade 3	319,629	225,482
Grade 4	346,197	232,425
Grade 5	332,090	223,122
Grade 6	258,736	185,518
Grade 7	233,195	165,498
Grade 8	231,073	157,321
Overall	**1,963,300**	**1,357,362**

School-Level Characteristics		
	Reading/Early Literacy	**Mathematics**
Number of Schools	6,954	3,955
School Type		
Public	85%	84%
Catholic and Other Private Schools	9%	10%
Unknown	6%	6%
School Locale		
Suburban	34%	34%
Urban	26%	25%
Rural/Non-Metro	20%	20%
Town	15%	15%
Unknown	5%	5%
Title I Schoolwide		
Yes	56%	55%
Unknown	44%	45%

Student Characteristics	Reading/Early Literacy	Mathematics
Race/Ethnicity		
Asian	3%	4%
Black or African American	8%	10%
Hispanic or Latino	16%	16%
American Indian or Alaska Native	1%	1%
Two or More Races	2%	2%
White	26%	29%
Other Race	<1%	<1%
Unknown (not entered)	43%	37%
Disability Status		
Students with Disabilities	1%	2%
Unknown (not entered)	99%	98%
Language Status		
English Language Learners	1%	1%
Unknown (not entered)	99%	99%
Testing Location in Spring 2021		
Star test taken in school building	77%	79%
Star test taken outside of school building	20%	19%
Unknown	3%	2%

Appendix B. Limitations

The first potential limitation of the study was that the sample was constrained to students taking Star tests in Fall 2019, Fall 2020, and Spring 2021. Appendix A summarizes the rationale for restricting our sample to students with a longitudinal testing history. An implication of this sampling strategy, which was reliant on scores from the previous fall, was that we were unable to present data for some grades, as the number of students who use Star Early Literacy prior to Kindergarten, and Star Math prior to grade 1, is relatively small. Another implication was that we were unable to include students who tested on Star for the first time in fall or spring of the 2020–2021 school year. Using this strategy, our sample may be somewhat biased towards certain types of schools or students who were able to maintain Star usage during the past two school years. Additionally, it is possible our results understate the negative impact of the pandemic, because students not captured in our sample may also have been relatively less engaged instructionally. Results may have differed had we included students who did not meet this test-taking criteria.

Related, as discussed in Appendix A, we further restricted our sample to flag and remove student records where low test-administration fidelity was suspected in either the fall or spring of the 2020–2021 school year. Specifically, students' observed growth from a prior period was evaluated against historical growth norms, and extreme values were trimmed. In applying these criteria, legitimate test records may have been removed, representing a potential limitation; however, even if valid records were omitted, this step likely resulted in a net improvement. Given concerns about test fidelity, particularly among out-of-school testers, applying caution and removing these records gave us increased confidence in the data.

A second limitation was incomplete demographic information for all students. Race/ethnicity data were available for just over 50 percent of our student sample, and English Language Learner and disability status was complete for about 1–2 percent of students. In contrast, school-level indicators such as type, locale, and Title I status were available for nearly all students. It is possible that our subgroup results could differ had we had complete demographic data for all students. As we broke down students into various groups based on demographics, smaller sample sizes, in combination with the sampling limitations noted above, may have yielded results not representative of the broader group. For example, results we presented for students at Title I schools in urban areas by race/ethnicity may not fully represent the experience of all students at urban Title I schools. Additionally, factors that cannot be easily measured, such as access to technology, may have been important in contributing to student success during the pandemic.

A third limitation was that although testing location was available for most students in spring, it was not available for tests from the fall. As a result, it was not possible to determine whether students testing in school in the spring also tested in school in the fall, or whether they switched testing locations due to schools reopening. Additionally, testing location data was collected from students' self-reported responses to a question at the beginning of each Star test, but it was not possible to confirm the accuracy of each response. Finally, it is important to acknowledge that testing location data explains only the student's physical location while taking the test but does not capture whether students were receiving primarily in-person, remote, or hybrid instruction. Assumptions that testing location serves as a proxy for mode of instruction, remote or in-school, are unsupported as it is possible that students taking tests outside of their school building were participating in a hybrid model of instruction and were testing on days where in-person instruction was not offered. As a result, the current study cannot address differing learning models and how these may influence student performance and growth.

About Renaissance

As a global leader in assessment, reading, and math solutions for pre-K–12 schools and districts, Renaissance is committed to providing educators with insights and resources to accelerate growth and help all students build a strong foundation for success. Renaissance solutions are used in over one-third of US schools and in more than 100 countries worldwide.

Learn more at www.renaissance.com.

RENAISSANCE

473755.0721
R63444