



Reading achievement declines during the COVID-19 pandemic: evidence from 5 million U.S. students in grades 3–8

Megan Kuhfeld¹ · Karyn Lewis¹ · Tiffany Peltier¹

Accepted: 5 August 2022 / Published online: 17 August 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

The COVID-19 pandemic has been an unprecedented disruption in students' academic development. Using reading test scores from 5 million U.S. students in grades 3–8, we tracked changes in achievement across the first two years of the pandemic. Average fall 2021 reading test scores in grades 3–8 were .09 to .17 standard deviations lower relative to same-grade peers in fall 2019, with the largest impacts in grades 3–5. Students of color attending high-poverty elementary schools saw the largest test score declines in reading. Our results suggest that many upper elementary students are at-risk for reading difficulties and will need targeted supports to build and strengthen foundational reading skills.

Keywords COVID-19 · Academic achievement · K-12 schools

Introduction

Since the onset of the COVID-19 pandemic, there have been widespread concerns about the accumulation of unfinished learning due to ongoing disruptions to schooling. All public schools in the U.S. closed during spring of the 2019–2020 school year, and many of schools operated in remote or hybrid models across much of the 2020–2021 school year. Given the magnitude of the disruptions due to the COVID-19 pandemic, it is not surprising that numerous studies have now documented impacts of COVID-19 on student test scores (Betthäuser et al., 2022; Hammerstein et al., 2021; Thorn & Vincent-Horn, 2021; West & Lake, 2021; Zierer, 2021). Although researchers have consistently found large declines in math test scores since the start of the pandemic (e.g., Halloran et al., 2021; Lewis et al., 2021), research on reading test scores has been more mixed. In fall 2020, reading test scores in grades

✉ Megan Kuhfeld
megan.kuhfeld@nwea.org

¹ NWEA, 121 NW Everett St., Portland, OR 97209, USA

3–8 were mostly consistent with historic averages (Curriculum Associates, 2020; Renaissance, 2020), though more substantial relative declines were reported by spring 2021 (Halloran et al., 2021; Lewis et al., 2021).

Although it is now clear that the pandemic has taken a toll on reading achievement, we still know relatively little about which students were hit hardest and what kinds of supports are needed for recovery. To address these gaps in our understanding, this research explores the impacts of the pandemic on reading achievement in a large national sample. We use reading test scores from fall 2021 (as well as fall data from the two prior years) from 5.2 million U.S. public school students to examine how reading achievement at the beginning of the school year has changed across the course of the COVID-19 pandemic thus far. Specifically, we address three research questions:

1. To what extent have reading test scores in the U.S. changed during the first two years of the pandemic (and how do these trends compare to what has been observed in math)?
2. Do the impacts of learning disruptions on reading achievement differ by grade level?
3. Which groups of students had the most and least change in test scores?

Grade differences

An additional important contribution of this research is examining how the impact of the pandemic may vary across grade. This is particularly relevant for research looking at reading outcomes given how reading skills typically develop. Reading comprehension is comprised of two major components, decoding and language comprehension (Gough & Tunmer, 1986). Decoding is a constrained skill (meaning it is comprised of a limited number of items and can be mastered in a relatively short amount of time). Students typically master decoding phoneme-grapheme correspondences and develop fluency while reading by the end of second grade (Hasbrouck & Tindall, 2017). Language comprehension, however, is an unconstrained skill (meaning it is comprised of an unlimited amount of information) in which students continue to develop throughout their lives (Paris, 2005). During the early grades students develop these components at different rates, with phonics and spelling instruction largely contributing to students' decoding ability (Tunmer & Hoover, 2019). In turn, decoding abilities provide the foundation for students to develop as fluent readers.

Given the sequential nature for how students develop decoding and language comprehension skills, it is likely that age and reading ability at the onset of the pandemic may be an important factor and younger students still in the process of acquiring early building blocks of reading may have been more impacted by pandemic disruptions. Consistent with this, initial evidence from students in K-2 tested during the pandemic indicates that a significantly higher proportion of students are at risk for reading problems compared to pre-COVID-19 students (Amplify, 2022; McGinty et al., 2021). Because most decoding-focused instruction takes place in the

early grades, we hypothesize that those most negatively impacted would be students enrolled in kindergarten through second grade during the onset of the COVID-19 pandemic (who would be expected to be in grades 2–4 by the 2021–2022 school year).

Differences by race/ethnicity and poverty

The evidence is clear that communities of color disproportionately bore the economic, social, and health consequences of the pandemic (Centers for Disease Control and Prevention, 2022). Studies have consistently found that Black, Hispanic, and Indigenous households were at increased risk of contracting and dying from COVID-19 and were also more likely to struggle with food insufficiency during the pandemic (CDC, 2021; Center on Budget and Policy Priorities, 2021). Furthermore, students in majority Black and Hispanic schools reported facing more obstacles to learning during the pandemic, including feeling depressed, stressed, or anxious and having concerns around their own health and the health of their family members (YouthTruth, 2021). Additionally, students in high-poverty households were also less likely to have adequate technical infrastructure, high-speed internet, and quiet learning spaces at home, which were important during remote learning periods (Coleman et al., 2021). Not surprisingly, the cumulative toll of these burdens is evident in education outcomes and achievement gaps across race/ethnicity and income groups that existed prior to the pandemic have widened. For instance, Black and Hispanic students and schools serving poorer communities made learning gains at lower rates during the pandemic compared with their White and higher-income peers (Curriculum Associates, 2021; Dorn et al., 2020).

In this research we also examine how reading achievement declines differ across groups and build on other studies by considering the intersection between race/ethnicity and poverty with regards to students' academic experiences in the pandemic. Race/ethnicity and socioeconomic status are not experienced in isolation but rather intertwined in ways that result in accumulated disadvantage (Nurius et al., 2015). Limited English proficiency status may also be related to factors such as ethnicity and poverty. Thus, a complete understanding of which students have been most impacted requires considering these factors in tandem, which has been largely missing thus far in studies of the impact of the pandemic.

Methodology

Sample

The data for this study are from the anonymized longitudinal student achievement database collected by Northwest Evaluation Association (NWEA), a research-based, not-for-profit organization based in the United States. School districts voluntarily choose to administer NWEA MAP® Growth™ assessments to their students, while students are typically given the option to opt out

of testing if they would like. Assessments typically administered in the fall (usually between August and November), winter (usually December to March), and spring (late March through June). The test scores are typically used to monitor elementary and secondary students' reading and math growth, but are also sometimes used for evaluating educational interventions, as a component of teacher evaluation systems, and as a part of admission decisions for special programs. For more information on how students and teachers use MAP Growth scores, see the MAP Growth technical documentation (NWEA, 2019).

The NWEA data also include demographic information, including student race/ethnicity, gender, and age at assessment. An indicator of student-level socioeconomic status is not available. We measure school poverty level using free or reduced priced lunch (FRPL) eligibility data from the 2019–2020 Public Elementary/Secondary School Universe Survey data file from the National Center for Education Statistics (NCES). We classified low-poverty schools as those with percentage FRPL eligibility less than 25% in 2019–2020, while high-poverty schools were schools with FRPL eligibility greater than or equal to 75% in 2019–2020.

In total, our sample consists of approximately 5.2 million students in grades 3–8 in approximately 12,000 U.S. public schools. We limited our sample of schools to a consistent set of schools that tested at least ten students in a given grade in fall 2019, fall 2020, and fall 2021. This sample restriction guards against the competing explanation that any differences we observe in achievement over time are potentially driven by systematic differences between schools that did and did not consistently test students in all three years. Descriptive information for the students in our sample by grade is provided in Table 1. Overall, the samples of students who tested in 2019 and of same-grade students who tested in fall 2021 were very similar in terms of gender and race/ethnicity, though the number of students tested in each grade was consistently larger in fall 2019. The average NWEA school was remote for approximately 20% of the 2020–2021 school year, though there was wide variability across states and school poverty levels, with high poverty schools spending about 5.5 more weeks in remote instruction than low- and mid-poverty schools (Goldhaber et al., 2022).

Descriptive information for the schools in our sample along with comparison information on the population of U.S. schools is provided in Table 2. Information about U.S. public schools was obtained from the 2019–2020 NCES Public Elementary/Secondary School Universe Survey data file. The schools in our sample represent roughly 12–15% of U.S. public schools in any given grade. Our sample closely matches the U.S. distribution of schools across various locales (urban, suburban, rural, and town). However, our sample reflects schools serving higher average percentages of White students (55% in our sample vs. 49% in the nation), lower average percentages of Hispanic students (20% vs. 26%), and slightly lower percentages of students eligible for free or reduced price lunch (FRPL) relative to national averages (53% vs. 56%).

Table 1 Student demographic characteristics by grade and term

Gr	Term	Counts			Race/ethnicity percentages							
		Students	Schools	Districts	Male	White	Black	Hispanic	Asian	AIAN	Multi-ethnic	Other race
All	All	5,183,849	11,923	3437	51.2	48.8	15.4	20.4	4.3	1.4	3.8	5.9
3	Fall 2019	546,718	7957	2835	51.2	49	16	19.3	4.3	1.4	4.1	5.8
4	Fall 2019	544,708	7877	2797	51.3	48.7	15.8	20.1	4.3	1.5	4	5.6
5	Fall 2019	554,836	7430	2767	51.1	48.8	15.7	20.3	4.2	1.5	3.8	5.6
6	Fall 2019	542,326	4538	2650	51.3	48.4	15.7	20.9	4.1	1.4	3.7	5.8
7	Fall 2019	529,660	3961	2541	51.3	48.4	15.7	21.3	3.9	1.4	3.6	5.7
8	Fall 2019	497,819	3820	2464	51.1	49.2	15.3	20.6	3.9	1.4	3.5	6.1
3	Fall 2020	523,997	7957	2835	51.1	48.8	15.3	19.7	4.6	1.3	4.1	6.1
4	Fall 2020	521,578	7877	2797	51.1	48.6	15.2	20.4	4.5	1.4	4	5.9
5	Fall 2020	518,332	7430	2767	51.2	48.9	15	20.4	4.5	1.4	3.8	5.9
6	Fall 2020	500,028	4538	2650	51	49.6	14.5	20.5	4.5	1.4	3.7	5.9
7	Fall 2020	497,326	3961	2541	50.9	49.4	14.6	20.9	4.3	1.3	3.6	6
8	Fall 2020	481,639	3820	2464	50.9	50.1	14.3	20.5	4.2	1.4	3.5	6.1
3	Fall 2021	518,020	7957	2835	51.1	48	16	19.9	4.4	1.4	4.2	6.1
4	Fall 2021	492,430	7877	2797	51.3	48.1	15.7	20.4	4.5	1.4	4.1	5.8
5	Fall 2021	474,299	7430	2767	51.3	48.7	15.8	19.9	4.4	1.5	3.9	5.8
6	Fall 2021	469,230	4538	2650	51.3	48.5	15.2	20.7	4.2	1.5	3.9	6
7	Fall 2021	469,702	3961	2541	51.3	47.9	15.7	21.2	4.2	1.4	3.7	6
8	Fall 2021	460,450	3820	2464	51.3	48.5	15.2	20.9	4.2	1.4	3.7	6.1

AIAN American Indian or Alaska Native. We limited our sample of schools to a consistent set of U.S. public schools that tested at least ten students in a given grade in fall 2019, fall 2020, and fall 2021. This sample restriction guards against the competing explanation that any differences we observe in achievement over time are potentially driven by systematic differences between schools that did and did not consistently test students in both years

Table 2 Sample school information relative to U.S. population of public schools

	Grade	Number of schools	Average School Enrollment	% FRPL	% White	% Black	% Hispanic	% Asian American	City	Rural	Suburb	Town
NWEA reading sample	3	7957	480	53%	54%	15%	20%	4%	28%	26%	36%	11%
NWEA reading sample	4	7877	484	53%	54%	15%	21%	4%	28%	26%	36%	10%
NWEA reading sample	5	7430	492	53%	53%	15%	21%	4%	28%	26%	36%	10%
NWEA reading sample	6	4538	568	52%	55%	16%	20%	3%	27%	31%	31%	12%
NWEA reading sample	7	3961	592	52%	55%	16%	20%	3%	26%	32%	29%	13%
NWEA reading sample	8	3820	594	51%	55%	16%	20%	3%	26%	33%	29%	13%
U.S. Public Schools	3	53,903	453	56%	48%	15%	26%	4%	30%	26%	33%	10%
U.S. Public Schools	4	53,665	453	56%	48%	15%	26%	4%	30%	26%	33%	10%
U.S. Public Schools	5	52,385	456	57%	47%	15%	26%	4%	31%	27%	33%	10%
U.S. Public Schools	6	37,355	482	57%	49%	15%	26%	4%	29%	31%	29%	11%
U.S. Public Schools	7	32,265	484	56%	50%	16%	24%	3%	27%	34%	27%	12%
U.S. Public Schools	8	32,507	486	56%	50%	16%	24%	3%	27%	34%	27%	12%

FRPL free or reduced priced lunch. The reported samples represent the set of schools that tested at least ten students in a grade in fall 2019, fall 2020, and fall 2021. The sources of the variables are the Common Core of Data (CCD) collected by the National Center for Educational Statistics. The U.S. public school population comparison for each grade was determined by limiting to the schools that were operational in the 2019–2020 school year and enrolled students in that grade level

Measure

Student test scores from the NWEA Measures of Academic Progress (MAP) Growth reading assessments, called RIT scores, were used in this study. MAP Growth is a computer adaptive test that precisely measures achievement even for students above or below grade level. Most tests are administered at school (in a classroom or computer lab), but many tests were administered remotely during the early part of the 2020–2021 school year. Each test begins with a question appropriate for the student's grade level, and then adapts throughout the test in response to student performance. Students respond to assessment items in order (without the ability to return to previous items), and a test event is finished when a student completes all the test items (typically 40–53 items). Each test takes approximately 40–60 min depending on the grade and subject area. MAP Growth scores are scaled using the Rasch item response theory (IRT) model and allow for both within-grade and across-grade level comparisons. The items on each assessment are aligned to state content standards (NWEA, 2019).

NWEA MAP Growth 2–12 is a measure of reading that measures comprehension of various elements of reading, including word meaning (e.g., word origins, semantics) and literacy and informational concepts (e.g., main ideas, inferences, purpose, text structure). This would correspond to a score for overall reading comprehension, as students need to decode the passages or words in order to answer the comprehension questions that follow. The assessments are highly reliable ($\alpha > 0.90$) and have strong correlations ($\rho > 0.70$) with other US state-specific assessments, such as the ACT Aspire, Partnership for Assessment of Readiness for College and Careers (PARCC), and Smarter Balanced Assessment Consortium (SBAC) assessments (NWEA, 2019).

Students receive test scores in each subject/term that are reported on the RIT (Rasch unIT) scale, which is a linear transformation of the logit scale units from the Rasch item response theory model. We also reported scores in standard deviation units, which are described in further detail in the following section. We primarily focus on reading results in this study, though we provide a comparison to math results using a consistent sample of schools.

Methods

To understand how overall reading achievement in fall 2020 and fall 2021 compared to prior to the pandemic (e.g., fall 2019), we standardized the fall 2020 and fall 2021 test scores relative to the mean and standard deviations (SDs) of the fall 2019 MAP Growth test scores (separately by grade level). The resulting estimate \bar{Z}_{21g} represents the standardized difference (in fall 2019 SDs) between the fall 2019 and fall 2021 means:

$$\bar{Z}_{21g} = \frac{\overline{\text{RIT}}_{21g} - \overline{\text{RIT}}_{19g}}{\text{SD}_{19g}}.$$

Standardized mean estimates were calculated in fall 2020 and fall 2021 for each grade/subject. The mean and SDs for each term used to calculate the standardized estimates are reported in Table 3. In all of the analyses presented, we compared test scores

in a single grade relative to same-grade peers from a previous school year and used the term “declines” to refer to changes across cohorts of students (rather than changes in an individual student’s trajectory across school years).

We also further disaggregated the results to examine trends across student subgroups. Specifically, we first compared trends in average test scores from fall 2019 to fall 2021 in low- and high-poverty schools. Second, we compared students by racial/ethnic group separately in low- and high-poverty schools. We translated all the subgroup RIT score means into standard deviation units based on the overall fall 2019 mean and SD. In 2019, the estimate \bar{Z}_{19sg} in grade g and subgroup s represents the difference in SDs between the fall 2019 subgroup mean and the overall mean in fall 2019:

$$\bar{Z}_{19sg} = \frac{\overline{\text{RIT}}_{19sg} - \overline{\text{RIT}}_{19g}}{\text{SD}_{19g}}.$$

In fall 2021, the estimate \bar{Z}_{21sg} in grade g and subgroup s represents the difference in SDs between the fall 2021 subgroup mean and the overall mean in fall 2019:

$$\bar{Z}_{21sg} = \frac{\overline{\text{RIT}}_{21sg} - \overline{\text{RIT}}_{19g}}{\text{SD}_{19g}}.$$

Results

Research question 1: trends in reading test scores

The top panel in Fig. 1 presents trends in average reading test scores (in RIT points) between fall 2019, fall 2020, and fall 2021. Additionally, test score changes are reported in the figure in SD units relative to the fall 2019 test score distribution. Students showed mostly similar reading performance in fall 2020 as before the pandemic (changes ranging from -0.02 to 0.05 SDs by grade). As a point of reference prior to the pandemic, average reading scores held constant (changes ranging from -0.02 to 0.01 SDs within each grade) between fall 2017, fall 2018, and fall 2019 (see Table 3). However, sizable drops in reading achievement occurred between fall 2020 and fall 2021, resulting in a total test score change of -0.09 to -0.17 SDs since the start of the pandemic. The bottom panel in Fig. 1 provides a comparison of the trends observed during the same period in math. In contrast, math test scores declined in a fairly linear fashion across the first and second year of the COVID-19 pandemic, with substantial math test score declines (-0.20 to -0.27 SDs) across the two-year period (fall 2019 to fall 2021).

Research question 2: differences by grade level

While all grade levels showed declines between fall 2020 and fall 2021, we were also interested in understanding which grade levels showed the largest effects in reading. Our analyses show the largest declines are evident for students who

Table 3 Sample sizes, means, and standard deviations by grade, subject, and term

Subject	Grade	Fall 2017*			Fall 2018*			Fall 2019			Fall 2020			Fall 2021		
		N	M	SD	N	M	SD	N	M	SD	N	M	SD	N	M	SD
Reading	3	842,241	187.4	17.0	—	—	—	546,718	187.3	16.8	523,997	188.1	17.4	518,020	184.4	17.7
Reading	4	828,996	197.5	16.7	809,723	197.6	16.6	544,708	197.8	16.6	521,578	197.7	16.7	492,430	195.1	17.4
Reading	5	839,039	204.9	16.5	829,581	204.9	16.3	554,836	205.3	16.4	518,332	204.9	16.5	474,299	202.8	17.1
Reading	6	809,928	210.3	16.4	827,356	210.5	16.1	542,326	210.2	16.2	500,028	210.6	16.1	469,230	208.4	16.9
Reading	7	773,826	214.7	16.5	763,480	214.8	16.1	529,660	214.0	16.5	497,326	214.7	16.5	469,702	212.4	17.0
Reading	8	—	—	—	725,371	218.7	16.2	497,819	217.9	16.7	481,639	218.5	16.8	460,450	216.3	17.3
Math	3	845,222	188.8	13.6	—	—	—	587,432	189.0	13.6	537,667	187.4	14.1	558,484	185.7	14.7
Math	4	831,191	200.6	14.3	812,113	200.5	14.5	597,244	201.0	14.2	549,148	198.4	14.3	564,971	197.2	15.3
Math	5	843,864	210.1	15.7	832,204	210.0	15.8	618,636	210.3	15.5	557,538	208.0	15.5	570,310	206.5	16.3
Math	6	814,015	214.9	16.1	833,825	214.8	15.8	611,635	214.9	15.3	537,406	213.0	15.1	552,265	211.7	15.7
Math	7	772,057	221.5	17.8	766,282	221.4	17.5	599,757	221.4	17.1	529,948	219.5	16.4	556,341	217.7	17.0
Math	8	—	—	—	720,323	226.9	18.7	553,611	226.6	18.4	478,817	224.6	17.9	512,793	222.1	17.9

N = number of students, M = mean, SD = standard deviation

*Data from fall 2017 and fall 2018 is provided as a comparison for pre-COVID trends, but it is pulled from a different sample of MAP Growth test takers [see Kuhfeld et al. (2022) for a description of the sample]

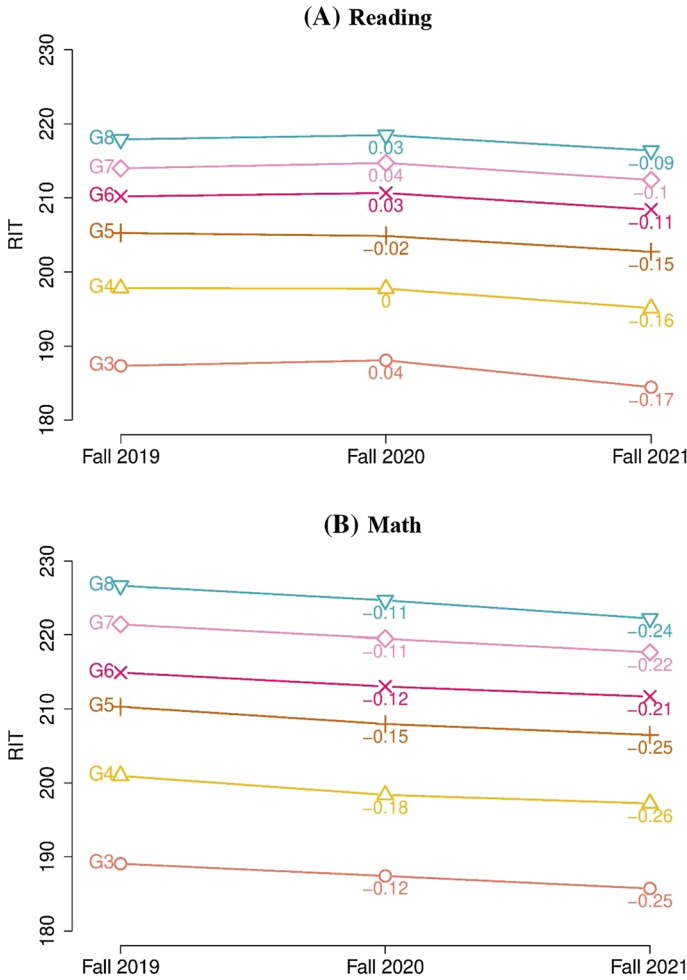


Fig. 1 Changes in MAP Growth test scores in fall 2020 and fall 2021 (relative to fall 2019). *Note.* Reported estimates are calculated based on subtracting the observed mean in each year by the fall 2019 mean and then dividing by the fall 2019 standard deviations (SDs) for each grade

were in grades 3–5 during fall 2021 (e.g., were in grades 1–3 when the pandemic began). The reading declines in the elementary school grades were approximately 1.5 times as large as the test score declines for students in middle school in fall 2021, while the math test declines were fairly similar across grade levels.

Research question 3: differences by school poverty and race/ethnicity

Figure 2 shows differences in average reading achievement between fall 2021 and fall 2019 disaggregated by student grade and school poverty level. This allows us, for example, to situate reading achievement for 3rd graders in high-poverty schools in fall 2021 (0.65 SD below the pre-pandemic overall mean) relative to the reading achievement of 3rd graders in high-poverty schools in fall 2019 (0.40 SD below the pre-pandemic overall mean) and calculate the difference between the two groups (0.25 SD drop). Across all grades and poverty levels, reading achievement dropped between fall 2019 and fall 2021, but the drops were considerably larger for students in high-poverty schools. This was especially notable in grades 3–5, where the achievement drops were 2.5 times larger for students in high-poverty schools compared with low-poverty schools.

Figure 3 displays differences in average reading achievement disaggregated by grade level, race/ethnicity, and school poverty level. The fall 2019 test scores reveal that there were sizable differences across racial/ethnic groups in average pre-pandemic performance within each school poverty level. Additionally, the pattern of test score drops between fall 2019 and fall 2021 is uneven across student groups. In both low- and high-poverty schools, Asian American and White students on average showed declines of a smaller magnitude relative to Hispanic, American Indian and Alaska Native (AIAN), and Black students.

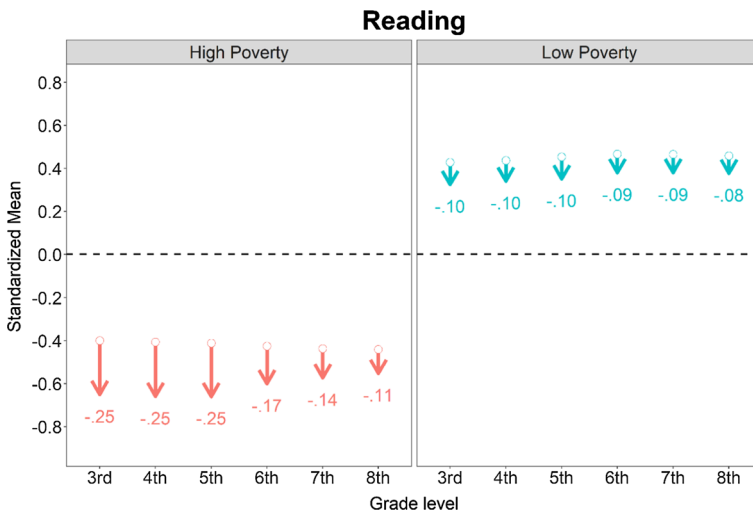


Fig. 2 Changes in standardized MAP Growth reading test scores between fall 2019 and fall 2021 by school poverty level. *Note.* The circles represent the average standardized test score for the pre-pandemic (fall 2019) cohort; the arrow tip represents the average standardized test score for the fall 2021 cohort; and the value outside the arrow indicates the change in averages between fall 2019 and fall 2021. For example, 3rd graders in high-poverty schools in fall 2019 were .40 SD lower than the pre-pandemic sample mean, which increased to .65 SD lower than the pre-pandemic sample mean by fall 2021 (a difference of .25 SD)

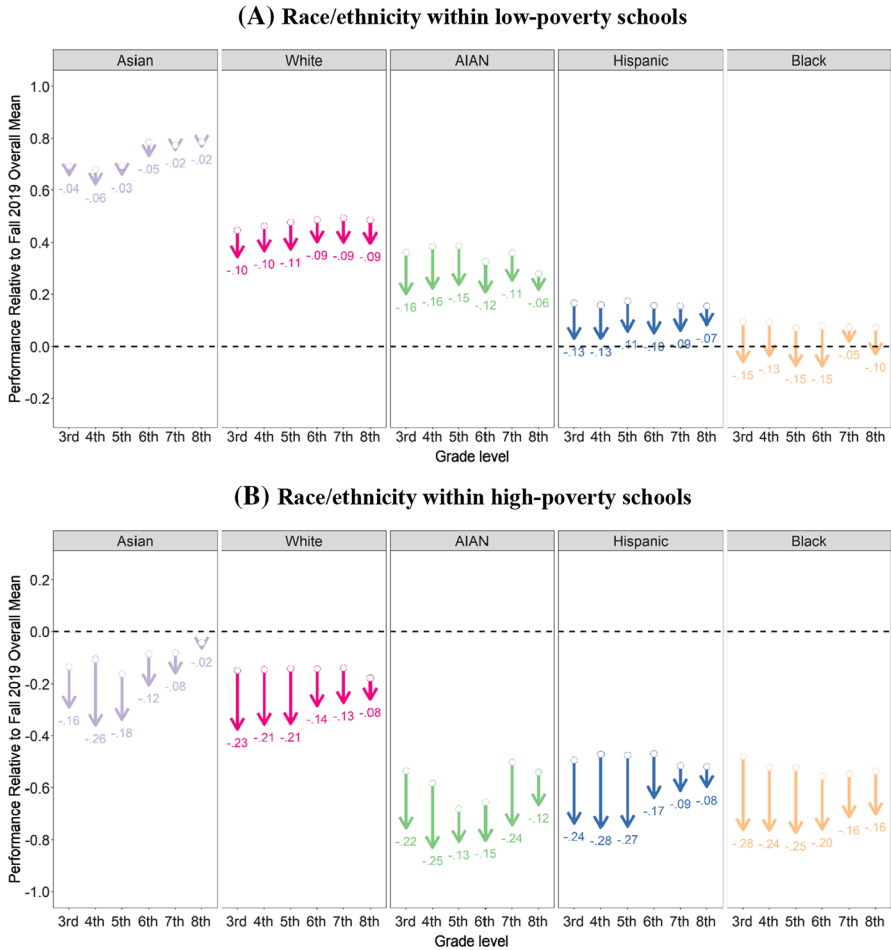


Fig. 3 Changes in standardized MAP Growth reading test scores between fall 2019 and fall 2021 by the interaction of race/ethnicity and school poverty level. *Note.* AIAN=American Indian or Alaska Native. The circles represent the average standardized test score for the pre-pandemic (fall 2019) cohort; the arrow tip represents the average standardized test score for the fall 2021 cohort; and the value outside the arrow indicates the change in averages between fall 2019 and fall 2021

Discussion

Our results add to the abundant evidence amassing showing the profound impacts of the pandemic on education outcomes. Consistent with other reports looking at interim assessments (Curriculum Associates, 2021; Renaissance, 2021) and end-of-year state tests (Halloran et al., 2021), we find that achievement declines relative to pre-pandemic averages are larger amongst students enrolled in high-poverty schools and students of color. Also consistent with other research showing larger educational

impacts of COVID-19 for the youngest students (West & Lake, 2021), our study shows differences in the magnitude of achievement declines by grade. However, our study also adds important nuance to this trend in that we find grade-level differences in the magnitude of achievement declines is dependent on school-poverty level. Specifically, we see differential patterns of achievement declines across grades when we compare high- vs low-poverty schools. Within high-poverty schools, elementary grade students showed larger achievement declines than older students, but in low-poverty schools achievement declines were of a roughly similar magnitude for all the grades in our study. As a result, gaps between students in low- and high-poverty schools disproportionately widened in the elementary school grades relative to middle school grades. The reading test score decline in high-poverty elementary schools was 2.5 times larger than in low-poverty elementary schools, compared to less than two times as large in high-poverty middle schools.

Our data cannot shine light on the mechanisms behind these differential trends. However, other evidence offers some insights that allow us to speculate about the underlying causes of these trends. First, there is mounting evidence that remote learning led to worse outcomes than in-person learning (Goldhaber et al., 2022; Halloran et al., 2021). Second, high-poverty schools were more likely to stay remote for longer (Camp & Zamarro, 2021; Parolin & Lee, 2021). Low-wage workers have significantly lower access to opportunities for telecommuting (Garrote Sanchez et al., 2021). Thus, caregivers in high-poverty households were probably more likely to continue to work outside the home and thus their children may have been less likely to be able to rely on the assistance of an adult when they encountered difficulties with virtual learning. This was potentially more detrimental for younger students who were less able to independently navigate online learning platforms. Simply put, students in high-poverty schools endured more virtual learning while simultaneously being less likely to have access to a caregiver to provide supplemental support and the end result may have been a learning environment that was particularly detrimental for the youngest students.

Implications for practice

Our findings suggest that after the COVID-19 pandemic began, elementary students began the year further behind in their reading development than previous cohorts. Because upper elementary classrooms are typically not resourced to teach early literacy skills, such as constrained decoding skills, school leaders must now consider providing new training, aligned resources, and personnel support to equip upper elementary teachers to provide differentiated, explicit instruction for classes of students who continue to exhibit difficulties in these areas. Additionally, high-quality, intensive intervention services will likely be needed for students entering significantly below their peers. Intervention in reading is crucial to ensure students can continue to access grade-level content in written materials during the instruction of other subject areas (e.g., math, science, history).

We recommend shifting schedules and instructional time to match students' needs as identified, on a broad level, by screening assessments such as NWEA

MAP Growth to determine students who are falling behind grade level expectations. Next, schools can administer further, more specific, diagnostic assessments to students who exhibit difficulties on the initial screener. This will help to determine if individual student difficulties are specifically in decoding skills and/or language comprehension and provide appropriate, targeted intervention, which is critical to begin to improve student reading outcomes and complete unfinished learning that may be due to the pandemic. Teachers may also need additional training and resources to bolster research-based practices, such as implementing and analyzing screening measures, identifying students' needs, grouping students for instruction, strategies for intensifying intervention, implementing high-quality instruction and intervention curricula, and ongoing coaching and support for data-based decision making. School leaders should also consider the addition of class-wide phonics, spelling, handwriting, and fluency intervention for cohorts of students now in the upper grades who may have missed out on valuable in-person learning centered on these constrained skills during their early elementary years.

Limitations

There are several important limitations to our study worth noting. Most importantly, we only included U.S. public schools that tested in fall 2019, fall 2020, and fall 2021. Schools with the resources to consistently test across this three-year span are likely different from schools that did not. Additionally, many U.S. schools stayed remote or hybrid for longer than schools in other countries, so the results presented here may not generalize to international contexts. Second, while we would like to examine the pandemic test score trajectories for students in kindergarten to 2nd grade, test comparability issues with the K-2 assessments prevent us from analyzing MAP Growth K-2 test scores collected during the pandemic (Kuhfeld et al., 2020b). Third, the number of students testing in a given grade changed over the course of the testing periods, which could impact the magnitude of the score drops we observe. If students were more likely to be missing in high-poverty schools and for students who were already low-achieving (e.g., the data are not missing at random), we may be underestimating the differences in the test score patterns in fall 2021. Fourth, we examined differences between schools classified as “low-poverty” and “high-poverty”, but it would be informative to further explore whether there is actually a non-linear relationship between poverty and achievement declines. Additionally, it would be worthwhile to explore how the pandemic differentially affected students who were initially struggling (bottom of the distribution) or high achieving (at the top of the distribution) prior to the pandemic. Finally, we do not have access to information about whether students had access to in-person, hybrid, or remote instruction through the course of the 2020–2021 school year. Without detailed information on the amount and types of instruction received, we are not able to adequately explain the disparities we observed by race/ethnicity and school poverty.

Conclusion

The COVID-19 pandemic represents an unprecedented interruption to students' lives and schooling experiences, so it is perhaps not surprising that large reading declines (0.09 to 0.17 SDs) were observed during this period. However, it is notable that these declines were the largest for elementary students and occurred primarily during the 2020–2021 period rather than directly following the spring 2020 school closures. More research is needed to unpack which students and schools were hardest hit by the COVID-19 pandemic and need additional supports and resources to rebuild and strengthen these key reading skills. Given the critical nature of early elementary school for the development of foundational reading skills, it is essential to continue to monitor reading achievement for young students and match recovery efforts with need. Our data suggest that further resources should be targeted to young students in high-poverty elementary schools to ensure students continue developing these foundational reading skills.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11145-022-10345-8>.

Declarations

Conflict of interest The authors declare no conflicts of interest.

References

- Amplify. (2022). Amid academic recovery in classrooms nationwide, risks remain for youngest students with least instructional time during critical early years. https://amplify.com/wp-content/uploads/2022/02/mCLASS_MOY-Results_February-2022-Report.pdf
- Bethhäuser, B. A., Bach-Mortensen, A., & Engzell, P. (2022). A systematic review and meta-analysis of the impact of the COVID-19 pandemic on learning. <https://doi.org/10.35542/osf.io/d9m4h>
- Camp, A., & Zamarro, G. (2021). Determinants of ethnic differences in school modality choices during the COVID-19 crisis. *Educational Researcher*, 51(1), 6–16.
- Center on Budget and Policy Priorities. (2020). Tracking the COVID-19 recession's effects on food, housing, and employment hardships. <https://www.cbpp.org/research/poverty-and-inequality/tracking-the-covid-19-recessions-effects-on-food-housing-and>
- Centers for Disease Control and Prevention. (2021). *Health equity considerations and racial and ethnic minority groups*. <https://www.cdc.gov/coronavirus/2019-ncov/community/health-equity/race-ethnicity.html>
- Coleman, V., Cookson, Jr., P. W., Garlow, S., Miller, B. K., Norris, S. J., & Reed, J. J. (2021). *Behind the screen: Schooling, stress, and resilience in the Covid-19 crisis*. Stanford Center on Poverty and Inequality, Federal Reserve Bank of Boston, and Federal Reserve Bank of Atlanta. <https://inequality.stanford.edu/sites/default/files/research/articles/covid-online-learning.pdf>
- Curriculum Associates. (2020). Understanding student needs early results from fall assessments. <https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-diagnostic-results-understanding-student-needs-paper-2020.pdf>
- Curriculum Associates. (2021). Academic achievement at the end of the 2020–2021 school year: Insights after more than a year of disrupted teaching and learning. <https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-understanding-student-needs-paper-spring-results-2021.pdf>
- Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020). *COVID-19 and learning loss—Disparities grow and students need help*. McKinsey & Company. <https://www.mckinsey.com/indus>

- tries/public-and-social-sector/our-insights/covid-19-and-learning-loss-disparities-grow-and-students-need-help
- Garrote Sanchez, D., Gomez Parra, N., Ozden, C., Rijkers, B., Violaz, M., & Winkler, H. (2021). Who on Earth can work from home? *The World Bank Research Observer*, 36(1), 67–100. <https://doi.org/10.1093/wbro/lkab002>
- Goldhaber, D., Kane, T., McEachin, A., Morton, E., Patterson, T., & Staiger, D. (2022). The Consequences of remote and hybrid instruction during the pandemic. NBER. https://www.nber.org/system/files/working_papers/w30010/w30010.pdf
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7(1), 6–10.
- Halloran, C., Jack, R., Okun, J., & Oster, E. (2021). Pandemic schooling mode and student test scores: Evidence from US states. NBER Working Paper No. 29497. <https://www.nber.org/papers/w29497>
- Hammerstein, S., König, C., Dreisoerner, T., & Frey, A. (2021). Effects of COVID-19-related school closures on student achievement—A systematic review. *Frontiers in Psychology*, 12, 4020. <https://doi.org/10.3389/fpsyg.2021.746289>
- Hasbrouck, J. & Tindal, G. (2017). *An update to compiled ORF norms* (Technical Report No. 1702). Eugene, OR, Behavioral Research and Teaching, University of Oregon.
- Kuhfeld, M., Lewis, K., Meyer, P., & Tarasawa, B. (2020b). Comparability analysis of remote and in-person MAP Growth testing in fall 2020b. <https://www.nwea.org/content/uploads/2020/11/Technical-brief-Comparability-analysis-of-remote-and-inperson-MAP-Growth-testing-in-fall-2020-NOV2020.pdf>
- Kuhfeld, M., Soland, J., Lewis, K., Ruzek, E., & Johnson, A. (2022). The COVID-19 school year: Learning and recovery across 2020–21. *AERA Open*, 8(1), 1–15.
- Kuhfeld, M., Tarasawa, B., Johnson, A., Ruzek, E., & Lewis, K. (2020a). Learning during COVID-19: Initial findings on students' reading and math achievement and growth. <https://www.nwea.org/content/uploads/2020a/11/Collaborative-brief-Learning-during-COVID-19.NOV2020.pdf>
- Lewis, K., Kuhfeld, M., Ruzek, E., & McEachin, A. (2021). Learning during COVID-19: Reading and math achievement in the 2020–21 school year. <https://www.nwea.org/content/uploads/2021/07/Learning-during-COVID-19-Reading-and-math-achievement-in-the-2020-2021-school-year-research-brief.pdf>
- McGinty, A., Gray, A., Partee, A., Herring, W., & Soland, J. (2021). Examining early literacy skills in the wake of COVID-19 spring 2020 school disruptions. University of Virginia. https://pals.virginia.edu/public/pdfs/login/PALS_Fall_2020_Data_Report_5_18_final.pdf
- NWEA. (2019). MAP® Growth™ technical report. https://www.nwea.org/content/uploads/2021/11/MAP-Growth-Technical-Report-2019_NWEA.pdf
- Nurius, P. S., Prince, D. M., & Rocha, A. (2015). Cumulative disadvantage and youth well-being: A multi-domain examination with life course implications. *Child and Adolescent Social Work Journal*, 32(6), 567–576.
- Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, 40(2), 184–202.
- Parolin, Z., & Lee, E. (2021). Large socio-economic, geographic and demographic disparities exist in exposure to school closures. *Nature Human Behavior*, 5(4), 522–528.
- Renaissance. (2020). How kids are performing: Tracking the impact of COVID-19 on reading and mathematics achievement. <https://www.renaissance.com/how-kids-are-performing/>
- Renaissance. (2021). How kids are performing. <https://www.renaissance.com/how-kids-are-performing/>
- Thorn, W., & Vincent-Lancrin, S. (2021). Schooling during a pandemic: The experience and outcomes of schoolchildren during the first round of COVID-19 lockdowns. *OECD Publishing*. <https://doi.org/10.1787/1c78681e-en>
- Tunmer, W. E., & Hoover, W. A. (2019). The cognitive foundations of learning to read: A framework for preventing and remediating reading difficulties. *Australian Journal of Learning Difficulties*, 24(1), 75–93.
- West, M., & Lake, R. (2021). How much have students missed academically because of the pandemic? A review of the evidence to date. Center on Reinventing Public Education. <https://www.crpe.org/publications/how-much-have-students-missed-academically-because-pandemic-review-evidence-date>

- YouthTruth. (2021). Students weigh in, part II: Learning & well-being during COVID-19. <https://youthtruthsurvey.org/students-weigh-in-part2/>
- Zierer, K. (2021). Effects of pandemic-related school closures on pupils' performance and learning in selected countries: A rapid review. *Education Sciences*, 11(6), 252. <https://doi.org/10.3390/educsci11060252>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.