Millions of Students are (Still) Above Grade Level:

Achievement and Achievement Variability in Mathematics and Reading Before and During COVID-19 in the United States

Karen E. Rambo-Hernandez[1]

Matthew Makel[2]

Noah Koehler[1]

[1] Texas A&M University, [2] University of Calgary

Millions of Students are (Still) Above Grade Level:

Achievement and Achievement Variability in Mathematics and Reading Before and During COVID-19 in the United States

Abstract

According to Peters et al. (2017), millions of students have already demonstrated they know the material slated to be taught that year. Consequently, grade-level standards are unlikely to be appropriately challenging for these students. In this paper, we conceptually replicate and extend this study. Using data from schools that administered the Renaissance Star assessment in fall 2018 and fall 2021, we quantify pre-COVID and mid-COVID average achievement and variance in achievement in mathematics and reading in fifth grade and estimate the grade level of instruction needed for students. Our results indicate that (a) achievement dropped during COVID-19 relative to pre-COVID, but the drop in mathematics was larger, and (b) achievement variability increased during COVID-19, but the variability in reading was slightly more pronounced. Further, our results replicated Peters et al.'s (2017) results—large numbers of students still performed above grade level, and substantial variability in achievement was present within schools.

Millions of Students are (Still) Above Grade Level: Achievement and Achievement Variability in Mathematics and Reading Before and During COVID-19 in the United States

Peters and colleagues (2017) used three state administrative data sets and one national data set to assess the prevalence of students who could demonstrate above-grade-level achievement. They found that many students (20% to 49% in English Language Arts and 14% to 37% in mathematics) scored above grade level. Based on these results, their titular question was whether millions of students should take a gap year. If students can demonstrate they already achieve above grade-level targets, they may not benefit from receiving instruction on material they have already learned. Whether gap years are the appropriate policy response to their results is up for debate. Regardless, the underlying fundamental reality Peters and colleagues assessed was the variability of student achievement within current grade levels. Another recent event (after Peters et al. published) that must be addressed when considering grade-level learning and variability is COVID-19. Numerous studies have shown that COVID-19-related school closures had massive effects on student learning and that these massive effects were not equitably distributed (e.g., Bailey et al., 2021; Kuhfeld & Lewis, 2022; Lewis & Kuhfeld, 2021; Lewis et al., 2022).

With the present study, we sought to conceptually replicate and extend Peters et al.'s (2017) work that estimated the magnitude of this variability in student readiness. Additionally, we sought to examine whether COVID-19-related disruptions were associated with any changes to the variability in achievement within grade levels. In this paper, we first provide a background on student variability. Next, we provide a broad theoretical framework. Finally, before analyzing new data, we review some of the work assessing COVID-19-related schooling closures and their consequential effects on student achievement before analyzing new data.

Variability in Grade-level Achievement

On the first day of kindergarten, students already differ from each other in terms of their learning readiness and needs. For example, Engel and colleagues (2013) analyzed the learning readiness of students entering kindergarten using Early Childhood Longitudinal Study–Kindergarten data. They reported that nearly all (95%) had already mastered basic mathematics concepts like identifying one-digit numbers, recognizing some geometric shapes, and counting to 10. Additionally, 62% demonstrated mastery of patterns and organizing objects by size; 25% understood place values and the value of coins; and 7% could add and subtract single-digit numbers. This variability in mathematical mastery at the beginning of kindergarten illustrates the variability of student learning needs that teachers face when they enter the classroom. That differences exist across individuals is hardly surprising. However, schooling broadly–and classroom instruction specifically– is often designed around grade-level standards (Polikoff, 2021), which do not account for the variability in learning readiness (Ansari, 2018; Engel et al., 2013; von Hippel & Hamrock, 2019).

As Engel et al. (2013) demonstrated, high variability within grade level in mathematics exists from the beginning of kindergarten. West and colleagues (2000) found similar variability in reading skills for kindergarten students more than a decade before Engel et al. (2013). These results suggest that achievement variability is the rule, not the exception.

Importantly, variability in achievement remains throughout schooling. As mentioned above, Peters et al. (2017) reported that many elementary and middle school students scored at least one year above grade level. Using data from the computer adaptive, vertically scaled NWEA MAP® test (NWEA, 2019) taken by students in approximately half of all US school districts (NWEA, n.d.), they estimated that about 10% of fifth-grade students scored at least four grade levels ahead in reading, and about 2.4% did so in mathematics. This suggests that variability in student achievement is both prevalent and substantial. Importantly, 84% of the variance in mathematics scores fell at within schools, meaning that broad academic variability is primarily present within schools and is not a function of different schools serving the needs of different students.

The Peters et al. (2017) paper analyzed variability within grade levels across entire states and the country. On the other hand, Rambo-Hernandez et al. (2021) examined within-classroom mathematics and reading variability. This additional layer is essential because it assesses the extent to which a student may vary from the classroom norm, not just a national norm. They found that in grade six, two-thirds of the achievement variability in mathematics was present in every classroom. The most common number of grade levels within a classroom was five. This means that a typical sixth-grade classroom might have students achieving at the fourth-grade level through the eighth-grade level in mathematics. Pedersen et al. (2023) found similar results with 2015 TIMSS data. 70% of sampled American classrooms included students scoring at four or five of the five possible international benchmarks in mathematics. These data suggest that variability is not simply a function of students in different locations having differing learning needs. It is students sitting in the same classrooms, each with remarkably different learning needs.

Consequences of Student Variability

Broadly speaking, more achievement variability is expected to be associated with greater difficulty for teachers to provide optimal learning opportunities for all students. If most students have similar learning needs, few will need large amounts of differentiation. Knowing the magnitude of within grade-level variability can help inform what is needed to maximize learning for all students in classrooms. For example, if a student knows how to count upon entering kindergarten but spends the first week of class being taught how to count, that instruction will not generate academic learning for that student. Learning occurs when the instruction offered matches the instructional needs demonstrated by a student. Vygotsky's Zone of Proximal Development (ZPD; 1978) asserts that students develop only in appropriately challenging environments; too little challenge leads to boredom, and too much challenge may lead frustration. If students are not provided opportunities for learning aligned with their demonstrated needs, then students are unlikely to maximize their growth.

Indeed, evidence suggests that achievement variability is connected to subsequent learning. As mentioned above, Engel et al. (2013) reported that 95% of entering kindergarteners could count to ten. Nevertheless, teachers in the same study reported spending *13 days per month* teaching this content. They spent over half the instructional days of the month on instruction that almost all their students had already mastered. Importantly, instructional time on this content was negatively associated with end-of-year test scores. Alternatively, the instruction devoted to teaching content few students already knew (e.g., currency, place value) was positively associated with end-of-year achievement. Students who spent more time in environments where they were engaged with material they had *already* learned performed more poorly than those in environments where they were engaged with material they had *yet* to learn, which supports the ZPD perspective.

This ZPD perspective is also supported by research showing that academic acceleration is effective for advanced learners (Steenbergen-Hu et al., 2016). When students who have mastered grade-level material are moved to an instructional context with older students, they showed substantial learning gains compared to same-age peers; the average achievement gain for accelerated students had an effect size of 0.70. In fact, the achievement of accelerated students was not statistically significantly different from their older peers, suggesting that accelerated students performed just as well as them.

Importantly, relative achievement effects on subsequent learning may vary across disciplines. Rambo-Hernandez et al. (2022) found that growth in mathematics was similar across all achievement levels from third through fifth grade. However, this was not the case in reading. In reading, relatively high-achieving students showed slower growth during the school year than their more typically performing peers. This trend was flipped in the summer, with relatively high-achieving students showing faster growth than their more typically achieving peers. This finding demonstrates that not only are there differences in the magnitude of the number of students achieving above grade level across disciplines, but that the effects on subsequent learning on relative high achievement may also differ across disciplines.

If the instruction enacted in classrooms is anchored by what is needed from the typical student in that classroom or school, then a student's achievement relative to local peers is what matters for subsequent learning, not absolute achievement itself. If one student performs typically for their school and the other performs well above their school's average, the former is more likely to have opportunities for growth. As such, between-student relative achievement is an essential factor when exploring which students have the best opportunities for growth. By looking only at state-level data, Peters et al. did not account for between-student relative achievement. We take such relative achievement into account in the present study.

COVID-19, Student Achievement, and Variability in Achievement

COVID-19 and its related school closures led to an upheaval of access to education and learning environments. Before examining its effects on achievement variability, it is important to situate the effects of COVID-19-related educational disruptions on achievement and achievement gaps. In estimating the magnitude of these negative COVID-19-related effects on achievement, an analysis of NWEA MAP® data showed that from Spring 2019 to 2021, median achievement in mathematics dropped by 9-11 percentile-rank points in grades three to five (Lewis & Kuhfeld, 2021). NWEA reported that high-poverty schools showed a percentile rank change of 14-16 points lower in mathematics in 2021 compared to 2019. In reading, median achievement in mathematics dropped by 6-7 percentile-rank points in grades three to five (Lewis & Kuhfeld, 2021), with high poverty schools dropping 10-11 points in percentile rank in 2021 compared to 2019. Similarly, Renaissance Learning, which is the publisher of the computer adaptive, vertically scaled Star assessments (see Renaissance Learning, 2023a, 2023b), reported that the effects of COVID-19 equated to approximately 6-8 weeks of lost instructional time on average in mathematics but only 3 weeks in reading (Renaissance Learning, 2021). Further, Haderlein and colleagues (2021) found that although school services for struggling learners, students receiving special education, and students who were English language learners increased from May to October in 2020, rates of gifted and talented service offerings continued to fall during that time. Prior to the pandemic, 13% of students participated in gifted and talented services. But that dropped to 7% in May 2020 and dropped further to 5% in October 2020. These changes to gifted services may have resulted in drops in overall achievement.

In examining how achievement gaps changed, a report on California students showed that earlier grades were more likely to be negatively affected and that students from low-income households were more likely to be negatively affected by COVID-19 (Pier et al., 2021). Overall, they estimated that students experienced a learning lag of approximately 12% in English/Language Arts between students in low versus high socioeconomic status households. In mathematics, although students from low and high achieving households grew much less than their pre-COVID-19 peers, students from low-achieving households showed the largest drops in learning. Collectively, these varying growth rates suggest there are important moderating variables on the effects of COVID-19 on student achievement, such as socio-economic status, academic discipline (i.e., mathematics versus reading), and grade level.

In contrast to the number of studies of changes in achievement during COVID-19, relatively fewer studies have examined the effects of COVID-19 on achievement variability. One such study was the NAEP Long-Term Assessment (U.S. Department of Education et al., 2022). On the National Assessment of Educational Progress (NAEP), the 90th percentile for nine-year-old students in reading in 2020 (pre-COVID-19) was 267, and in 2022 (mid-COVID-19) the 90th percentile was 265. However, the 10th percentile was much lower— in reading in 2020, the 10th percentile was 164, but in 2022 was 155. In mathematics, the results were similar. The 90th percentile in 2020 was 286 and in 2022 was 283, and the 10th percentile in 2020 was 191 and in 2022 was 178. Results from other countries looked slightly different. In Germany, reading scores narrowed immediately following the onset of the pandemic, and this narrowing was primarily due to a drop in achievement from the highest-performing students (Schult et al., 2022). In the same study, student mathematical operations assessment scores widened, i.e., the 95th percentile increased while the 5th percentile fell after the onset of the pandemic. Thus, COVID-19-related educational disruptions appeared to impact variability in student achievement. However, within and across countries, the results are inconsistent regarding how mathematics and reading variability were impacted and whether one was affected more than the other, which would be helpful information when tailoring COVID-19 recovery efforts.

Present Study

In the present study, we sought to conceptually replicate the Peters et al. (2017) study that quantified how many students scored above grade level in mathematics or reading using state administrative test data and the NWEA MAP® test before the onset of educational disruptions due to COVID-19. To test the generalizability of prior findings, we purposefully altered two aspects of the prior study: (a) the years assessed (fall 2018 and fall 2021) and (b) the assessment used (Renaissance Star data). The years assessed were altered to assess (a) whether similar results would be found across time and (b) whether COVID-19-related school disruptions were associated with any changes to grade-level achievement variability. The assessment used was altered to assess the generalizability of achievement variability across different assessments. Such changes and generalizability checks are standard components of conceptual replications (Makel & Plucker, 2015).

Prior to analyzing the data, we preregistered predictions (https://aspredicted.org/JQ3_3TF). To set the stage for examining the variability in mathematics and reading achievement for the pre-COVID-19 and mid-COVID-19 cohorts, we first examined the differences in achievement for the pre-COVID-19 cohort and the mid-COVID-19 cohort. Next, we examined changes in variability between the pre-COVID-19 cohort and the mid-COVID-19 cohort and whether variability differed by school poverty status, operationalized by the US Title 1 designation which indicates a high poverty school. Finally, we examined whether mathematics or reading had more variability in student scores as a result of COVID-19-related educational disruptions. Specifically, we hypothesized:

Baseline Achievement Hypothesis

1. The pre-COVID-19 cohort would have higher achievement than the mid-COVID-19 Cohort.

Change in Variability Hypotheses

2. The variability of achievement would be larger for the mid-COVID-19 cohort than the pre-COVID-19 cohort.
3. The ratio of the variability in achievement of the mid-COVID-19 cohort to the pre-COVID-19 cohort would be related to Title 1 status and average school achievement, and that average school achievement would moderate the relationship between Title 1 status and the variability ratio.

Academic Discipline and Achievement Changes in Variability Hypotheses

4. The mid-COVID-19 cohorts would have a larger drop in achievement and a greater increase in variability in mathematics than reading.
5. For grade-level variability:
    - A larger percentage of students would be below grade level in mathematics and reading in the mid-COVID-19–19 cohort compared to the pre-COVID-19 cohort;
    - a smaller percentage of students would be above grade level in mathematics in the mid-COVID-19 cohort compared to the pre-COVID-19 cohort;
    - a larger percentage of students would be above grade level in reading in the mid-COVID-19 cohorts compared to the pre-COVID-19 cohort.

Method

Participants

We observed fifth-grade student scores in mathematics and reading from those who participated in the Renaissance Star test in fall 2018 and fall 2021. We first paired and retained schools in the dataset that were observed in fall 2018 (pre-COVID cohort) and fall 2021 (mid-COVID-19 cohort). Next, we removed schools that did not include at least ten students at each observation. The mathematics sample contained 4,815 schools, with 294,831 students in the pre-COVID cohort and 258,049 students in the mid-COVID-19 cohort. The reading sample contained 5,214 schools, with 318,146 students in the pre-COVID-19 cohort and 280,943 students in the mid-COVID-19 cohort.

Measures

The Renaissance Star assessment is a vertically scaled, computer-adaptive assessment. The reading assessment is designed to assess reading comprehension, and the mathematics assessment covers numbers and operations, geometry and measurement, algebra and data analysis, statistics, and probability (Renaissance Learning, 2022). The Star assessment was designed to be an interim assessment and is typically given up to three times per academic year (fall, winter, and spring). Because of the nature of the assessment, scores are not as susceptible to floor or ceiling effects often found in state-standardized yearly assessments. For the norming sample at the fifth grade level, the alternate form ($\rho_m = .85$, $\rho_r = .86$), split-half ($\rho_m = .92$, $\rho_r = .94$), and generic reliability ($\rho_m = .92$, $\rho_r = .94$) were high for mathematics and reading (Renaissance Learning, 2023a; Renaissance Learning, 2023b). It is worth noting that this assessment is similar in scope and application to one of the assessments used in the Peters et al. (2017) study, namely the NWEA MAP®.

Of importance to users of the Star test is its validity, especially as it applies the assessments' predictive ability to accurately predict performance on end-of-year state assessments, such as the Smarter Balanced Assessment (see https://smarterbalanced.org/). To this end, Renaissance conducted a concordance study to approximate which Star scores are aligned with Smarter Balanced performance levels, namely Levels 1-4 (Renaissance Learning, 2022). Smarter Balanced reports that students at levels 3 and 4 are considered on-track (Smarter Balanced, 2022). For the fifth-grade students in the concordance study, concurrent Star scores were correlated at .88 and .85 with Smarter Balanced scores for mathematics and reading, respectively. Further, prior studies have demonstrated that the classification accuracy of Star scores for predicting Smarter Balanced assessments was strong, 88% for mathematics and 85% for reading (Renaissance Learning, 2023a; Renaissance Learning, 2023b).

Plan of Analysis

Baseline Achievement Hypothesis

Hypothesis 1. We conducted paired sample *t*-tests of the school means of achievement observed mid-COVID achievement to pre-COVID-19 separately for mathematics and reading.

Change in Variability Hypotheses

Hypothesis 2. We tested separately whether the variances of mathematics and reading scores in schools during COVID-19 were greater than those observed before COVID-19 by calculating the ratio of mid-COVID-19 variance to pre-COVID-19 variance for each school. Although we originally planned only to bootstrap the results to account for the anticipated skewed distribution, the ratios were so skewed that we transformed each ratio by calculating the natural log of each ratio. Thus, a school that had the same variance in mid-COVID-19 as it did in pre-COVID-19 would have a variance ratio of 1, which would then be equal to 0 in the natural log scale. Next, we compared whether the average natural logs for mathematics and reading were different from 0 using a one-sample *t*-test. We bootstrapped the standard errors and reported the bias-corrected 95% confidence intervals as planned.

Hypothesis 3. Separately for mathematics and reading, we regressed the natural log of the ratio of mid-COVID-19 variance to pre-COVID-19 variance for each school (created in Hypothesis 2) on school Title 1 status, school average achievement prior to COVID-19 disruptions, and their interaction. The predictor variables included: (a) Title 1 status, which was coded 0.5 for Title 1 and -0.5 for non-Title 1 schools, (b) achievement at each school, which was grand mean centered, and (c) the interaction of the two. We bootstrapped the standard errors and reported the bias-corrected confidence intervals.

Academic Discipline and Achievement Changes in Variability

Hypothesis 4. For schools that gave both mathematics and reading assessments before and during COVID-19, we calculated which academic discipline— mathematics or reading— had a greater drop in achievement and increase in variability. To test which academic discipline had the greater drop, we created a *z*-score for achievement for each mid-COVID-19 school relative to their school's pre-COVID-19 mean and standard deviation. We created separate z-scores for mathematics and reading. Then, we conducted a paired sample *t*-test of school *z*-scores for mathematics and the school *z*-scores for reading for the mid-COVID schools. To test whether the variances were different, we used a one-sample *t*-test to compare the difference of the ratio of two natural logs created in hypothesis 2 for mathematics to reading. Again, if the difference was greater than 0, then mathematics had a greater increase in variance, and vice-versa if the difference was less than 0.

Hypothesis 5. First, to determine the proportion of students below and above grade level, we used the mathematics and reading Star scores corresponding to met expectations (also referred to as on-track) in their respective grade levels in Smarter Balanced Assessment (Renaissance Learning, 2022). For example, if students scored at or above the fourth-grade met expectations score (e.g., $\geq$ 1022 for reading) but below the fifth-grade met expectations score (e.g., < 1044 for reading), they were assumed to need fifth-grade instruction. Those below the fourth-grade met expectations score were deemed to need below fifth grade instruction, and those above the fifth-grade met expectations were expected to need above fifth grade instruction. We repeated this for the students in fall 2018 and fall 2021. To determine if the proportions changed, we compared the proportions using a paired sample *t*-test for below and above grade level for mid-COVID to pre-COVID in schools. If the differences differed from 0, then the proportion of students above (or below, depending on the comparison) was different for the mid-COVID schools compared to the pre-COVID schools.

Finally, we further disaggregated the proportions of students below and above grade level by assigning student performance to the grade level they were estimated to need instruction. A student was expected to need third-grade instruction (or lower) if their score was below the third-grade level met expectations minimum score (e.g., 993 for reading). Students were estimated to need fourth-grade instruction if their score fell between the third-grade minimum score (e.g., $\leq$993 for reading) and below the fourth-grade minimum score (e.g., 1022 for reading). For above-grade-level students, students were binned on whether their score was above the fifth-grade level met expectations threshold but not the sixth grade (needs sixth-grade instruction), was above the sixth-grade met expectations threshold but not the seventh grade (needs seventh-grade instruction), or was above the seventh-grade met expectations threshold (needs eighth grade or higher instruction). See Table 1 for a breakdown of the thresholds. For example, a student who scored 1047 on the reading assessment was said to have met fifth-grade expectations and be estimated to need instruction at the sixth-grade level (e.g., 1047 is above the fifth-grade reading score for met expectations [1044] but below the sixth-grade level met expectations [1068]). *It is important to note that we used this approach* ***not*** *to recommend that these students move grade levels to be aligned with their scores but rather to estimate the range of instructional needs in fifth grade in schools. Given that each student's score contains some error, a single score should not be used to determine services for any one student. However, in the aggregate, these estimates may help elucidate the range of academic needs of students in a single grade level.*

Results

Descriptive Statistics

First, even though these are the same schools observed before and during COVID-19 educational disruptions, we want to note that the number of students who tested in the mid-COVID-19 cohort was smaller than the number of students who tested prior to COVID-19. To describe the range of achievement prior to and during the COVID-19 educational disruptions, we first calculated the average number of students, school means, school variances, school standard deviations, and the range of achievement for mathematics and reading for each school twice— fifth-grade fall 2018 (pre-COVID-19) and fall 2021 (mid-COVID-19). The mid-COVID-19 average school scores were approximately 13 points lower in mathematics and 5 points lower in reading than their pre-COVID-19 counterparts. Further, the average standard deviation within schools in the mid-COVID-19 cohorts also appeared to increase slightly— in mathematics, about 1.5 units from pre-COVID-19 to mid-COVID-19, and in reading, about 4 units from pre-COVID-19 to mid-COVID-19 (see Table 2 and Figure 1).

Further, we examined the descriptive statistics by cohort and Title 1 status (see Table 1A in the supplemental materials). Across both mathematics and reading, Title 1 schools had lower average school scores than non-Title 1 schools observed at the same time point. Further, the drop in scores between the pre-COVID-19 and the mid-COVID-19 cohorts was, at least descriptively, greater for Title 1 schools than non-Title 1 schools. Additionally, across mathematics and reading, the average standard deviation within schools appeared to be greater during COVID-19 for both Title 1 and non-Title 1 schools compared to pre-COVID-19, but Title 1 schools appeared to have a larger increase in their standard deviation relative to pre-COVID-19. Notably, Title 1 schools' reading standard deviation increased by nearly 6 units, while non-Title 1 schools' standard deviation only increased by approximately 1.5 units.

Next, we examined how the average standard deviation changed as a function of the school's achievement quintile and COVID-19 cohort status. In mathematics across COVID-19 cohorts, as average achievement at the school increased, the average standard deviation decreased—indicating that higher achieving schools tended to be slightly more homogeneous (see Figure 2, left panel; Table 2A in the supplemental materials). There was little difference between the average standard deviations of the COVID-19 cohorts for the two lower quintiles. However, descriptively, quintiles 3, 4, and 5 had slightly larger standard deviations in the mid-COVID-19 cohort relative to the pre-COVID-19 cohort (see Table 2A in the supplemental materials).

In reading, regardless of COVID-19 cohort and like mathematics, as school average achievement increased, the standard deviation of achievement at the school decreased— indicating that higher achievement schools were more homogeneous than lower achieving schools (see Figure 2, right panel).

Unlike mathematics, the standard deviation in reading was consistently larger in the mid-COVID-19 cohorts relative to their pre-COVID-19 counterparts, and the difference was most pronounced in lower-achieving schools (see Table 2A in the supplemental materials).

Baseline Achievement

Hypothesis 1

As predicted, the school mathematics average scores for the mid-COVID-19 schools were statistically significantly lower than the same schools observed pre-COVID-19 ($mean_{diff}$= -13.29, t(4814)= -42.42, 95% CI [-13.91, -12.68], Hedge's g = -.41). Similarly, the average school reading scores for the mid-COVID-19 schools were also statistically significantly lower than the same schools observed pre-COVID-19 ($mean_{diff}$= -5.12, t(5213)= -17.00, 95% CI [-5.71, -4.53], Hedge's g = -.14).

Change in Variability

Hypothesis 2

Next, we examined the ratio of the variance in scores within schools during COVID-19 to the variance in the same schools before COVID-19 educational disruptions. We took the natural log of each ratio and compared whether the average natural log was different from 0.

In mathematics, the bootstrapped result of the natural log of the ratio of the mid-COVID-19 to pre-COVID-19 variance was 0.066 (95% bias-corrected CI [0.052. 0.078]). This result supported the hypothesis that the variance was greater in school mathematics scores during COVID-19 compared to when they were observed prior to COVID-19. In short, this translates to an average increase in the variance by a factor of 1.07 ($e^{0.066} = 1.07$) in mid-COVID-19 schools in mathematics achievement compared to pre-COVID-19 schools, which translates to an increase of approximately 1.03 in standard deviation units.

In reading, bootstrapped results indicated that the natural log of the ratio of the mid-COVID-19 to pre-COVID-19 variance was 0.10 (95% bias-corrected CI [0.09, 0.12]), which indicates that the variance in the mid-COVID-19 schools was greater than in pre-COVID-19 schools. Further, this translates to an average increase in the variance by a factor of 1.11 ($e^{0.10} = 1.11$) in mid-COVID-19 schools in reading achievement compared to pre-COVID-19 schools. which translates to an increase of approximately 1.05 in standard deviation units.

Hypothesis 3

Next, we predicted that variability in school achievement would be related to Title 1 status and average school achievement (assessed prior to COVID-19), and that the relationship between Title 1 status and variability would be moderated by school achievement. Using the natural log-transformed ratio calculated for Hypothesis 2 as the dependent variable, we regressed the mathematics and reading transformed ratios separately (Table 3).

In mathematics, Title 1 status and prior school achievement were statistically significant predictors, but the moderation was not. Regardless of Title 1 status, as prior school achievement increased, the natural log of the variance ratio was expected to increase (schools with prior achievement 1 SD above average: $e^{.06+.003(55.14)} = 1.25$, schools with prior achievement 1 SD below average: $e^{.06+.003(-55.14)} = 0.90$). Thus, schools with above-average achievement saw increases in variance during COVID-19, while schools with below-average achievement saw shrinkage in variance during COVID-19. Similarly, controlling for prior achievement, Title 1 schools demonstrated a greater increase in variance than non-Title 1 schools (odds ratio for Title 1 schools: $e^{.06+.08(.5)} = 1.11$, odds ratio for non-Title 1 schools: $e^{.06+.08(-.5)} = 1.02$). However, the total variability explained by the model was small ($R^2 = .025$); thus, these differences may not be practically meaningful.

In reading, only Title 1 status was uniquely predictive of the ratio of variance during COVID-19 to pre-COVID-19. Title 1 schools had a greater increase in variance than non-Title 1 schools during

COVID-19 educational disruption (odds ratio for Title 1 schools: $e^{.10+.12(.5)} = 1.17$, odds ratio for non-Title 1 schools: $e^{.10+.12(-.5)} = 1.10$). However, the R^2 for the model was only slightly more than 1%; thus, this may not be a meaningful difference between school types.

Academic Discipline and Achievement Changes in Variability

Hypothesis 4

Next, we hypothesized that the mid-COVID-19 cohorts would have a larger drop in achievement and a greater increase in variability in mathematics than reading. There were 4,750 schools with both mathematics and reading scores. As hypothesized, the average mathematics *z*-score (z_{mean}= -0.256, *SD*= 0.42) was statistically significantly lower than the average reading z-score (z_{mean}= -0.098, *SD*= 0.35; $diff_{mean}$ = -0.158, *SE*= 0.005, 95% *CI* [-0.167, -0.148]).

Next, we examined the change in variance during COVID-19 relative to pre-COVID-19 in the same schools to see if mathematics had a greater increase in variability compared to reading. Again, we used the natural log transformation of each ratio and a paired sample *t*-test to compare the natural log of the mathematics ratio of variance to the reading variance ratio. The results were different from what we hypothesized. The variance in mathematics during COVID-19 relative to pre-COVID-19 (nl(Mathematics ratio)= 0.06, SE= .007, 95% CI [0.05, 0.08]) was smaller than the variance in reading during COVID-19 relative to pre-COVID (ln(reading ratio) = 0.10, SE= .52, 95% CI [.09, -.11], difference = –0.04, 95% CI[-0.056, -0.26], *t* (4749) = -5.36).

Hypothesis 5

Finally, we examined the estimated grade levels of instructional need for students and schools both pre- and mid-COVID-19 (Table 4). The left side of Table 4 summarizes the proportion of all *students* in each cohort who were at each grade level. The right side summarizes the proportion of students in each *school* that were at each grade level. Because the two sides report similar proportions, this provides evidence that typical schools have a wide range of academic needs at the fifth-grade level. For schools observed mid-COVID-19 relative to pre-COVID-19, the proportion of students below grade level— both at the student level and at the average proportion within schools— increased. In contrast, the proportion of students above grade level decreased during COVID-19 relative to pre-COVID-19 rates. Of note, the grade level shifts were most evident in the proportion of students who were not at grade level— both mathematics and reading results indicated nearly 10% more students were below grade level in the mid-COVID-19 cohort relative to the pre-COVID cohort. Additionally, the proportion of students estimated three or more grade levels ahead was slightly smaller but non-negligible in mid-COVID (mathematics pre: 2.84%, mid: 2.22%; reading pre: 15.47%, mid: 14.68%).

The proportion of students who needed below, at, and above fifth-grade instruction in mathematics and reading is illustrated in Figure 3. The proportion of students on- and above grade level decreased in mathematics for the mid-COVID-19 cohort relative to the pre-COVID-19 cohort; however, the proportions below, at, and above grade level were much more stable in reading across the pre- and mid-COVID cohorts. Thus, reading changes in the proportion of students below, at, or above grade were much smaller than the swings observed in mathematics. Further, the variability between schools in their average mathematics scores appeared to have increased for the schools when they were observed mid-COVID-19 relative to pre-COVID (ICC pre = .21, ICC mid= .24), but the variability between schools remained stable in reading (ICC pre = .20, ICC mid= .20).

In mathematics, both the proportion of all students and the average proportion in schools of students who needed below grade level instruction was large before COVID-19 and was even larger for the mid-COVID-19 cohort. The average proportion in schools of students that were below grade level increased for schools observed during mid-COVID-19 by approximately 7% ($mean_{diff}$ = .07, $t(4814)$= 36.43, 95% CI [.07, .08], Hedge's g = .39). On the other hand, the average proportion of students who were above level decreased for schools observed during mid-COVID-19 by 2% ($mean_{diff}$ = -.02, $t(4814)$= -22.54, 95% CI [-.029, -.024], Hedge's g = -.27).

Similarly, in reading, the proportion of all students who were below grade level and the average proportion in schools of those below grade level was larger during mid-COVID-19 than pre-COVID-19. The average proportion in schools of students who needed below fifth-grade level instruction increased for schools observed during mid-COVID-19 by 1.5% ($mean_{diff}$ = .015, $t(5213)$= 8.11, 95% CI [.011, .08], Hedge's g = .07). On the other hand, the average proportion of students who were above level in schools decreased schools observed during mid-COVID-19 by 1.2% ($mean_{diff}$ = -.012, $t(5213)$= -22.54, 95% CI [-.015, -.001], Hedge's g = -.06).

Discussion

In this study, we sought to conceptually replicate and expand upon Peters and colleagues' (2017) findings that reported how many students scored above grade level in reading and mathematics. As part of the conceptual replication, we purposefully altered the assessment to make sure that results extended to other assessments. To extend prior results, we also analyzed the proportion of students performing below grade level and assessed whether achievement changed across two different cohorts occurring before and in the midst of COVID-19-related educational disruptions. Several results are worth noting. First, our results successfully replicated the main results reported by Peters et al. (2017); at the beginning of the school year, many students performed above grade level, and this substantial variability in achievement was present within schools.

Second, as we predicted, average achievement in both mathematics and reading was lower for the mid-COVID-19 cohort than the pre-COVID-19 cohort (g = -.41 and -.14, respectively). This decrease in average achievement aligns with several prior studies (e.g., Bailey et al., 2021; Lewis & Kuhfeld, 2021; Lewis et al., 2022). Similarly, as predicted, the overall variance in achievement within a grade level increased in both mathematics and reading (1.07 and 1.11, respectively). These increases in variability demonstrate that the spread of achievement within a grade level increased despite the decrease in average achievement. This increase means that teachers had students with learning needs that were more variable during the COVID-19 pandemic than they had prior to the pandemic.

Third, higher levels of poverty were weakly associated with changes in achievement variability within schools. Prior research relying on Star data and research with students in California reported that students who were socioeconomically disadvantaged showed greater learning loss than their non-economically disadvantaged peers (Pier et al., 2021), but this study did not examine if variability in achievement was impacted. In our study, high poverty status (i.e., Title 1) was statistically significantly associated with changes in variance across schools but was small and thus likely did not uncover any meaningful relationships in mathematics or reading. This lack of strong associations seems to illustrate that while COVID-19 negative impacts on achievement appeared to be stronger for Title 1 schools than non-Title 1 schools, the increased variability in achievement was consistent across all schools—regardless of Title 1 status.

Fourth, following the onset of COVID-19, the drop in mathematics was larger than the drop in reading. However, in contrast to our prediction, the change in variance during COVID-19 relative to pre-COVID-19 was larger in reading than in mathematics. The smaller change in variability in mathematics may have been due to the more teacher-dependent nature of learning mathematics relative to reading (Alexander et al., 2001; Cooper et al., 1996). For example, once students have learned the building blocks of reading, extending these skills to more complex text may be more independently accessible to students. While in mathematics, students may depend more on their teachers to direct them to the next, more complex set of skills to master. Because teacher access was restricted to some degree during COVID-19, mathematics achievement may have dropped more but also varied less than reading achievement. Further, this drop in mathematics achievement and students who performed above level could be due to fewer students being served through gifted and talented programs during COVID-19 (Haderlein et al., 2021).

Finally, replicating and extending the Peters et al. (2017) analysis that reported the percentage of students performing above grade level, we found that more students performed off grade level than on grade level. In fact, for the mid-COVID-19 cohort, fewer than one in four students performed at grade level in mathematics and fewer than one in seven did so in reading. Instead, the majority of students were below grade level for both cohorts in both disciplines. At the same time, similar to Peters et al. (2017), a large proportion of students performed above grade level, namely 9% in mathematics and 35% in reading in the mid-COVID-19 cohorts, and sometimes multiple years above grade level.

This replication of prior findings further demonstrates that variability in student achievement is the norm— not the exception. The magnitude of the number of students this variability represents is incredibly large. If our results generalize to the approximate 3,550,000 fifth-grade students that were in schools in 2021 (National Center for Educational Statistics, 2022) across the United States (an inference that we cannot directly support), this would imply that over 300,000 fifth-grade students in mathematics and over 1.2 million students in reading can demonstrate above-grade-level performance. Further, the number of students performing several years above grade level is also quite large. If our results were generalized to the entire U.S. fifth-grade population, the number of students performing at least three years above grade level would be nearly 80,000 in mathematics and over half a million in reading. This estimate is not a trivial number of students. If they are performing above grade level, the essential educational follow-up question concerns whether they are receiving instruction that matches their demonstrated achievement levels.

Importantly, our results showed there is more variability in student achievement within schools than between schools. For example, the ICCs ranged from .20 to .24. *Thus, approximately 20% of the variability in achievement is due to differences between schools while approximately 80% of differences in achievement are due to variability between students within schools*. Thus, the large variability in student achievement is *not* due to students within schools being of similar ability, while the average achievement between schools varied dramatically. The large variability in achievement is mostly due to students with wide-ranging academic needs being served in any given school. Increased variability expands the amount of juggling instructors are required to perform while teaching, thus increasing the number of students whose learning needs may not be met in the typical classroom.

Limitations and Future Research

This study had limitations that we want to note. First, the differences between the cohorts unrelated to COVID-19 could be driving the results (e.g., demographics and access to resources). For example, although we used data from the same schools in the pre- and mid-COVID-19 cohorts, the demographics may have shifted over time as not all families sent their children back to school and other families may have relocated after the onset of the COVID-19 pandemic. Next, cohort differences could be related to who participated in testing during COVID-19 and who did not— more than 10% fewer students tested in mid-COVID-19 compared to the pre-COVID-19 cohort. Finally, determining academic needs based on one assessment is not precise and subject to error. We binned students into estimated grade levels, which should be viewed cautiously. Like all scores, these scores contain some margin of error, and the grade level estimated was meant to be illustrative, not deterministic of academic need. Future research is needed to assess whether the trends observed here generalize to other grade levels and the extent to which the large proportions of students performing outside their grade level, and the associated large variability, affects their subsequent growth. Specifically, how are variability in classrooms/schools and performing above-grade level associated with future learning?

Conclusion

Similar to Peters et al. (2017), our results suggest that students performing at a grade level that differs from their assigned grade level is not rare. It is normative. In fact, with our addition of assessing students performing below grade level, our results suggest that students performing at the grade level in which they are assigned is a rare event. Further, this range of student achievement levels is not due to extreme achievement divides across different schools; there is wide variability in student achievement within schools. Large numbers of students showing wide variability in achievement across multiple domains within schools suggest that substantial modification may be needed from grade-level instruction to maximize growth for all students, especially as schools address learning loss due to COVID-19 educational disruptions.

References

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2001). Schools, achievement, and inequality: A seasonal perspective. Educational Evaluation and Policy Analysis, 23, 171–191. https://doi.org/10.3102/01623737023002171

Ansari, A. (2018). The persistence of preschool effects from early childhood through adolescence. *Journal of Educational Psychology, 110*(7), 952. https://doi.apa.org/doi/10.1037/edu0000255

Bailey, D. H., Duncan, G. J., Murnane, R. J., Yeung, N. A. (2021). Achievement gaps in the wake of COVID-19. *Educational Researcher, 50*(5), 266–275. https://doi.org/10.3102/0013189X211011237

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, *66*(3), 227-268. https://doi.org/10.3102/00346543066003227

Engel, M., Claessens, A., & Finch, M. A. (2013). Teaching students what they already know? The (mis)alignment between mathematics instructional content and student knowledge in kindergarten. *Educational Evaluation and Policy Analysis, 35*(2), 157–178. https://doi.org/10.3102/0162373712461850

Haderlein, S. K., Saavedra, A. R., Polikoff, M. S., Silver, D., Rapaport, A., & Garland, M. (2021). Disparities in educational access in the time of COVID: Evidence from a nationally representative panel of American families. *AERA Open, 7*(1), 1-21. https://doi.org/10.1177/23328584211041350

Kuhfeld, M. & Lewis, K. (2022). *COVID-19 in the early elementary years: A comparison of achievement in spring 2019 and spring 2022*. https://www.nwea.org/research/publication/covid-19-in-the-early-elementary-years-a-comparison-of-achievement-in-spring-2019-and-spring-2022/

Lewis, K. & Kuhfeld, M. (2021). *Learning during COVID-19: An update on student achievement and growth at the start of the 2021-2022 school year*. https://www.nwea.org/research/publication/learning-during-covid-19-an-update-on-student-achievement-and-growth-at-the-start-of-the-2021-22-school-year/

Lewis, K., Kuhfeld, M., Langi, M., Peters, S., & Fahle, E. (2022). *The widening achievement divide during COVID-19*. https://www.nwea.org/uploads/2022/11/CSSP-Brief_Widening-achievement-divide-COVID-19.pdf

Makel, M. C. & Plucker, J. A. (2015). An introduction to replication research in gifted education: Shiny and new is not the same as useful. *Gifted Child Quarterly, 57*(3), 157–164. https://doi.org/10.1177/0016986215578747

National Center for Educational Statistics (2022). Digest of Education Statistics: Enrollment in public elementary and secondary schools, by level and grade: Selected years, fall 1980 through fall 2031. https://nces.ed.gov/programs/digest/d22/tables/dt22_203.10.asp?current=yes

NWEA. (2019). MAP® Growth™ technical report. https://www.nwea.org/research/publication/map-growth-technical-report/ .

NWEA. (n.d.). Case Studies. https://www.nwea.org/case-studies/

Pedersen, B., Makel, M. C., Rambo-Hernandez, K. E., Peters, S. J., Plucker, J. A. (2023). Most mathematics classrooms contain wide-ranging achievement levels. *Gifted Child Quarterly, 67*(3), 220-234. https://doi.org/10.1177/00169862231166074

Pier, L., Hough, H. J., Christian, M., Bookman, N., Wilkenfeld, B., & Miller, R. (2021, January 25). *COVID-19 and the educational equity crisis: Evidence on learning loss from the CORE Data Collaborative* [Commentary]. Policy Analysis for California Education. https://edpolicyinca.org/newsroom/covid-19-and-educational-equity-crisis

Peters, S. J., Matthews, M. T., Rambo-Hernandez, K. E., Makel, M. C., & Plucker, J. A. (2017). Should millions of students take a gap year? Large numbers of students start the school year above grade level. *Gifted Child Quarterly, 61*(3), 229–238. https://doi.org/10.1177/0016986217701834

Polikoff, M. (2021). *Beyond Standards: The Fragmentation of Education Governance and the Promise of Curriculum Reform.* Harvard Education Press.

Rambo-Hernandez, K. E., Makel. M. C., Peters, S., J., & Worley, C. (2022). Differential return on investment: Academic growth based on initial achievement. *British Journal of Educational Psychology, 92*(3), 817–842. https://doi.org/10.1111/bjep.12479

Rambo-Hernandez, K. E., Peters, S. J., Plucker, J. A., Makel, M. C., Pedersen, B. A. (April, 2021). *How Academically Diverse Is the "Grade-Level" Classroom?* Paper presented at the 2021 American Educational Research Association Annual Conference: Virtually.

Renaissance Learning (Winter, 2021). *How kids are performing: Tracking the midyear impact of COVID-19 on reading and mathematics achievement*. https://www.renaissance.com/how-kids-are-performing-winter-2020/

Renaissance Learning (2022). Relating Star Reading and Star Math to California Assessment of Student Performance and Progress (CAASPP) (Smarter Balanced). Renaissance Learning. https://renaissance.widen.net/view/pdf/sm7wh7p2k8/R44908.pdf?t.download=true&u=zceria

Renaissance Learning (2023a). Star Assessments™ for Math Technical Manual. Renaissance Learning. https://star-help.renaissance.com/hc/en-us/articles/11583043402267-Star-Assessments-for-Math-Technical-Manual

Renaissance Learning (2023b). Star Assessments™ for Reading Technical Manual. Renaissance Learning. https://star-help.renaissance.com/hc/en-us/articles/12542471051803-Star-Assessments-for-Reading-Technical-Manual

Schult, J., Mahler, N., Fauth, B., & Lindner, M. A. (2022). Did students learn less during the COVID-19 pandemic? Reading and mathematics competencies before and after the first pandemic wave. *School effectiveness and school Improvement, 33*(4), 544-563. https://doi.org/10.1080/09243453.2022.2061014

Smarter Balanced (2022). Smarter Balanced Validity Research. Retrieved February 28, 2023, from https://validity.smarterbalanced.org/scoring/

Steenbergen-Hu, S., Makel, M. C., Olszewski-Kubilius, P. (2016). What one hundred years of research says about the effects of ability grouping and acceleration on K-12 students' academic achievement: Findings from two second-order meta-analyses. *Review of Educational Research, 86*(4), 849-899. https://doi.org/10.3102/0034654316675417

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP). (2022). *1971-2022 Long-Term Trend (LTT) Reading and Mathematics Assessments*. https://www.nationsreportcard.gov/highlights/ltt/2022/

von Hippel, P. T., & Hamrock, C. (2019). Do test score gaps grow before, during, or between the school years? Measurement artifacts and what we can know in spite of them. *Sociological Science, 6*, 43. http://dx.doi.org/10.15195/v6.a3

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman., Eds.; A. R. Luria, M. Lopez-Morillas, & M. Cole [with J. V. Wertsch], Trans.). Harvard University Press. (Original manuscripts [ca. 1930-1934])

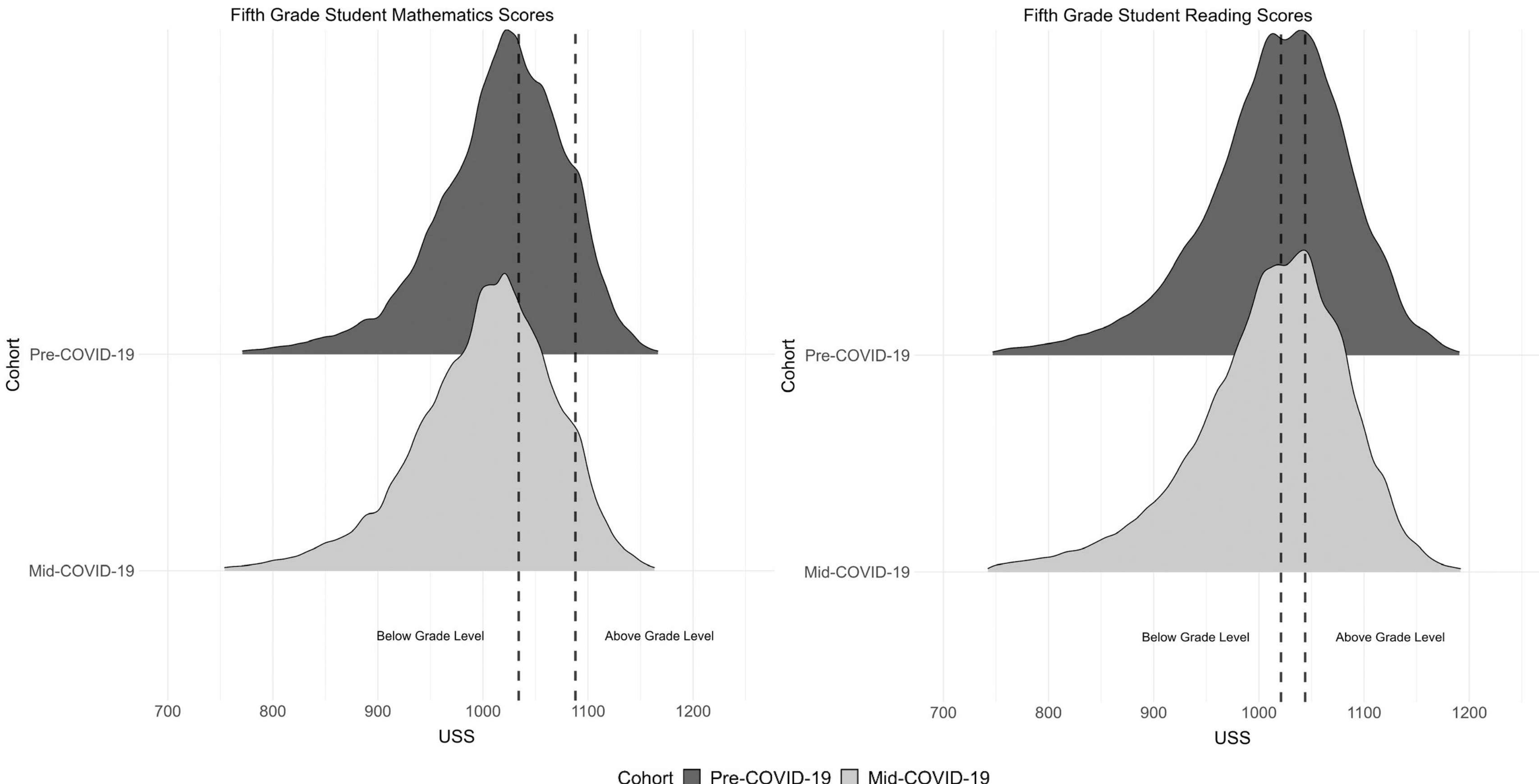
West, J., Flanagan, K. D., & Germino-Hausken, E. (2000). *America's kindergartners: Findings from the Early Childhood Longitudinal Study, kindergarten class of 1998-99, fall 1998. Jerry West, Kristin Denton, Elvie Germino-Hausken.* U.S. Dept. of Education, Office of Educational

Research and Improvement, National Center for Education Statistics.

https://nces.ed.gov/pubs2000/2000070.pdf

Figure 1

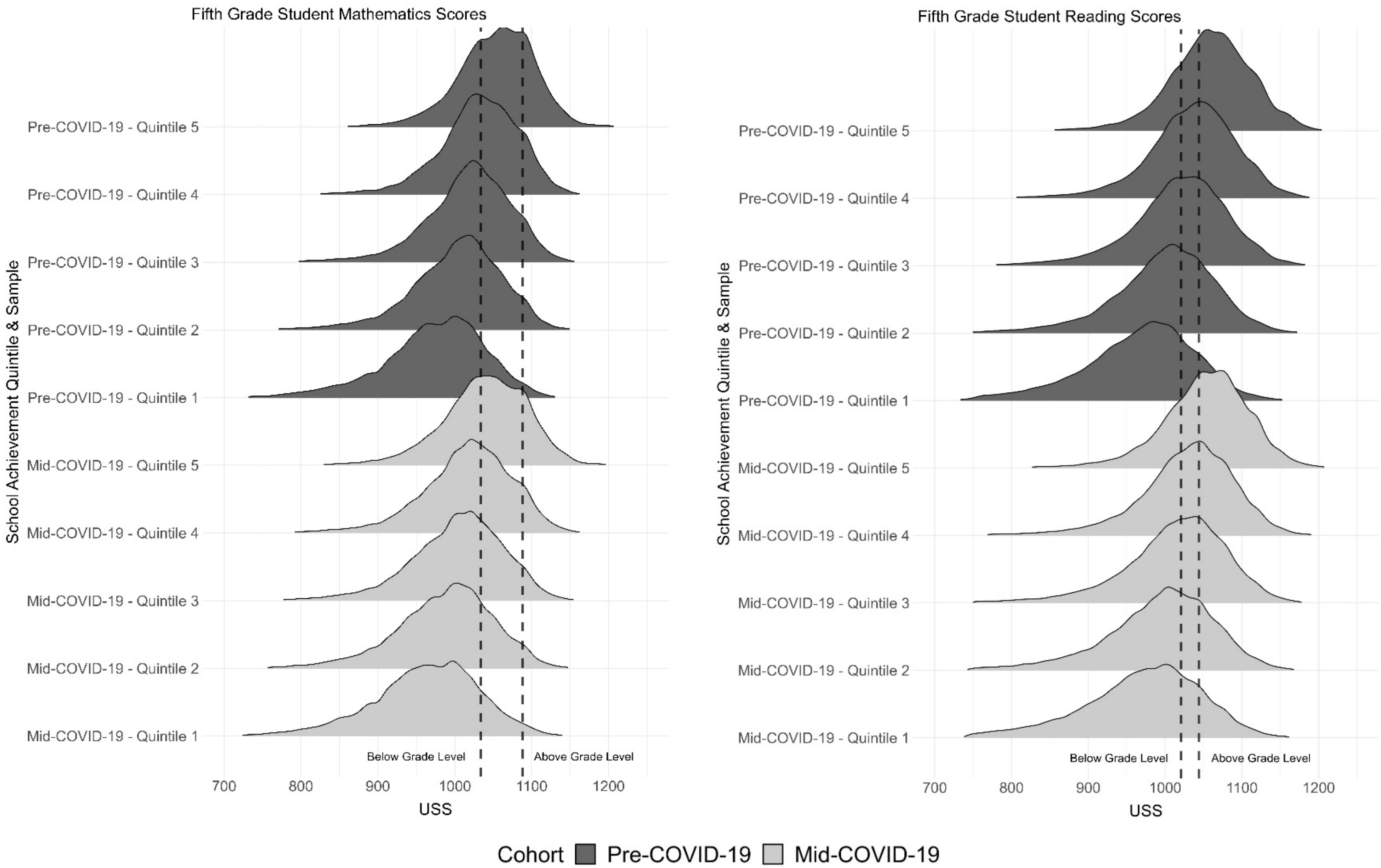
Distributions of student scores by pre-COVID-19 and mid-COVID-19 cohort by academic discipline



Notes: USS is the score metric for Renaissance Star. Students who scored at the fifth grade level are between the dashed lines.

Figure 2

Distributions of student scores by pre-COVID-19 and mid-COVID-19 cohort by academic discipline by school achievement Quintile



Note: USS is the score metric for Renaissance Star. Students who scored at the 5th grade level are between the dashed lines.

Figure 3

Estimated grade level need for fifth grade students by discipline and COVID-19 cohort

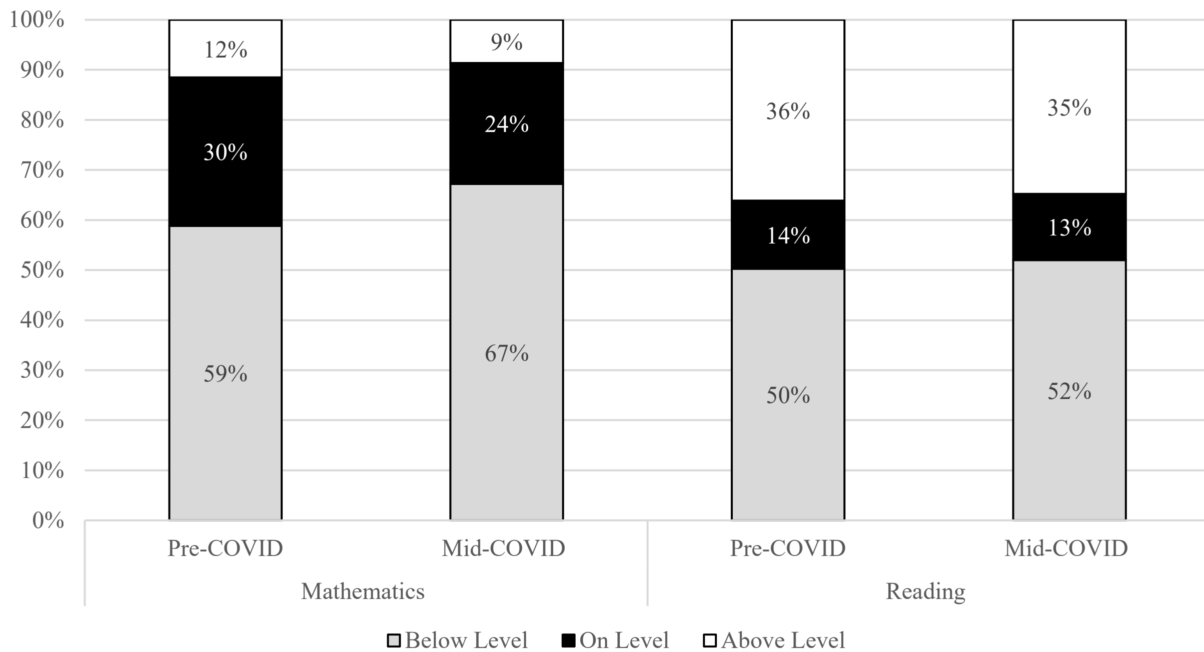


Table 1

Minimum Scores on Renaissance Star Required to be considered Level 3 (e.g., on track or met expectations) on Smarter Balanced for mathematics and reading spring scores for grades 3-7

Grade Level	Mathematics	Reading
Below 3rd	<988	<993
3rd	988	993
4th	1035	1022
5th	1088	1044
6th	1103	1068
7th	1118	1082

Table 2

Summary statistics of school level descriptive statistics by discipline and cohort

	Pre-COVID (Fall 2018)				Mid-COVID (Fall 2021)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Mathematics (n_{schools} =4,815)								
Sample Size	61.23	59.65	10	1,569	53.59	52.33	10	1,185
School Mean Score	1,014.71	30.44	856.00	1,141.36	1,001.41	34.84	840.10	1,163.56
School Variance	3,175.03	1,338.01	312.00	12,595.32	3,389.47	1,428.57	291.80	19,617.00
School SD	55.14	11.61	17.66	112.23	56.98	11.96	17.08	140.06
School Range	258.28	77.05	62	653	258.27	77.84	46	676
Reading (n_{schools}=5,214)								
Sample Size	61.02	59	10	1570	53.88	53.92	10	1,385
School Mean Score	1,014.25	35.01	842.00	1,125.63	1,009.13	38.04	764.60	1,123.24
School Variance	4,122.82	1,941.21	322.99	25,133.36	4,760.84	2,651.19	329.88	29,055.84
School SD	62.63	14.13	17.97	158.54	66.67	17.79	18.16	170.46
School Range	299.31	98.21	61	667	315.38	115.85	54	639

Table 3

Multiple regression results of the natural log of the ratio of variance for schools mid-COVID to pre-COVID for reading and mathematics on Title 1 Status, Prior Achievement, and their Interaction

	Mathematics			Reading		
	Coefficient	Bootstrapped SE	BC 95% CI	Coefficient	Bootstrapped SE	BC 95% CI
Intercept	0.06	0.01	(0.05, 0.08)	0.10	0.01	(0.08, 0.16)
Title 1 Status	0.08	0.02	(0.04, 0.11)	0.12	0.02	(0.08, 0.16)
Prior School Achievement	0.003	0.0003	(0.002, 0.003)	0.0001	0.0003	(-0.0005, 0.0007)
Title 1 by School Ach.	0.001	0.001	(-0.0001, 0.002)	0.001	0.001	(-0.0005, 0.002)
R^2	.025			.013		

Notes: The dependent variable is the ln(mid-Covid $\text{variance}_{\text{schoolj}}$/pre-Covid $\text{variance}_{\text{schoolj}}$), Standard Error (SE), Biased corrected (BC), Confidence Interval (CI), Achievement (Ach).

Table 4

Estimated Grade Level for students by discipline and COVID-10 Cohort

	Summarized at the Student Level					Summarized at the School Level					
Below, at, or above 5th grade	Pre (%)	Mid (%)	Estimated Grade level	Pre (%)	Mid (%)	Below, at, or above 5th grade	Pre (%)	Mid (%)	Estimated Grade level	Pre (%)	Mid (%)
Mathematics											
Below	58.85%	67.18%	3rd or lower	27.89%	37.10%	Below	60.01%	67.71%	3rd or lower	28.60%	37.07%
			4th	30.96%	30.08%				4th	31.42%	30.64%
On	29.63%	24.18%				On	29.26%	24.19%			
Above	11.52%	8.64%	6th	5.62%	4.18%	Above	10.73%	8.10%	6th	5.36%	4.04%
			7th	3.06%	2.25%				7th	2.88%	2.09%
			8th or 8+	2.84%	2.22%				8th or 8+	2.48%	1.97%
N	294,831	258,049					4,815	4,815			
Reading											
Below	50.26%	51.99%	3rd or lower	33.14%	35.40%	Below	49.80%	51.32%	3rd or lower	32.76%	34.68%
			4th	17.12%	16.59%				4th	17.04%	16.64%
On	13.62%	13.26%				On	13.70%	13.38%			
Above	36.12%	34.74%	6th	14.01%	13.40%	Above	36.50%	35.30%	6th	14.20%	13.69%
			7th	6.64%	6.67%				7th	6.69%	6.76%
			8th or 8+	15.47%	14.68%				8th or 8+	15.61%	14.85%
N	318,146	280,943					5,214	5,214			

Notes: *Pre* refers to the schools observed pre-COVID-19 (fall 2018). *Mid* refers to the schools when observed after the onset of COVID-19 (fall 2021).