ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 21-391

Changing Patterns of Growth in Oral Reading Fluency During the COVID-19 Pandemic

Benjamin W. Domingue
Stanford University

Heather J. Hough
Stanford University

David Lang
Stanford University

Jason Yeatman
Stanford University

Education has faced unprecedented disruption during the COVID-19 pandemic; evidence about the subsequent effect on children is of crucial importance. We use data from an oral reading fluency (ORF) assessment—a rapid assessment taking only a few minutes that measures a fundamental reading skill—to examine COVID's effects on children's reading ability during the pandemic in more than 100 U.S. school districts. Effects were pronounced, especially for Grades 2–3, but distinct across spring and fall 2020. While many students were not assessed in spring 2020, those who were seemed to have experienced relatively limited or no growth in ORF relative to gains observed in other years. In fall 2020, a far more representative set of students was observed. For those students, growth was more pronounced and seemed to approach levels observed in previous years. Worryingly, there were also signs of stratification such that students in lower-achieving districts may be falling further behind. However, at the level of individual students, those who were struggling with reading prior to the pandemic were not disproportionately impacted in terms of ORF growth. This data offers an important window onto how a foundational skill is being affected by COVID-19 and this approach can be used in the future to examine how student abilities recover as education enters a post-COVID paradigm.

Acknowledgements
Domingue, Lang, and Yeatman have consulting relationships with Literably; however, work on this analysis was not funded by Literably. Lang was supported in part by the Institute of Education Sciences (R305B140009). The authors thank Habib Moody, Sherrie Reed, Jeannie Myung, Libby Pier, and H. Alix Gallagher for comments on this manuscript as well as Literably and the districts for sharing the data. PACE working papers typically are circulated for discussion and comment purposes; the related policy brief has undergone the peer-review process that accompanies official PACE publications.

VERSION: April 2021

Suggested citation: Domingue, Benjamin W., Heather J. Hough, David Lang, and Jason Yeatman. (2021). Changing Patterns of Growth in Oral Reading Fluency During the COVID-19 Pandemic. (EdWorkingPaper: 21-391). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/v9qa-tm74

Introduction

There is widespread concern among parents, educators, and policy makers alike that the COVID-19 pandemic will result in substantial deficits in student learning (Kuhfeld, Soland et al., 2020; Dorn et al., 2020). Particularly when it comes to formative skills such as reading and math, school closures may result in learning-loss with compounding impacts over time. Initial evidence regarding losses has been mixed. Some evidence focusing on results from standardized assessments (e.g., the NWEA MAP[1]) that focus on a suite of skills suggest relatively limited losses, especially in reading (Kuhfeld, Tarasawa et al., 2020; Renaissance Learning, 2020). Other evidence suggests either larger losses or substantial heterogeneity across, for example, student age (Huff, 2020; Bielinski et al., 2020) or language status (Pier et al. 2020). However, the limitations inherent to these early analyses (i.e., a reliance on at-home administrations of lengthy standardized tests typically given in schools) as well as the scale and severity of the problem—virtually every district in the country shut down in spring 2020 and experienced substantial disruptions to standard practice in fall 2020—suggest more evidence is needed so as to both refine our understanding of the problem and to identify the specific student skills that may be impacted. For example, it may be that learning growth largely recovers after the period of COVID-related school closures or, alternatively, students may face compounding challenges in the months and years to come demanding long-term attention from policymakers.

Despite its urgency, learning loss can be difficult to study. Beyond the fact that assessments are typically done in a classroom setting (something clearly not possible through much of the pandemic period), there are two particular challenges associated with accurate understanding of the learning effects of COVID-19. One key issue is selection, i.e., which students do we observe? Understanding selection is crucial as neither exposure to distance learning nor access to the digital resources required to engage in online education (and associated educational measurements) are evenly distributed (Herold, 2020; Parolin & Lee, 2020). One possibility is that we are producing biased estimates of the size of COVID-related learning losses given that we are not observing the performance of students with least access to online educational resources. Even amongst those students with access, there are concerns that measures requiring substantial amount of student time may be influenced by parents helping their younger learners and that other aspects of at-home administration of the measures may distort the resulting scores. A second issue is that many of the standardized educational outcomes we observe are for older students, in particular those in Grades 3 and above (e.g., CAASPP[2]). While such findings are useful, they leave open the possibility that we are not observing potential changes in learning for younger learners, a group that may be especially sensitive to the shift to remote instruction.

---

[1] https://www.nwea.org/map-growth/

[2] https://www.caaspp.org/

Therefore, we focus on potential deficits in a foundational skill, oral reading fluency (ORF; Fuchs et al., 2001). ORF is the ability to fluently read text aloud. ORF depends on more basic single word decoding skills, but also requires words to be fluently read in the context of sentences. ORF is highly predictive of comprehension—that is, substantial meta-analytic evidence exists connecting ORF to other measures of reading (Baker et al., 2008; Reschly et al., 2009) and, arguably, is the best overall measure of reading competency in the early grades (Fuchs et al., 2001). Learning to read is one of the first major challenges that children face when they begin elementary school. Over the first couple years of elementary school children are expected to make the transition from learning to read to reading to learn. By third and fourth grade, math is taught through word problems, reasoning skills through discussing text, and children are expected to be able to gain knowledge about the world through reading. Thus, children who fall behind developing reading skills can quickly find themselves struggling to keep up throughout their coursework and there is thus concern that inadequacies in reading instruction during the pandemic might have cascading effects for years to come.

We assess the effects of the pandemic on reading development using data from Literably who provide an ORF assessment based on first recording students' readings of texts presented to them on a device and then using a combination of human transcription and speech recognition to score these recordings. Literably assessments are typically delivered in classroom settings but were readily transitioned to distance education as they require fairly minimal technology and time. Previous work on the human rater component suggests that this approach can be used to generate scores that are appropriately predictive of downstream outcomes (e.g., standardized test scores; Townsend & Domingue, 2018; Literably, Inc., 2018).[3] In total, we use data from nearly 100,000 students who attend schools in over one hundred school districts spread across 22 states who collectively provide over 250,000 measures of ORF.

Our focus on ORF measurements collected continuously throughout the pandemic offers several benefits that are important in light of the challenges enumerated above. First, our analysis focuses on longitudinal within-person change; we focus inference on the expected change within an individual and thus avoid bias due to differences in the composition of those who provide scores at different points in time. This approach helps to minimize bias in our estimates resulting from selection; however, our results do not necessarily generalize to all students (a fact we further discuss below). Second, the ORF measure used here is both rapid—perhaps allowing for measurement of more children than a more expansive assessment would be able to measure—and familiar (and low-stakes) to children—thus minimizing the possibility of parental aid during the measurement process. Third, ORF measures are highly reliable and are strongly associated with other types of indicators of a student's developing reading proficiency (Fuchs et al., 2001), making them an exceptional measure for tracking growth over time. For example, despite the fact that they are assessed relatively quickly, Literably scores are associated with alternative measures of student achievement (i.e., standardized test scores; Townsend & Domingue, 2018). Finally, we are able to study younger children, especially those

[3] More information about the measure can be found at https://literably.com/learn

in Grades 1–2, who are less frequently measured in other assessment scenarios but who spend those years developing an academic skill of fundamental importance: learning to read.

We examine patterns of growth in ORF across different academic years. We focus on five questions. (1) How did growth look overall during the COVID-19 pandemic? We observed significant disruptions to learning in spring 2020 but stronger growth in fall 2020. Based on these results (and the fact that growth may be expected to differ across fall and spring even absent COVID-19), we then turn to more targeted questions. (2) What happened to growth in student ORF during the spring of 2019–20 when schools were effectively shut down due to COVID-19? During that period, we observe less growth than anticipated and even no growth for some grades suggesting that reading skills may have largely stopped developing for some students when schools closed (although we also emphasize that relatively few students were observed in that spring). (3) What happened to growth in student ORF during fall 2020 when students returned at the start of an unusual school year? During this period, we observed stronger rates of growth than in the previous spring but some evidence that growth was stronger in higher-achieving districts, especially in Grades 2–3. (4) Were there disparities in growth across districts and as a function of prior performance? While we found some evidence that students in higher-achieving districts were growing faster—a concerning development that we did not observe in previous years—we did not observe differences in student growth as a function of prior ability. (5) While missingness was not as bad in 2020–21 as in 2019–20, it is still of concern. How might this missingness bias our results? We find that our results are relatively robust given the observed levels of missingness in 2020–21. However, those missing students may be suffering more substantial losses in learning as a function of COVID-19 and, even amongst the observed students, the rate of learning observed in fall 2020 won't ameliorate the loss prompted by the initial disruption to schooling.

Results

Results are based on measures of ORF—measured as the words correct per minute (WPM), or the number of words read correctly divided by the time of the audio file (which are typically between 60 and 120 seconds long)—taken from approximately 100,000 students over the last several years (see Table 1). These students come from 111 school districts in 22 states.[4] Students in these districts tend to be of higher socioeconomic status and to perform at a higher-level on standardized tests than do students from the nation's districts as a whole (see Figure A1 in the Appendix; comparisons based on data from the Stanford Education Data Archive in Reardon et al., 2019). However, the students that we study here are in districts that have relatively high levels of school closures in fall 2020 (based on cell phone tracking data; Parolin & Lee, 2020). Students are assessed 2.9 times per year on average. Data is primarily collected in first through fourth grade (with additional data from the end of kindergarten and Grades 5–7). Data collection is performed intermittently throughout the year (see Figure A2)

[4] Districts were allowed to opt out of the analysis. All student data was provided in a totally anonymized form.

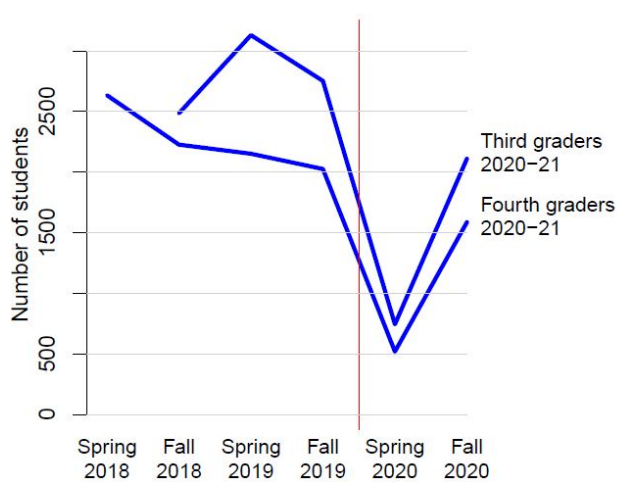
and, crucially, the composition of students providing data at any point in the year is variable (e.g., students who provide scores over the summer tend to have relatively low measures of ORF). Further description of the data can be found in the Appendix.

Table 1. Number of Students and Readings as a Function of Academic Year.

Academic year	Number of Students	Number of Readings
2017–18	8036	30235
2018–19	14531	71639
2019–20	22697	80594
2020–21	58354	121062

As with most assessments delivered remotely, not all students were able to participate, i.e., there was selection. To illustrate the degree of the selection problem, consider Figure 1. This plots the number of students with scores observed in different fall (calendar months 9–12) and spring (months 3–6) periods of the noted academic years. To be included in either cohort, a student must be observed in first grade in the first year (2017–18 or 2018–19) of the given cohort. The precipitous decline in observations in spring 2020 (i.e., the COVID-19 spring) is apparent. However, there is also a substantial increase in the number of observations in fall 2020 for these cohorts. We use these facts—a relative sparsity of testing in spring 2020 with a fuller, but potentially incomplete, return to testing in fall 2020—to structure our interpretation of results below.

Figure 1. Number of Students in Two Cohorts of First-Graders Observed at Different Periods.



Given these issues, we deploy a model that focuses on longitudinal within-person change (details on our modeling approach are in the Appendix). This model addresses expected performance within-person over time rather than differences in performances between individuals. As a first example of this approach, we use it to estimate baseline levels of ORF growth in 2018–19, an academic year unaffected by COVID-19. Here, we estimate growth as words per minute per day (WPM/day). Given natural changes of reading ability across the gradespan (i.e., ORF grows more slowly as students age (Hasbrouck & Tindal, 2006)), we compute values separately for each grade. Based on this model (i.e., Equation 1 in the Appendix), Table 2 shows the expected growth of ORF for students in a given grade. Growth decelerates throughout the grades we observe starting at roughly 0.12 WPM/day in kindergarten and declining to below 0.07 WPM/day in the higher grades. We can also translate this to the number of days before a student is expected to have an ORF score one word higher. In kindergarten and first grade, it starts at roughly 8 days and grows from there. Students in Grades 2 and 3 are expected to have higher scores after 10–14 days while older students take even longer to see score increases. Given the patterns observed here and the patterns associated with data collection (see Figures A3 and A4 in the Appendix), we focus analysis on Grades 1–4.

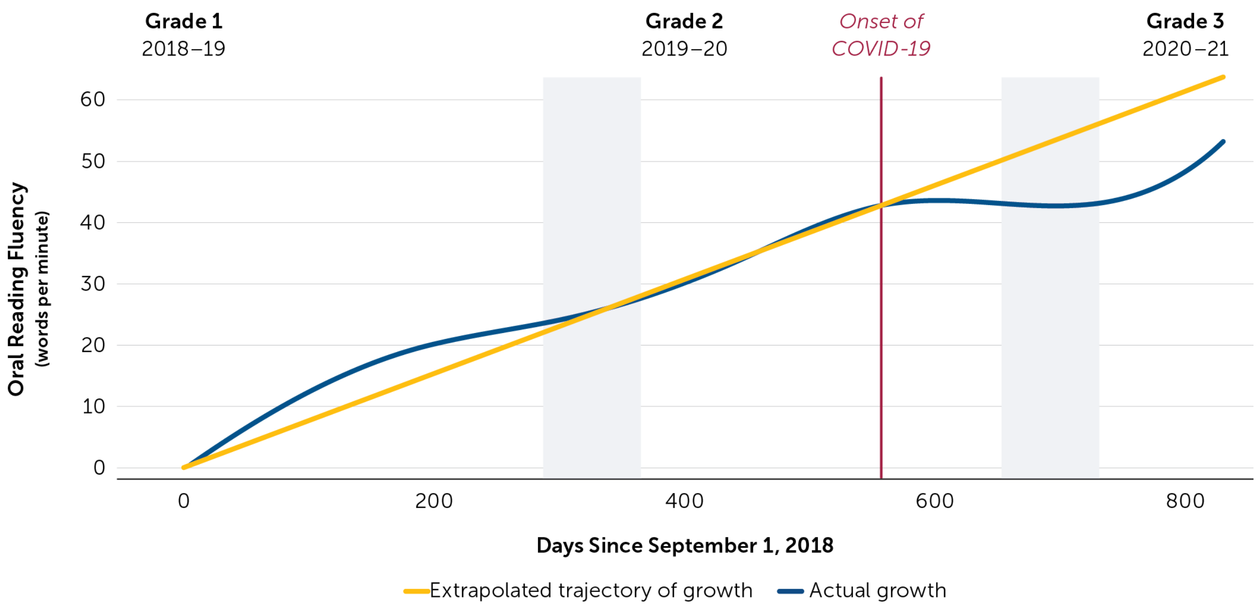
Table 2. Oral Reading Fluency Growth Rates (Based on 2018–19 Data).

Grade	Words per minute/day	Days(+1 word per minute)
K	0.121	8
1	0.118	8
2	0.091	10
3	0.068	14
4	0.056	17
5	0.052	19

The Pandemic Affected Growth

When we track students over multiple years, we observe COVID-related declines in the rate of growth of ORF. We use flexible models—Equation 2 in the Appendix, which utilizes B-splines (Friedman et al., 2001) to allow for a nonlinear analysis of growth—to descriptively examine growth trends for students. We start with a long-run view of growth (Figure 2) amongst students we observe for a relatively long period and for whom the COVID-19 pandemic occurs during Grades 2–3. Prior to the pandemic, these students observed relatively consistent growth (allowing for some fluctuation during summer, demarcated as blue regions). However, the disruptions to learning due to the COVID-19 pandemic are apparent in the form of a sudden flattening of the blue curve around the onset of the COVID-19 pandemic (i.e., the vertical red line marking March 2020). We emphasize this disruption by extrapolating total growth from the start of first grade through the onset of the pandemic (the black line). Note the emergence of a gap in spring 2020 and its subsequent shrinking in the fall of the following academic year; these are patterns we follow up on in subsequent analyses. Recall that the model relies entirely on time since the start of 2018–19; in particular, we do not specify a structural break in growth, for example, associated with the onset of the pandemic.

Figure 2. Growth Curve in Oral Reading Fluency for First-Graders Beginning in 2018–19 and Extending Through Fall 2020–21.

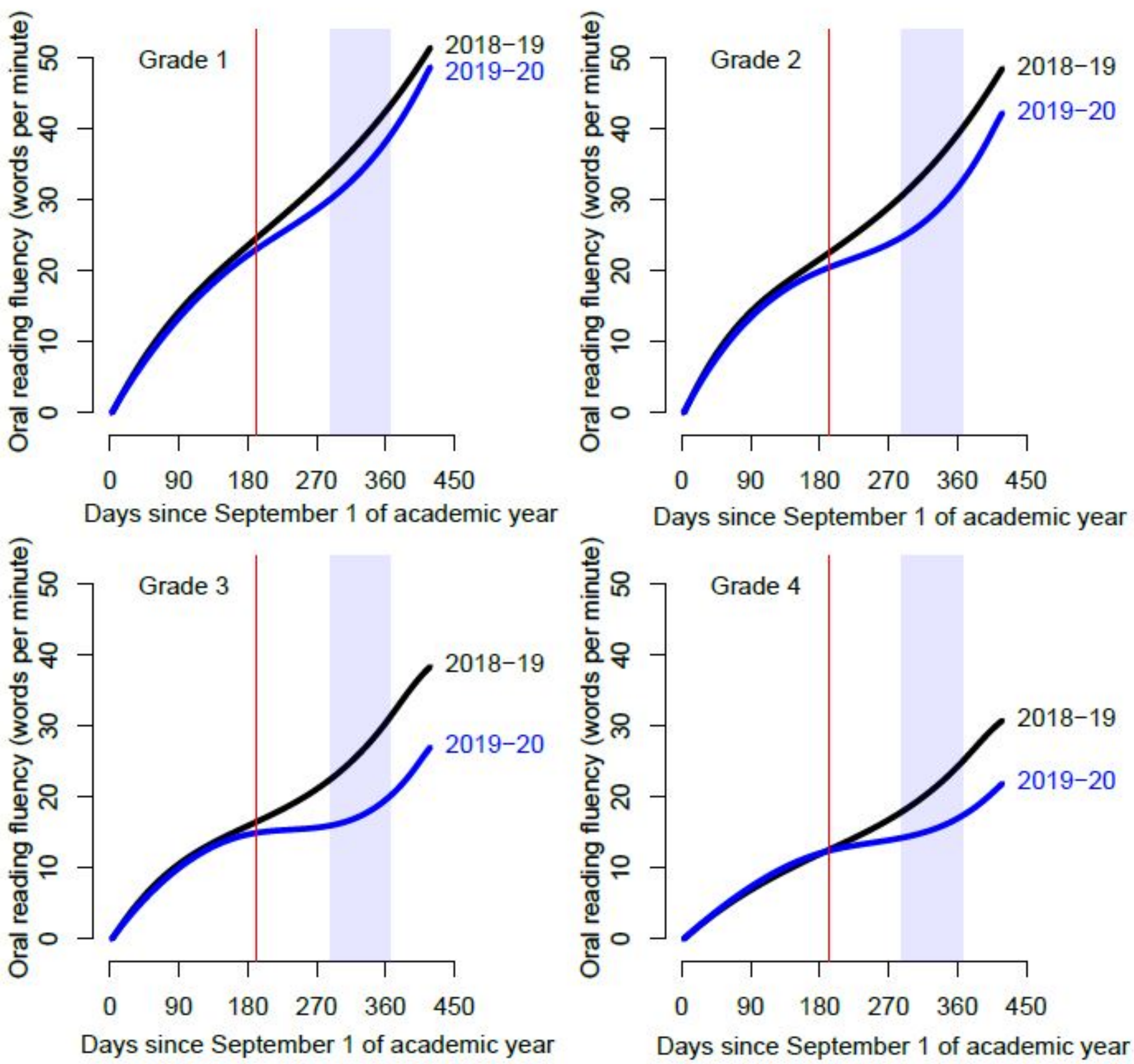


Note. Vertical line represents approximate point in year at which COVID-19 pandemic lead to widespread disruption (March 11, 2020; red line). Summers are roughly demarcated via blue background. Black line extrapolates previous growth through pandemic.

Growth during the COVID-19 period seems weak, but we now seek something to which it can be compared. This problem—identification of a counterfactual—is a key problem of causal inference (Holland, 1986). What level of ORF growth would we have observed for these students in spring 2020 and fall 2020 absent COVID-19? A first attempt was shown via the black line in Figure 2 but, given that we know growth decelerates as students age, this is a suboptimal comparison. We attempt to generate more refined counterfactuals in subsequent analyses via targeted comparisons. One concern is that some students may not be observed during the pandemic (if they, for example, lack the digital resources needed to engage in remote instruction). We attempt to address this by focusing largely on students that we consistently observe pre- and post-COVID onset. One intuitive comparison involves growth in a given season of an academic year to growth observed in that season in a different academic year. We might compare growth in spring 2020 to that in spring 2019 in Figure 2. That comparison from Figure 2 compares growth of the same students across grades; to avoid this comparison of students at different ages, we focus on comparing growth in the same grade from different academic years. However, this raises a new problem: due to changes in the composition of students receiving scores in a given academic year, comparisons may be between different types of students. To address this, we focus on comparisons between students observed consistently across time.

We turn now to direct comparisons between different cohorts focusing on the 15 months following the start of a given academic year (Figure 3). We focus on 2018–19 and

Figure 3. Comparison of Growth Curves by Grade in Oral Reading Fluency for 2018–19 and 2019–20 Based on Students Who Provided at Least One Score During the COVID-19 Pandemic.



Note. Vertical line represents approximate point in year at which COVID-19 pandemic led to widespread disruption (March 11, 2020; red line). Summer is roughly demarcated via blue background.

2019–20 cohorts but follow these students into the subsequent academic year. We analyze grades separately given the different rates of growth observed in Table 2. Prior to March of a given academic year, growth rates were comparable in each cohort. However, there is a clear impact on learning—note the flattening of the blue line relative to its counterpart—associated with COVID-19. We stylize the patterns as follows: there is (a) no estimated gain in ORF during the months prior to summer following the onset of the COVID-19 pandemic, (b) flat growth during summer (recall that relatively few scores are collected during that time; see Figure A3 in the Appendix), and (c) a return to growth in the following fall.

Based on the patterns in Figure 3, we now turn to two targeted questions focusing on concentrated windows of time. We first ask about the learning loss observed in spring of 2020 associated with the sudden school closures that March. We anticipate relatively sharp changes in growth. We then turn to fall 2020, when growth in ORF seems to have again taken place, and focus on comparisons of growth in 2020–21 to that of previous academic years. Consideration of these concentrated time windows will allow for more refined analyses given the differences between them in, for example, the degree to which we observe a selected sample (i.e., Figure 1). In these analyses, we will parametrize growth as linear (which, we argue, is a reasonable assumption given patterns observed in Figures 2 and 3). We omit further analysis of summer given that relatively little data is collected in that period.

ORF Scores Did Not Show Anticipated Gains in Spring 2020

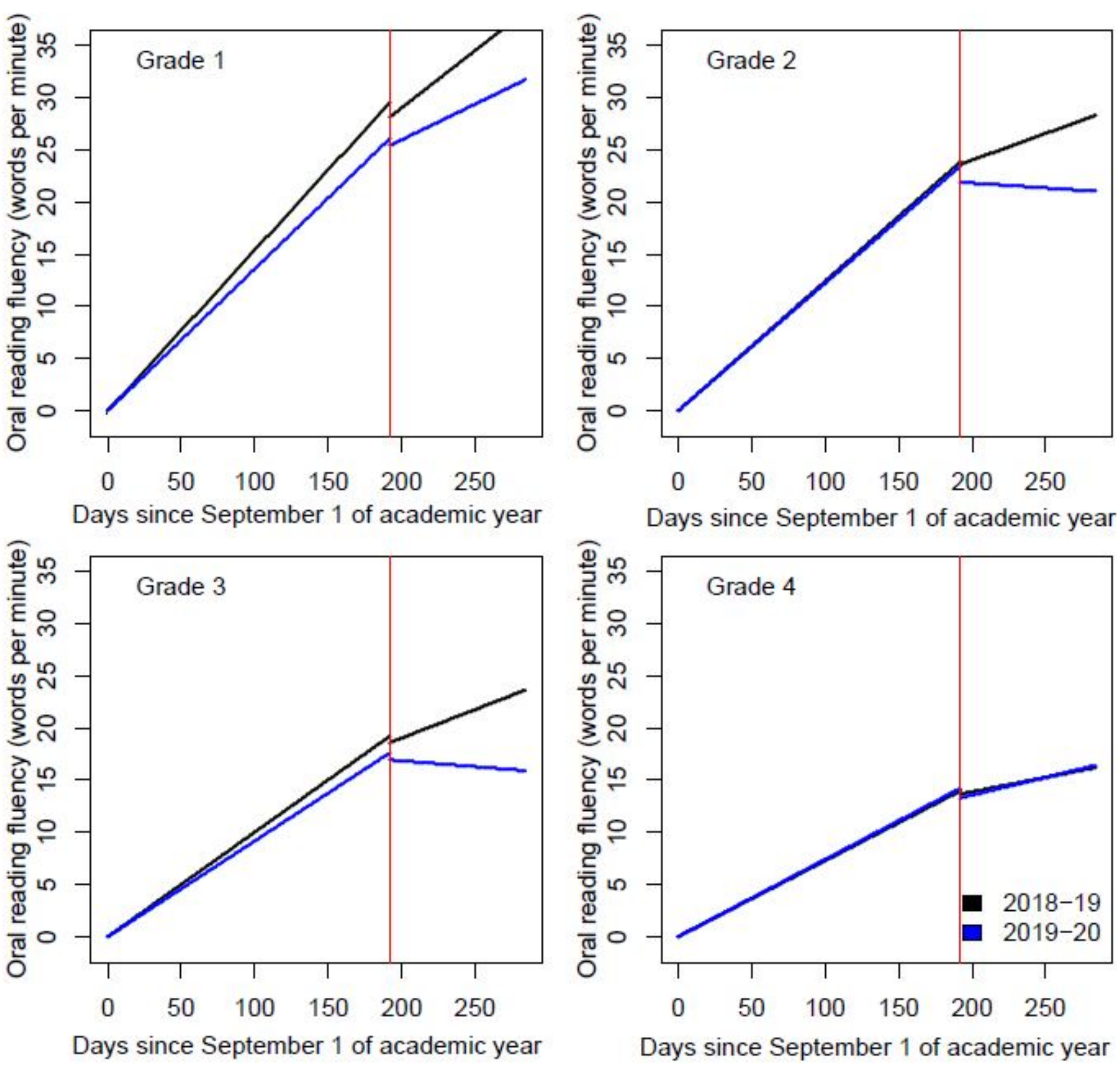
Observed ORF growth during spring 2020 was weak following the onset of the COVID-19 pandemic and we suspect that even this weak growth may be a substantial overestimate given the level of missingness observed that spring. We emphasize that most students were not observed in spring 2020. Of the students we observed in 2019–20, we only observe 30% of them post-COVID onset; see Table 3 which builds on Figure 1 in documenting substantial attrition in spring 2020. In contrast, of the students in 2018–19, we observe more than 90% of those students that spring. COVID-19 resulted in a substantial degree of selection in spring 2020.

Table 3. Proportion of Students Observed in 2019–20 Post-COVID Onset.

	Observed Post-COVID Onset?	
Grade	No	Yes
1	0.69	0.31
2	0.73	0.27
3	0.71	0.29
4	0.67	0.33

Based on the results in Figure 3, we summarize changes in a given academic year succinctly by allowing for piecewise linear growth at different points in the academic year (i.e., we allow for different linear growth in ORF before/after the onset of the COVID-19 pandemic; Equation 3 in the Appendix). Results are in Figure 4 (coefficient estimates in the Appendix). Focusing on the first part of the year prior to March (at which point the 2019–20 cohort became affected), growth is relatively comparable across cohorts in all grades. For 2018–19 students, growth continues through the rest of the year but at a slower pace. Such a deceleration is consistent with, perhaps, instructional and behavioral changes associated with the coming summer break.

Figure 4. Estimated Growth in 2018–19 and 2019–20 Based on Piecewise Linear Model.



In 2019–20, spring growth in Grades 1–3 tended to be flatter than that observed in spring 2019. Consider Grade 3. There, prior to March, growth in ORF is consistent across the academic year. In 2018–19, 3rd graders see an expected gain of 0.10 WPM/day versus 0.09 WPM/day in 2019–20 (see Table A3 in the Appendix). In 2018–19, when COVID-19 was a nonissue, growth declines following March but is still positive at 0.055 WPM/day. For 2019–20 students, however, growth is slightly negative during the spring; students expect to lose -0.011 WPM/day. Thus, while growth in the fall across academic years is similar, we see quite distinct patterns during the spring. These results suggest that the unprecedented interruption in schooling during 2019–20 has led to real, tangible losses in ORF gains for students in Grades 2–3 (growth estimates across academic years following the structural break were not statistically different for Grades 1 and 4). The learning loss for second and third graders would leave them 7.3 and 7.7 WPM behind their expected level respectively (representing 26 percent and 33 percent of the expected yearly gains).

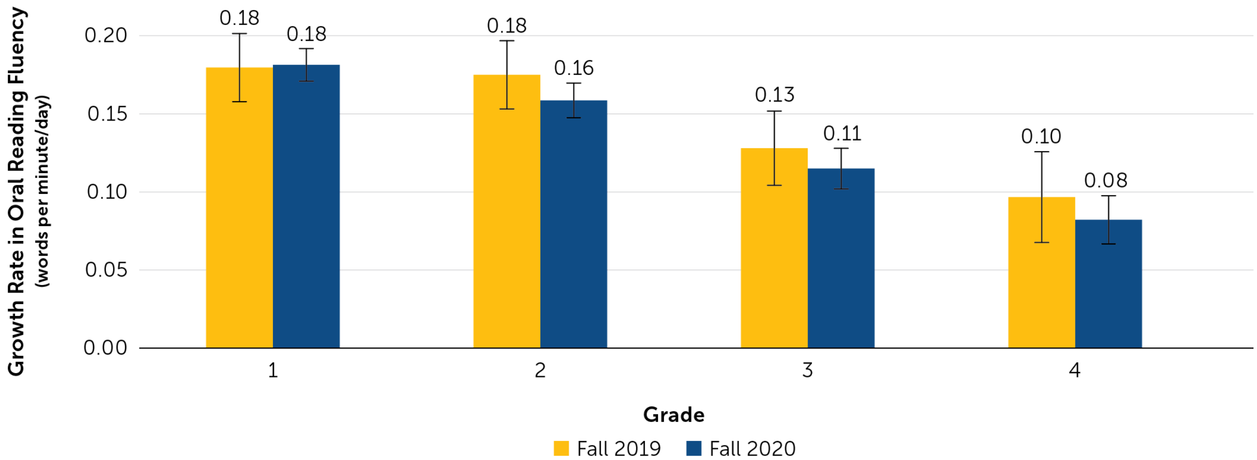
We again emphasize that these observations only pertain to the roughly 30% of students observed in spring 2020. For the remaining 70%, given that they were unobserved during this period, we cannot know how they would have done. However, we can note some systematic differences between the observed and unobserved students. In particular, students not observed in spring 2020 grew *slower* in a non-affected period (see Figure A5 in the Appendix). We would thus suggest interpreting the estimates shown here as upper bounds on the true rate of growth in spring 2020 given that unobserved students are likely to be those experiencing the most extreme educational disruptions.

Growth in ORF Largely Returned to Normal in Fall 2020

In spring 2020, school shutdowns were sudden; educators did not have time to prepare for remote instruction and students were unaccustomed to this mode of instruction. In contrast, while remote instruction in fall 2020 offered novel challenges for educators, it was at least expected and more familiar for both students and educators. Thus, it may be associated with different patterns of growth than were observed in spring 2020. We first look at linear growth rates (i.e., Equation 1 in the Appendix) in 2020–21 versus 2019–20 for the first 3 months of school based on those students observed in 2020–21.

Results are shown in Figure 5. Growth rates are fairly similar within-grade across the two cohorts. Growth is nearly identical in Grade 1 across the two years and slightly lower in fall 2020 compared to 2019–20 for Grades 2–4. Compared to the results for spring 2020, this evidence is reassuring in suggesting that students are developing crucial reading skills at rates comparable to previous years despite the hardships associated with education in 2020–21. However, questions remain. One question is whether returns to growth are evenly distributed. A second question has to do with the impact of selection on these results. While missingness was less of a problem in fall 2020 than spring 2020, this is still an issue of concern. We address these issues in the following sections.

Figure 5. Growth in First 90 Days of 2019–20 and 2020–21 for Those Students Who Are Observed in 2020–21.

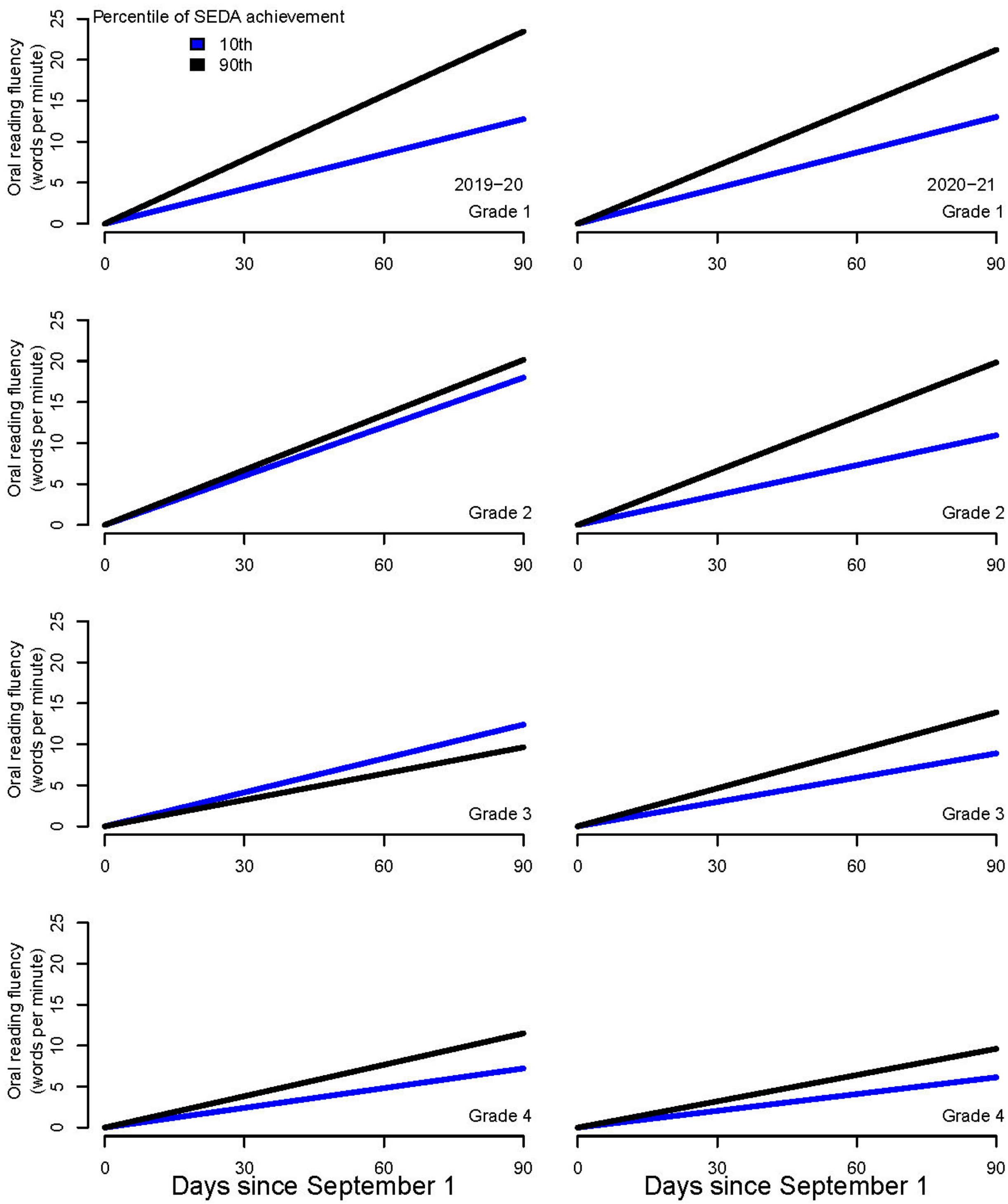


Novel Differences in Growth Across District Emerged

Given the challenges associated with 2020–21 and the fact that resources required to thrive in remote instruction are not universally available, we are concerned about different levels of growth across school districts. We thus consider a stratified analysis of growth in 2020–21 based on the mean achievement in the district. To quantify achievement, we use data from SEDA (Reardon et al., 2019) which uses federal reporting of performance on state-mandated assessments to develop an overall index of student achievement within a district (Reardon et al., 2017). In the pre-COVID fall 2019, growth in ORF tended to be similar in high- and low-achieving districts (see Figure 6). The one exception was in Grade 1 where high-achieving districts grew faster; ORF growth was 0.20 WPM/day on average but a one-*SD* increase in the district's achievement was associated with an increase of 0.04 WPM/day (*SE* = 0.011; point estimates and standard errors are shown in Table A2 in the Appendix). These results can be seen in Figure 6 wherein we compare growth for districts at the 10th and 90th percentiles of achievement; note that growth is largely similar in low- and high-achieving districts in 2019–20.

In contrast, in 2020–21, we observe accelerated growth in Grades 1–3 in high-achieving districts. Base rates of growth for districts with mean achievement are 0.18, 0.16, and 0.12 WPM/day respectively for Grades 1–3. A one-*SD* increase in achievement is associated with increases of 0.03 (*SE* = 0.006), 0.04 (*SE* = 0.007), and 0.02 (*SE* = 0.008). If we instead consider the socioeconomic status of the district, results are similar; students in more affluent districts tend to exhibit more rapid ORF growth (Table A2 in the Appendix). Especially in Grades 2–3, COVID-19 may be introducing additional inequality in reading levels across school districts. In contrast, we also consider differences in current year growth as a function of prior year score. Although these estimates are based on smaller samples, differences are muted (Table A2 in the Appendix). We do not observe increasing differences between relatively high- and low-ORF students in fall 2020.

Figure 6. Fall Oral Reading Fluency Growth in Low- and High-Achieving (Based on Percentiles of SEDA Achievement) School Districts.



The Impact of Missingness on Fall 2020 ORF Growth Estimates

Given the nature of disruptions due to the COVID-19 pandemic, we are concerned about the role of missingness on our growth estimates. We focus on fall 2020 given that missingness was so severe in spring 2020 (for spring 2020, we encourage interpretation of growth estimates as upper bounds on true growth). We focus on three questions: How many students are missing? Who are they? What is the effect of this missingness on fall 2020 ORF growth estimates?

How many students are missing? Table 4 shows the proportions of students observed across consecutive academic years.

Table 4. Pattern of Observation Across Two Academic Years.

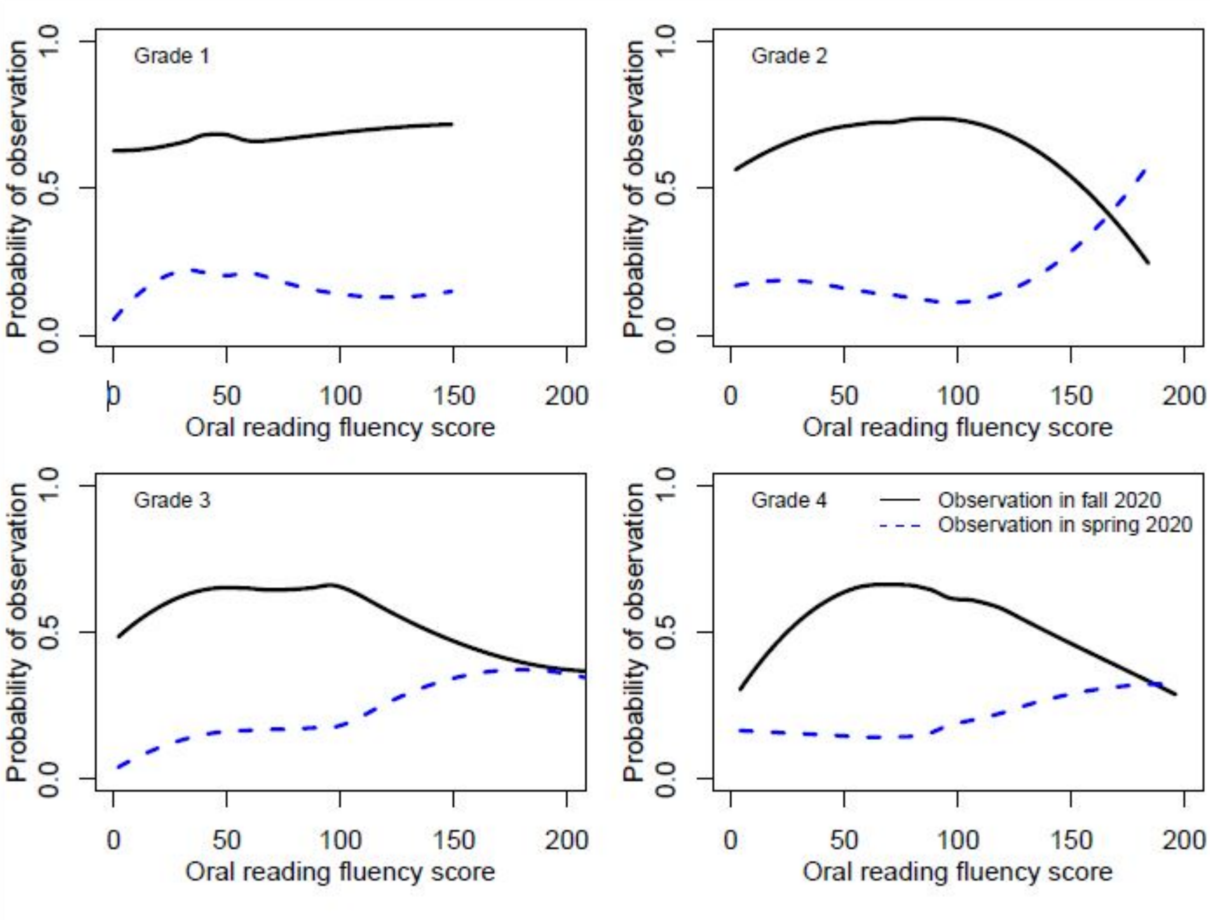
	Students observed in Y1		
Grade	*N*	Percent missing in Y1	*N*(new) Y2
2018–19 → 2019–20			
1	2406	22.4	2031
2	3123	23.9	2273
3	3653	30.8	1657
4	3254	28.5	2695
2019–20 → 2020–21			
1	3302	26.9	7237
2	3899	36.2	7843
3	4649	40.5	7429
4	4184	40.5	7216

Note. The first set of results pertains to the pattern of observations in fall of 2018–19 and 2019–20 (both unaffected by COVID-19). The second set of results follows students from fall 2019–20 to 2020–21 (fall of 2020–21 is affected). The second and third columns show the size of the student sample in the first year (Y1) for each grade and the percentage of students not observed in Y2. The final column shows the number of new students in the subsequent grade in Y2.

For example, of the 2406 students observed in Grade 1 of 2018–19, (100 – 22.4 =) 77.6% were also observed in 2019–20 (Grade 2). However, in Grade 2 of 2019–20, a new group of 2031 students was also observed (i.e., these were students not observed in Grade 1 of 2018–19). We emphasize two important facts. First, of those observed in the first year, roughly 5–10% more students were not observed in the calendar year affected by COVID-19 (2020) as compared to academic year 2020–21. To give a specific example, 26.9% of the students observed in Grade 1 of 2019–20 were not observed in 2020–21, this is 4.5% higher than the 22.4% not observed in 2019–20. Compared to the roughly 70% of students not observed in spring 2020 (Table 3), this is reassuring. However, it is still the case that this missingness may lead to bias in our estimates; we attempt to account for this in making adjustments to our estimates as explained below. Based on the estimates shown here, we assume that approximately 5–10% of the students are unobserved in fall 2020. Second, there were large shares of new students observed in fall 2020 (i.e., that were not observed before this point) largely due to the incorporation of new school districts into the Literably system. We include these students in analyses given that they represent a relatively large number of students; however, these new students may differ in key ways from students observed in earlier years, suggesting that findings (especially Figure 6) need to be interpreted accordingly. Note also that we are unable to benchmark the level of missingness in these districts and assume it is of similar magnitude to the previously observed districts.

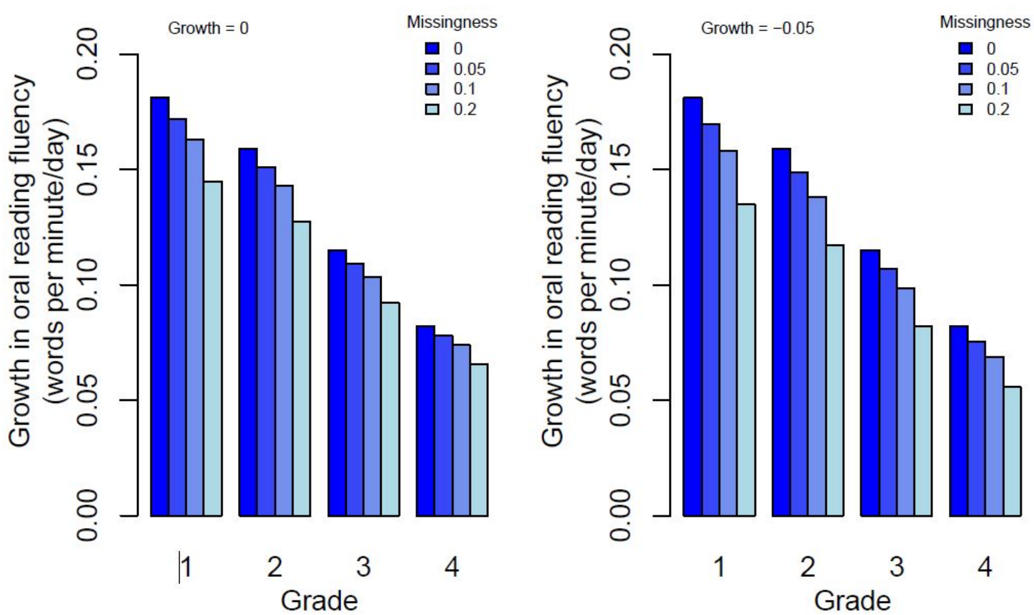
Who is missing? We examine this question by looking at probability of observation in a given period (fall or spring of 2020) as a function of the fall 2019 score, see Figure 7. There were relatively weak trends in selection as a function of ORF in fall 2019. This was true for observations in both the spring and fall of 2020. Fitted probabilities for observation in fall 2020 are approximately 0.6 but note that this accounts for both missingness due to COVID-19-related selection and generic "churn" (i.e., note that between 20–30% of students are missing due to this kind of churn in unaffected years, see Table 4). The Appendix contains additional characterizations of missing data; observed students tend to grow more rapidly than unobserved students in non-affected periods.

Figure 7. Estimated Probability of Observation in Fall or Spring 2020 Based on Mean Fall 2019–20 Oral Reading Fluency (for Those Students in Districts Observed in the Given Period).



What is the effect of missingness? We can adjust our growth estimates in Figure 5 based on different assumptions about the proportion of students we are missing and their growth rates, see Figure 8. We assume that unobserved students exhibit either flat or slightly negative (-0.05) growth in their ORF during fall 2020 (although, of course, we do not observe this). The assumption of growth of -0.05 WPM/day for unobserved students is fairly extreme, recall that growth in in fall 2019 was 0.09 WPM/day while in spring 2020 was only -0.01 WPM/day; we are assuming that students are losing ORF abilities much more rapidly than observed in spring 2020 and roughly half as fast as they gained them during a period of non-COVID affected instruction. Based on the values in Table 4, we suspect roughly 10% of students are unobserved in 2020–21 but consider a range of values between 0–20%. Adjusted estimates are shown in Figure 8. When no students are missing, growth corresponds to that observed in Figure 5. As larger proportions of students are missing, estimates decline towards zero. For example, the estimate of 0.18 for growth in Grade 1 declines to 0.16 if 10% of the students are missing and have no growth. Absent relatively extreme assumptions (e.g., negative growth for large proportions of missing students), these declines are relatively modest. However, that does not change our concern about the missing students. Their learning may have been materially altered by the pandemic, we have no way of knowing.

Figure 8. Adjusted Estimates for Fall 2020 After Accounting for Different Proportions of Missing Respondents.



Note. We adjust estimates based on finding the expected value for a probability distribution that is a mixture of the implied normal distribution for the estimate (i.e., centered at the observed estimate with standard deviation being the standard error) plus a point mass at the unobserved growth rate (0 or -0.05); the mixture rate is such that the point mass is observed with probability p. Estimates based on no missingness (p = 0) are identical to those in Figure 5.

Discussion

The COVID-19 pandemic has resulted in a clear disruption to "business as usual" in K–12 education. Researchers have begun to monitor incoming data for signs of learning loss. Results have been mixed (Kuhfeld et al., 2020; Renaissance Learning, 2020; Huff, 2020; Bielinski et al., 2020). Here, using measures based on a foundational skill (ORF), we demonstrate clear loss of learning for younger students—particularly in Grades 2–3—during spring 2020. This is consistent with similar reporting of lower-than-expected ORF in fall 2020 (May, 2020). However, learning gains are occurring at a rate similar to that observed in earlier years in fall 2020 although there is evidence for heterogeneity across districts. Below we further unpack and contextualize our findings and what they tell us about growth in student ORF amidst COVID-19.

In spring 2020, schools closed suddenly and educators were left scrambling to create an entirely new paradigm for virtual instruction from scratch for students who not only had a range of educational needs but also a broad range of access to both the internet and necessary devices needed to engage in virtual instruction. This abrupt change was associated with a real decline in growth of ORF for students. In fact, we suspect our estimates of learning losses (Figure 4) are underestimates of the true effects given that we only observed a fraction (roughly 30%) of students in spring 2020.[5] Given that observed students had at least some access to the relevant technology (i.e., the technology needed to take the test), observed students are perhaps those who experienced less substantial impacts to their learning as compared to those we don't observe. In particular, the observed growth in spring 2020 for Grades 1 and 4 could substantially distort what is occurring for unobserved students. These losses in spring 2020 are concerning and future research should attempt to monitor whether the impact on ORF growth caused by this disruption has long-term consequences (i.e., how do long-term trajectories of growth look in the coming years?).

By fall 2020, educators had time to prepare modified instructional plans—i.e., use of new technological aids, identification of the most at-risk students, distribution of additional resources needed for students to access virtual resources, a limited return to in-person instruction in some places—that they would pursue and these plans seem to have been relatively effective in spurring ORF growth (Figure 5). However, we emphasize three reasons for caution in interpreting these results. First, we estimate 5–10% of students were not observed in fall 2020 and these students might be experiencing more troublesome learning losses. Second, our comparison to 2019–20 is perhaps not the one we should care about. In Figure 5, we compare growth of students in fall 2020 to the growth observed in prior years. Students in these prior years may make for a poor counterfactual in one crucial sense: they experienced

[5] We do not attempt to adjust these estimates for missingness given the scale of the missingness problem here and instead prefer to suggest that they be interpreted as upper bounds on the true decline in ORF growth.

normal instruction in the spring of the prior year. In particular, we would not interpret our results as necessarily implying that reading instruction in fall 2020 was as effective as typical instruction would have been; the unique combination of COVID-related circumstances experienced by the students make identification of an appropriate counterfactual—especially for fall 2020—challenging. Third, and perhaps most crucially, gains are unequal across schools (Figure 6) and may be introducing new skill gaps in reading. While these factors suggest areas for concern, we think it encouraging that many students do seem to be developing oral reading skills in 2020–21.

Above and beyond the issue of selection we have emphasized throughout, we acknowledge additional limitations. First, we do not investigate possible heterogeneity as a function of how the district is operating in 2020–21 (e.g., are classes in-person or remote?). Differences in learning mode are also associated with differences in, for example, student demographics (Parolin & Lee, 2020). In future work, we hope to examine variation in ORF growth patterns as a function of the nature of education (e.g., in-person? remote? hybrid?) in 2020–21. Second, the school districts we observe are not a random sample; compared to all school districts in the U.S., we observe districts that are relatively high-achieving and that have relatively higher levels of closures in September 2020. Taken together, these facts perhaps suggest that our growth estimates for fall 2020 may not entirely generalize to unobserved districts (even omitting the other potential sources of concern about fall 2020 estimates). Third, our approach conflates COVID-related changes in ORF growth with those due to at-home administration of ORF measures; for example, one potential explanation for the observed change from spring to fall of 2020 could be an increase in familiarity with the digital environment for many students. Fourth, results for ORF need not generalize to other subjects. In fact, given the centrality of reading to elementary education, we would strongly suspect that they do not.

COVID-related learning losses have the potential to harm a large number of students. Understanding these learning losses is thus important. Our findings suggest that reading skills were substantially impacted by the COVID-19 disruption in the spring; even if growth was improved in fall 2020 it may not be sufficient to make up for that loss. Further, the pandemic is not over. School disruptions will continue through 2020–21. The analytic platform we have described here can be used to monitor changes in learning associated with both further disruptions and abnormalities in learning environments during the current academic year but also to monitor whether students return to normal levels of ORF following the resumption of typical educational activities. In particular, future work will aim to examine differential growth associated with variation in policies implemented across various districts. Such evidence will be useful in understanding which policies are succeeding and which students continue to need assistance.

Appendix

Data

Students in our study come from districts that use the Literably ORF assessment. We focus on *N* = 111 districts that agreed to participate (districts were given a choice to opt out of this study). The average district contributed 885 students and 4205 scores. There is substantial variation across districts; the IQR for number of students spanned 86–806 while the IQR for the number of scores spanned 185–2366.

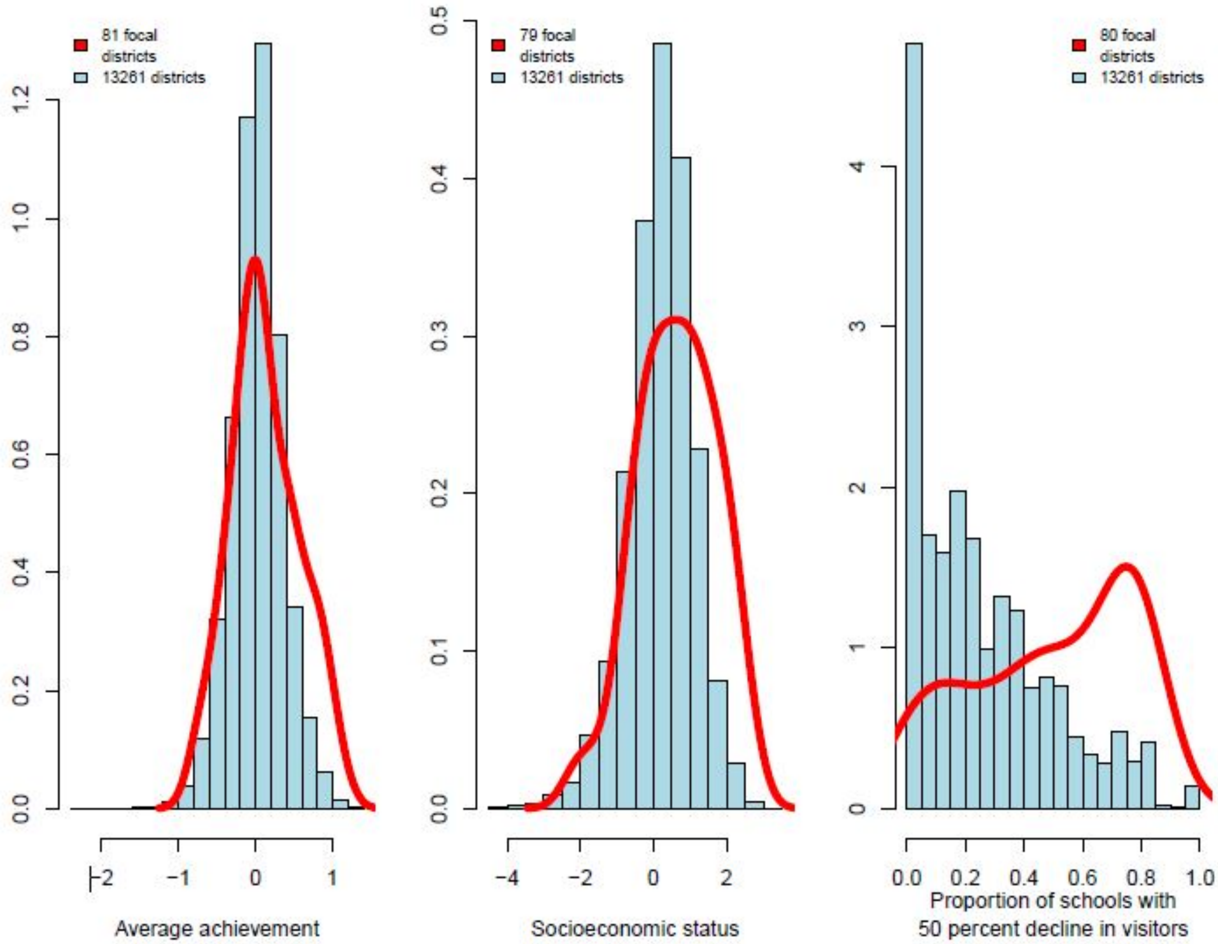
Via the district ID, we link the student-level data to additional datasets. Some districts could not be linked to the external data; precise numbers vary across the different external datasets, but we are able to match roughly 80 of the districts that collectively contain roughly 95% of the ORF scores. We focus on three comparisons (see Figure A1):

- First, we compare the average academic performance of our sample of districts to the full set of U.S. districts using data from the Stanford Education Data Archive (SEDA) (Reardon et al., 2019). We use an overall index of achievement on standardized tests.[6] The focal districts tended to be somewhat higher achieving than U.S. districts; the mean achievement across the focal districts was at the 62nd percentile of the distribution of achievement for all U.S. districts.
- Second, we compare the average socioeconomic status (SES) again using SEDA (Reardon et al., 2019). Focal districts tended to be higher status; the mean focal district was at the 69th percentile of the distribution of status for all U.S. districts.
- Third, we compare levels of school closure amongst our sample to U.S. districts using data that tracks school closures by looking at year-on-year changes in school visits (based on mobile phone data, Parolin & Lee, 2020). We characterize a school as mostly closed in a given month of 2020–21 if there is more than a 50% decline in the number of visitors from 2019–20 to 2020–21. We focus on the proportion of elementary schools in a district that meet this definition in September 2021. Our districts display higher levels of closure than do U.S. districts in general; the average district in our sample has a level of closure at the 82nd percentile of the full distribution in U.S. districts. This is due to the geographic locations of our districts and is a function of the fact that some states have high levels of school closure (i.e., > 70% while others do not (i.e., < 20).

---

[6] In particular, we use the variable described as: Geo Dist Mean SEDA EDFacts Test-Based Achievement Pooled Across Subjects (Math & ELA), Ordinary Least Squares (OLS) estimate, Cohort Scale (CS).

Figure A1. Comparisons of Focal Districts to All U.S. Districts.



In total, from the 111 districts, we focus on 98,210 students each of whom provides 4.8 scores on average. Variation in the sample sizes across grades is captured in Table A1. Smoothed scatterplots with local regression trajectories (i.e., LOESS; Cleveland et al., 1992) trajectories are shown in Figure A2. Scores are somewhat less frequently collected during the winter break and then there is a clear gap in collection during summer. Note that the LOESS trends are relatively flat; this is due to the fact that different students are assessed at different times (i.e., relatively low performing students are assessed during summer) and is something we address in our analyses via inclusion of a person fixed effect.

A key issue is that students have differential access to ORF assessments following the onset of the COVID-19 pandemic. Figure A3 shows the number of scores collected in each month of 2018–19 and 2019–20. The relative paucity of readings in month 9 (May) for 2019–20 shows the problem with a naive comparison of scores collected pre- and post-COVID onset.

Table A1. Number of Students and Scores/Student for Each Grade Across Entire Dataset.

Grade	Number of students	Scores/student
K	8195	2.67
1	19661	3.87
2	22471	3.74
3	24847	3.25
4	24726	2.92

Figure A2. Scatterplot of Scores as a Function of Time Since the Beginning of 2018–19.

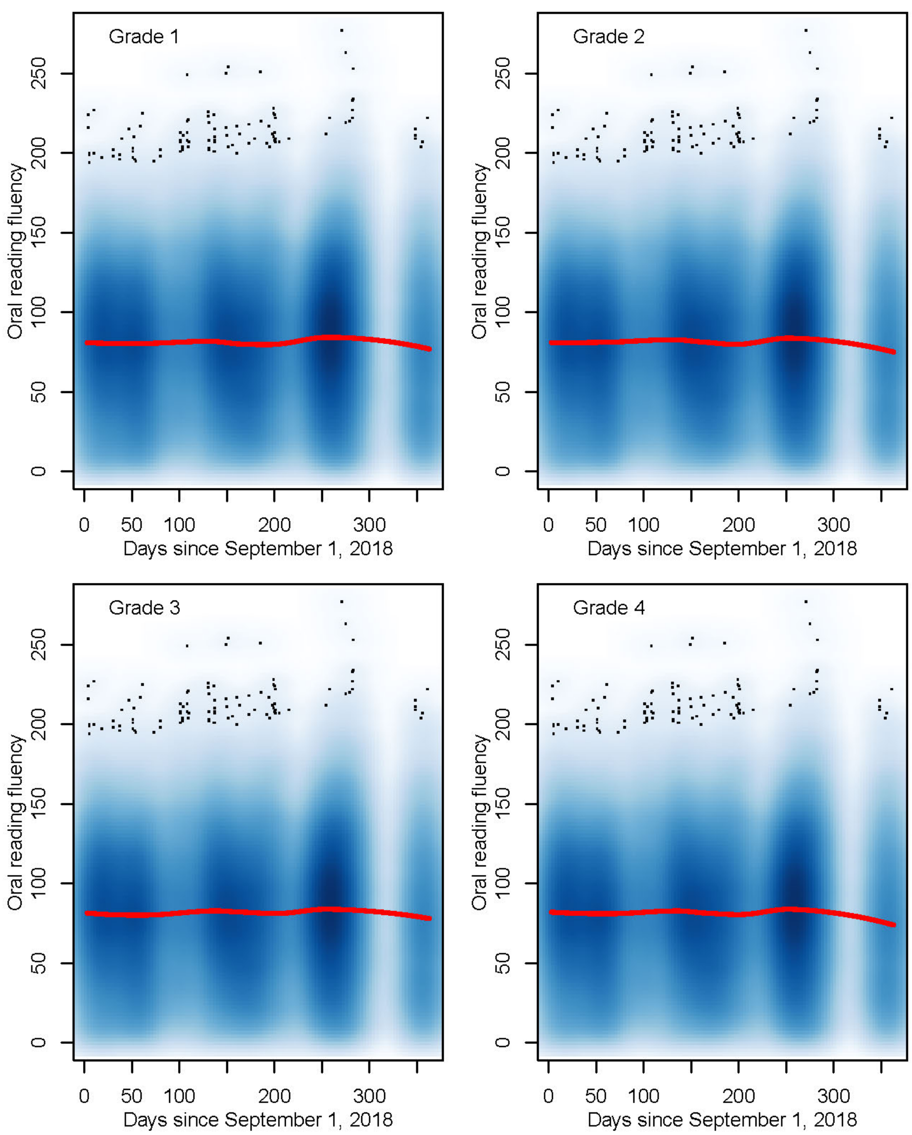
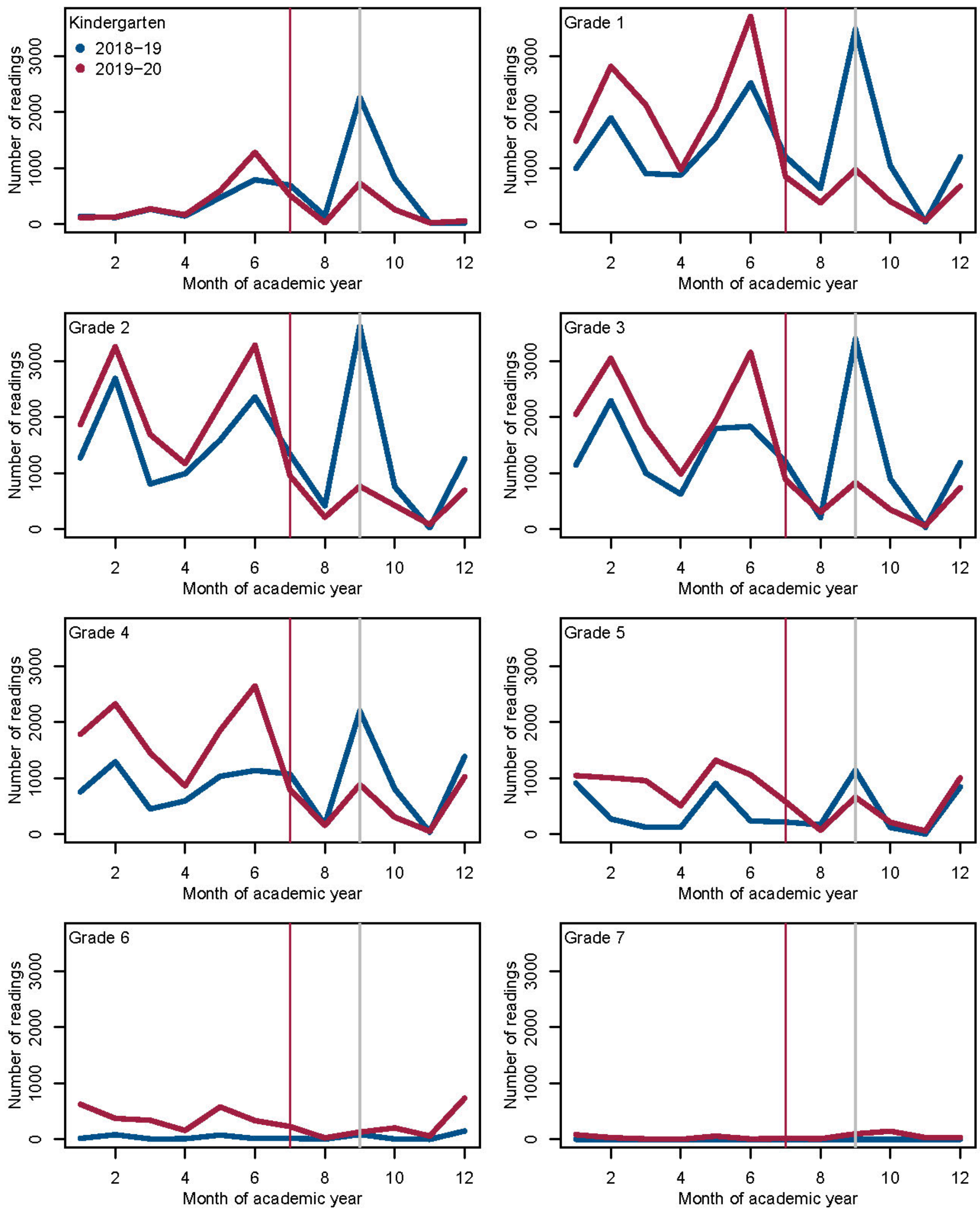


Figure A3. Number of Scores per Month in 2018–19 and 2019–20.



Note. The decline in the number of scores in 2019–20 following the outbreak of the COVID-19 pandemic is notable. We include Grades 5–7 so as to illustrate the decrease in data collected in those grades.

Analyzing Growth

Our approach to analyzing growth is based on identifying expected changes in ORF score as a function of time after controlling for person and book fixed effects. Specialized approaches are needed to estimate models with large numbers of fixed effects. Here, we use the R package `lfe` (Gaure, 2013). To estimate growth in ORF for a given cohort, we first calculate the time since some point (i.e., the start of the school year, which we assume is September 1 of a given academic year). For student i, we denote the time of the *j*-th observation of that student as t_{ij} and we call the score y_{ij}. In some cases, we treat time linearly,

$$y_{ij} \sim \mathrm{Normal}\left(\beta_t t_{ij} + \gamma_i + \lambda_b, \sigma^2\right). \text{ [1]}$$

So as to exclude individual-specific performance differences and differences specific to texts, we also include fixed effects for person, γ_i, and text, λ_b. Thus, estimates of β_t tell us the expected growth in ORF per day for a student.

In other cases, we allow for nonlinear effects of time. To account for potential nonlinearities in student growth in ORF—in particular, we assume that there might be highly nonlinear trends in the 2019–20 and following due to COVID-19—we use B-splines (Friedman et al., 2001). As used here, B-splines are a map from from R^1 to R^K where *K* is specified by the user (for figures in the main text, we use *K* = 7 for Figure 2 and *K* = 5 for Figure 3). We then model score *j* for individual *i* when reading text *b* as

$$y_{ij} \sim \mathrm{Normal}\left(\sum_k \beta_k B(t_{ij})_k + \gamma_i + \lambda_b, \sigma^2\right). \text{ [2]}$$

In such cases, we do not focus on estimates β_k but instead examine fitted trajectories of growth based on $\sum^k \hat{\beta_k} B(t)$ for some appropriate choice of t.

We also consider a parametric analysis based on piecewise linear models. This model allows for different linear growth when we split the time frame of interest into different blocks, denoted with T:

$$y_{ij} \sim \mathrm{Normal}\left(\sum_T (\delta_T + \beta_T t_{ij}) + \gamma_i + \lambda_b, \sigma^2\right). \text{ [3]}$$

Interest will be in comparison of the β_T coefficients both across time blocks T within a cohort (e.g., comparing fall and spring ORF growth for the 2019–20 cohort) and within-T across cohort (e.g., comparing spring 2019 and spring 2020 growth) as well as construction of fitted growth trajectories.

A Focus on Selection

As emphasized throughout, students observed during the pandemic are not a random subsample of students observed in, for example, fall 2019. Below we consider three analyses meant to further assess differences between observed and unobserved students. In Figure A4, we examine pre-COVID means for those observed and not observed in spring 2020 when missingness was most pronounced. We then turn to analyses of growth (Figures A5 and A6).

Figure A4. Means in Fall 2019 for Those Students Observed and Not Observed in Spring 2020.

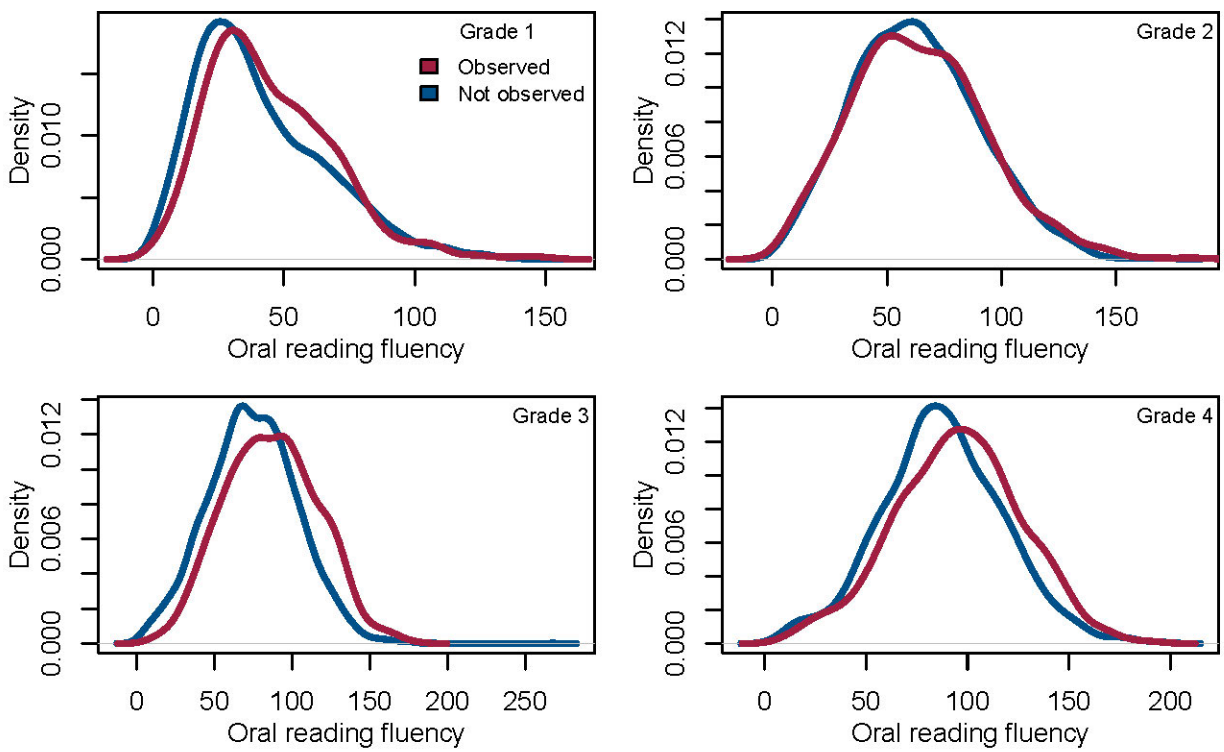
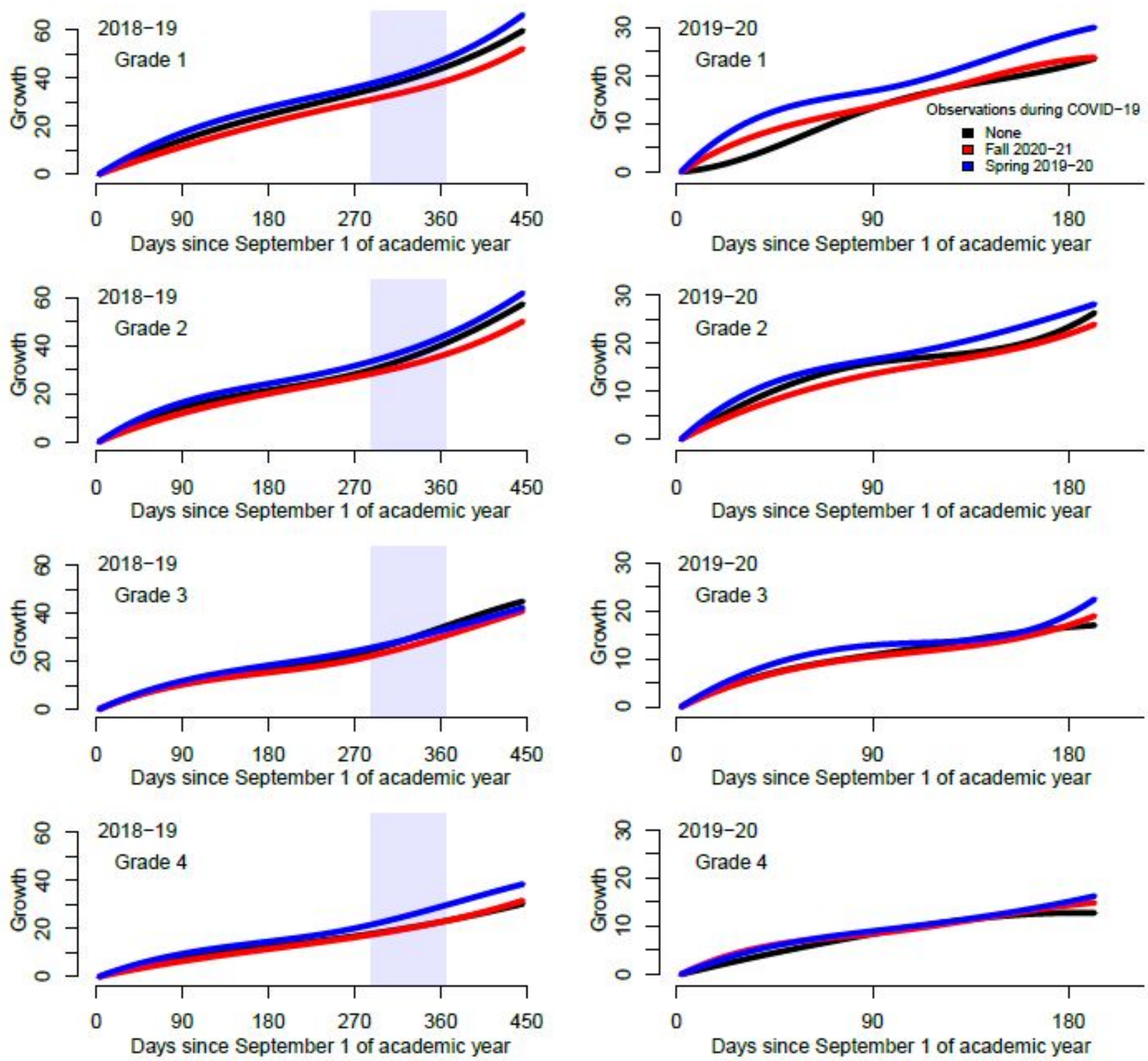
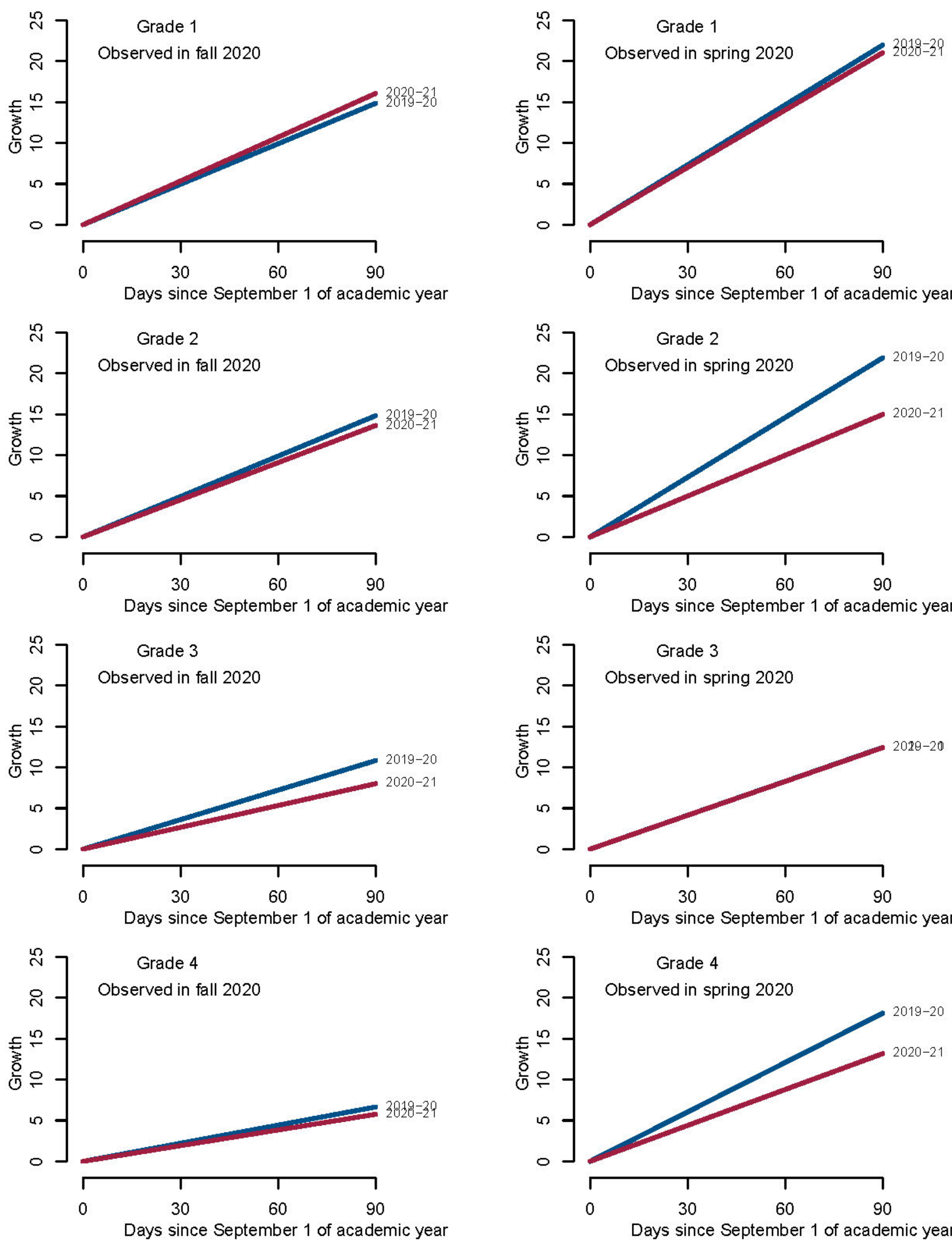


Figure A5. Comparison of Pre-Pandemic Growth.



Note. Growth in 2018–19 & pre-COVID in 2019–20 for students with different patterns of observed cores during the pandemic.

Figure A6. Comparison of Post-Pandemic Growth During Fall 2020.



Note. Graphs on the left indicate analysis of respondents observed during fall 2020 (but not spring 2020). Graphs on the right indicate analysis of respondents observed during spring 2020.

We first consider differences in means for students who do and do not provide scores in spring 2020, see Figure A4. Observed students tend to be somewhat higher-achieving than unobserved students (c.f., Grade 2). However, differences are relatively modest. Given this, we now turn to analyses of growth as a function of missingness.

We next consider differences in pre-COVID growth as a function of COVID-related missingness. For students for whom we have data in fall 2019, we divide them into three categories: those who provide no data post-COVID onset, those who provide data in fall 2020, and those who provide data in spring 2020. Figure A5 compares growth amongst these three groups *focusing on the pre-COVID period* when they are all observed. Students observed in spring 2020 tend to show the most growth. These trends suggest, perhaps running counter to the results from Figure A6, that selection is occurring for a non-random set of students.

We now ask about the level of growth for students in fall 2020 as a function of whether they were observed in spring 2020. Figure A6 shows linear growth trends for 2019–20 and 2020–21 students *during the COVID-19 period* (note that we do not consider the third group from Figure A5 since clearly they are not observed during the pandemic). We break students with 2019–20 scores (N = 38,016) into two groups: those observed during the pandemic spring 2020 (right, N = 9,893) and those only observed in fall 2020 (left, N = 12,804).[7] Focusing first on growth rates across the columns, students observed in spring 2020 typically grow faster than those observed only in 2020–21 (i.e., lines tend to be steeper in the panels on right). This is consistent with results from Figure A5. In Grade 1, returns to growth were equal across groups. In Grades 2 and 4, those not observed in spring 2020 seemed to show more return to pre-existing growth patterns than did those observed in spring 2020. In Grade 3, this pattern was reversed.

Moderation Analyses

We estimated several moderation models. To do so, we extended Equation 1 to include an interaction term,

$$y_{ij} \sim \mathrm{Normal}\left(\beta_1 t_{ij} + \beta_2 t_{ij} M_{ij} + \gamma_i + \lambda_b, \sigma^2\right).$$

for some moderator M_{ij}. As moderators, we consider the school district's mean test score achievement (SEDA) and socioeconomic status (SES) using the SEDA data (Reardon et al., 2019); these moderators were standardized across the sample of districts. For those students that had data available, we also considered differences in growth as a function of prior-year mean ORF

[7] A third group, N = 15,319, were not observed post-COVID onset. This group consisted of more older students than the others; for example, 46% of this group was in Grades 5–6 while only 5% and 16% of the other two groups were composed of students from these grades.

(for scores collected pre-COVID in the case of 2020–21); we standardize prior-year ORF means across the analytic dataset.

Estimates are in Table A2. As discussed in the main text, there is some evidence for novel differences when we compare the growth of low- and high-achieving districts in 2020–21 as compared to the previous year. We similarly observe growth differences as a function of district SES in Grades 1–2 in 2020–21. However, pre-COVID growth trends as a function of district SES in 2019–20 are more complex. Turning to the analyses based on the prior year ORF mean, first note the substantial loss of sample. This is due to the increase in data collected by Literably over time. In 2020–21, we observe no evidence for pronounced differences in growth as a function of prior-year ORF mean.

Table A2. Estimates for Moderation Models Based on Fall Oral Reading Fluency Estimates.

		2019–20						2020–21					
Moderator	Grade	Beta_1	*SE*	Beta_2	*SE*	*N* score	*N* student	Beta_1	*SE*	Beta_2	*SE*	*N* score	*N* student
SEDA	1	0.196	0.010	0.043	0.011	7492	3347	0.179	0.005	0.032	0.006	18831	8095
SEDA	2	0.211	0.010	0.009	0.010	8425	4008	0.161	0.006	0.035	0.007	19805	9175
SEDA	3	0.124	0.011	-0.011	0.011	8849	4740	0.120	0.007	0.019	0.008	19062	9797
SEDA	4	0.102	0.012	0.017	0.011	7206	4343	0.083	0.008	0.014	0.009	16479	9786
SES	1	0.196	0.011	0.041	0.017	6697	3017	0.168	0.006	0.041	0.008	18359	7835
SES	2	0.202	0.010	0.022	0.014	7367	3598	0.153	0.006	0.028	0.008	19193	8818
SES	3	0.130	0.011	-0.048	0.015	8035	4347	0.117	0.007	0.007	0.009	18446	9458
SES	4	0.072	0.015	0.072	0.017	6306	3934	0.079	0.009	0.011	0.012	15774	9429
ORF	1	0.223	0.031	0.013	0.025	4475	2164	0.159	0.025	-0.020	0.019	3161	1301
ORF	2	0.142	0.017	-0.053	0.016	6590	3173	0.138	0.016	-0.018	0.015	6010	2847
ORF	3	0.104	0.013	-0.020	0.017	6520	3516	0.090	0.012	-0.013	0.014	6037	2904
ORF	4	0.085	0.014	0.009	0.016	6088	3666	0.074	0.016	-0.009	0.021	5067	3158

Table A3. Coefficient Estimates for Piecewise Linear Growth Models for Spring 2020.

Grade	Academic year	Pre-break growth	*SE*	Intercept at break	*SE*	Post-break growth	*SE*	*p* value, test of difference in post-break growth between 2018-19 and 2019-20	*N* student
1	2018–19	0.154	0.007	28.126	1.299	0.110	0.019	0.149	3252
	2019–20	0.136	0.005	25.375	1.293	0.068	0.022		4777
2	2018–19	0.124	0.006	23.602	1.028	0.051	0.015	0.029	4616
	2019–20	0.122	0.005	21.888	1.310	-0.010	0.023		4369
3	2018–19	0.100	0.007	18.561	1.218	0.055	0.020	0.017	3370
	2019–20	0.092	0.005	16.937	1.241	-0.011	0.020		4688
4	2018–19	0.073	0.008	13.651	1.338	0.027	0.019	0.837	2905
	2019–20	0.074	0.005	13.274	1.695	0.034	0.026		4304

Note. This table refers to Figure 4 in main text.

References

Baker, S. K., Smolkowski, K., Katz, R., Fien, H., Seeley, J. R., Kame'Enui, E. J., & Thomas Beck, C. (2008). Reading fluency as a predictor of reading proficiency in low-performing, high-poverty schools. *School Psychology Review 37*(1), 18–37. doi.org/10.1080/02796015.2008.12087905

Bielinski, J., Brown, R., & Wagner, K. (2020). *COVID slide: Research on learning loss & recommendations to close the gap.* Illuminate Education. https://www.illuminateed.com/download/covid-slide-research-on-learning-loss-recommendations-to-close-the-gap/

Cleveland, W. S., Grosse, E., & Shyu, W. M. (1992). Local regression models. In *Statistical Models in S*, J. M. Chambers & T. J. Hastie (Eds.). Wadsworth & Brooks/Cole.

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020). *COVID-19 and student learning in the United States: The hurt could last a lifetime.* McKinsey & Company. https://webtest.childrensinstitute.net/sites/default/files/documents/COVID-19-and-student-learning-in-the-United-States_FINAL.pdf

Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning*. Springer. https://web.stanford.edu/~hastie/ElemStatLearn/

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading 5*(3), 239–56. https://doi.org/10.1207.S1532799XSSR0503_3

Gaure, S. (2013). OLS with multiple high dimensional category variables. *Computational Statistics & Data Analysis 66,* 8–18. doi.org/10.1016/j.csda.2013.03.024

Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher 59*(7), 636–44. doi.org/10.1598/RT.59.7.3

Herold, B. (2020). The disparities in remote learning under coronavirus (in charts). *Education Week.* www. Edweek.org/ew/articles/2020/04/10/the-Disparities-in-Remote-Learning-Under-Coronavirus.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association 81*(396), 945–60. https://www.tandfonline.com/doi/abs/10.1080/01621459.1986.10478354

Huff, K. (2020). *National data quantifies impact of COVID learning loss; raises questions about at-home testing.* Curriculum Associates. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/ca-impact-of-covid-learning-loss-fall-2020.pdf

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher 49*(8), 549–65. https://doi:10.3102/0013189X20965918.

Kuhfeld, M., Tarasawa, B., Johnson, A., Ruzek, E., & Lewis, K. (2020). *Learning during COVID-19: Initial findings on students' reading and math achievement and growth.* NWEA Research. https://www.nwea.org/content/uploads/2020/11/Collaborative-brief-Learning-during-COVID-19.NOV2020.pdf

Literably, Inc. (2018). *Literably technical manual*. https://s3.amazonaws.com/literably-assets/literably_technical_manual_latest.pdf.

May, H. (2020). The literacy skills of Utah first graders have dropped statewide. *The Salt Lake Tribune*. www.sltrib.com/news/education/2020/12/11/literacy-skills-utah/.

Parolin, Z., & Lee, E. K. (2020). *Large socio-economic, geographic, and demographic disparities exist in exposure to school closures and distance learning.* OSF Preprints. https://osf.io/cr6gq

Pier, L., Hough, H. J., Christian, M., Bookman, N., Wilkenfeld, B., & Miller, R. (2020). *COVID-19 and the educational equity crisis: Evidence on learning loss from the CORE Data Collaborative.* https://edpolicyinca.org/newsroom/covid-19-and-educational-equity-crisis

Reardon, S. F., Shear, B. R., Castellano, K. E., & Ho, A. D. (2017). Using heteroskedastic ordered probit models to recover moments of continuous test score distributions from coarsened data. *Journal of Educational and Behavioral Statistics 42*(1), 3–45. doi.org/10.3102/1076998616666279

Reardon, S. F., Ho, A. D., Shear, B. R., Fahle, E. M., Kaolgides, D., Jang, H., Chavez, B., Buontempo, J., & DiSalvo, R. (2019). *Stanford education data archive* (Version 3.0). http://purl.stanford.edu/db586ns4974.

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

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology 47*(6), 427–69. https://doi.org/10.1016/j.jsp.2009.07.001

Townsend, J. B., & Domingue, B. W. (2018). Predictive validity of Literably Oral Reading Fluency. https://s3.amazonaws.com/literably-assets/literably_validity_report_1.pdf.