K-8 Pandemic Learning Trends: Technical Appendix

Data Description

The analyses were conducted using three main data sources, all of which were prepared by and accessed in partnership with Curriculum Associates (CA).

The first two sets of data contain anonymized individual-level information for the universe of i-Ready Diagnostic users who took the reading or math assessment from SY2016-17 to SY2018-19 and from SY2020-21 to SY2021-22 and were enrolled in grades 1 through 8 during the sample period. We omit records from SY2019-20 due to data limitations.[1] CA's i-Ready Diagnostic tests are computer-delivered formative assessments typically administered three times per school year (fall, winter, and spring quarters). Unlike summative assessments (such as mandated state tests), the i-Ready Diagnostic is a "low-stakes" evaluation intended to serve as a metric of student progress that school leaders and teachers can use to inform practice.[2] The variables used in the analyses include scale scores, on-grade placement levels, growth measures, and location of diagnostic assessment.

We merged individual student diagnostic records with school-level information for the universe of i-Ready schools who administered at least one of the i-Ready diagnostics (reading and math) during the school years 2017-18, 2018-19, and 2020-21. These data were obtained by CA from the National Center for Education Statistics (NCES). The variables used in the analyses include demographic characteristics such as the percent of students by race/ethnicity and sex, and income-related measures such as the percent of students who qualified for free or reduced-price lunch, and indicators for Title I status and funding.

Analysis Sample

The analyses were conducted using spring i-Ready Diagnostic scores and were limited to schools with nonmissing demographic and income-related variables for the school years 2017-18 and/or 2018-19 – that is, at least one of the years preceding the initial pandemic-related school closures for which we have school demographic records. This selection criteria restricted the sample to the set of schools who were i-Ready Diagnostic users in at least one subject (reading and math) and one year prior to SY2019-20, and it was necessary to define school-level characteristics as time-invariant measures capturing averages over pre-pandemic school years. The latter was done to avoid confounding differences in achievement by school types (e.g., defined by the percent of students of color) with enrollment changes following the COVID-19 pandemic (Dee et al., 2021; Musaddiq et al., 2022).

---

[1] Although a small portion of schools did complete assessments in spring 2020, we do not include these data in our analysis because of the inability to generalize these findings to the nation as a whole.

[2] i-Ready assessments are typically positively correlated with "high-stakes" tests such as mandated state exams, with stronger correlations in elementary grades. See Curriculum Associates (2022) for an example of a correlation study.

The sample was further restricted to students who took the spring i-Ready Diagnostic in a school setting. In an analysis of over 900 schools using i-Ready, CA reported a positive correlation between in-home test taking and diagnostic scores in the fall of SY2020-21, which they speculated to be related to the possibility that students "received extra support from parents and others" (Curriculum Associates, 2020). Other comparable studies likewise find large score improvements among at-home test takers suggesting a pattern across a broader set of assessment types (see e.g., Kuhfeld et al. 2020). Considering the findings from multiple national analyses, we omit students who reported taking the i-Ready Diagnostic at home. Approximately 11 percent of observations in the full sample reported testing at home and were therefore dropped from the analyses. Note that this sample selection condition does not imply that instruction took place in-person, it simply indicates that the assessments were administered in a school environment. All observations in the sample prior to SY219-20 were assumed to take the diagnostic in a school setting.

Lastly, following Goldhaber et al. (2022), we restricted the sample to schools that had nonmissing test scores for at least 10 students in a grade-year combination. This condition was used to minimize the number of schools that might use i-Ready as a small-scale intervention rather than a school-wide diagnostic tool.

After applying these sample conditions, the final analysis data consisted of two cross-sectional datasets (one for each subject: reading and math) covering students in grades 1 through 8 who took the spring i-Ready Diagnostic assessment at least once over the sample period. Table A1 summarizes the number of students, schools, and observations in each analysis sample.

Table A1: Number of Schools, Students, and Total Observations
by Grade Levels and Subject (2017-2022)

	Grade Level	Number of Schools[3]	Number of Students[4]	Total Observations
Reading Sample	Grades 1 - 2	14,968	3,939,538	4,870,116
	Grades 3 - 5	16,591	5,613,595	7,760,235
	Grades 6 - 8	10,343	3,931,613	5,045,750
Math Sample	Grades 1 - 2	16,654	4,437,862	5,464,162
	Grades 3 - 5	18,162	6,166,190	8,541,554
	Grades 6 - 8	11,427	4,316,415	5,543,367

[3] The schools in the reading and math samples overlapped substantially -- 80.97 percent in grades 1 and 2; 81.87 percent in grades 3 through 5; and 81.42 percent in grades 6 through 8.
[4] The students in the reading and math samples also overlapped substantially (although at a lower rate than schools in the reading sample) – 72.09 percent in grades 1 and 2; 75 percent in grades 3 through 5; 66.73 percent in grades 6 through 8.

Variable Definitions

Learning Outcomes
Our main outcome variables are spring i-Ready Diagnostic scores in reading and math. We use these as our preferred measure of student performance to account for a full year of learning after the initial COVID-related school disruptions. Therefore, the first post-pandemic measure of learning outcomes corresponds to the end of the 2020-21 school year when nearly 30 percent of states in the country required in-person instruction.[5]

School Characteristics
We use school-level demographic and income-related information from SY2017-18 and SY2018-19 to construct time-invariant variables of the average share of students of color by school and indicators for whether schools were ever identified as Title I over the sample period. Using the distribution of non-white students by school, we computed the 75th and 25th percentiles and used these cutoffs to classify schools as serving the highest and lowest proportions of students of color, respectively.

- Majority non-white schools: Schools that served a high proportion of students of color were identified as such if they were schools within the top quartile of the share of non-white students.

- Majority white schools: Schools that served the lowest proportions of students of color were identified as such if they were schools within the bottom quartile of the share of non-white students.

- Low-income schools: Schools were identified as "low-income" if they were designated as Title I schools for at least one year from SY2017-18 to SY2018-19.

- High-income schools: Schools were identified as "high-income" if they were never designated as Title I schools in any year from SY2017-18 to SY2018-19.

Descriptive Statistics

Tables A2 and A3 below show summary statistics for the schools included in the reading and math analysis samples, respectively. Each table contains information for the racial/ethnic composition of the school, a host of income-related variables (e.g., median household income and percent of students eligible for free or reduced-price lunch), location of the school, and the percent of students who took i-Ready in a school setting. Column 1 shows the school averages for the set of students in first and second grades. Column 2 shows the school

[5] https://www.edweek.org/leadership/map-where-are-schools-closed/2020/07

averages for the set of students in third through fifth grades. Lastly, column 3 shows the school averages for the set of students in sixth through eighth grades.

On average, students in the reading sample attended schools where the percent of white students ranged between 41 – 44 percent. Hispanic students were the second-largest demographic group accounting for roughly 29 percent of students in the sample followed by Black students (16-18 percent) and Asian students (approximately 5 percent). Schools in the sample enrolled on average 57-60 percent of students who qualify for free or reduced-price lunch and 52-66 percent were designated as Title I. Lastly, nearly 30 percent of the schools were located in urban settings. The descriptive statistics of the schools used in the analysis of the math trends are shown in Table A3 and are similar to those for reading.

Table A2: Descriptive Statistics Reading Sample Grades 1 – 8

VARIABLE NAME	Grades 1-2	Grades 3-5	Grades 6-8
Pct white	41.70	41.45	43.91
Pct Black	18.54	18.18	16.99
Pct Asian	5.42	5.39	5.29
Pct Hispanic[*]	28.46	29.22	28.69
Pct Non-white	58.02	58.25	55.77
Pct FRL	58.67	59.07	56.98
Share Title I Status	0.66	0.66	0.52
Share Urban	0.28	0.29	0.27
Median Household Income	68,213	68,165	68,635
Pct In-School Testing	89.93	89.67	85.93

Table A3: Descriptive Statistics Math Sample Grades 1 – 8

VARIABLE NAME	Grades 1-2	Grades 3-5	Grades 6-8
Pct white	43.39	43.15	44.85
Pct Black	17.83	17.69	17.37
Pct Asian	5.14	5.06	4.91
Pct Hispanic[*]	27.76	28.32	27.70
Pct Non-white	56.22	56.46	54.66
Pct FRL	58.14	58.45	57.18
Share Title I Status	0.66	0.66	0.52
Share Urban	0.29	0.29	0.29
Median Household Income	68,007	67,930	67,900
Pct In-School Testing	89.89	90.02	86.16

[*] Here, we use the term "Hispanic" to be consistent with the terminology used in the NCES data.
[*] Here, we use the term "Hispanic" to be consistent with the terminology used in the NCES data.

Because we use cross-sectional data, the set of schools included in our reading and math samples were not constant over the span of the years referenced in the dataset. Therefore, comparing pre- and post-pandemic scores, as we do in the briefs, do not allow one to fully disentangle observed changes over time from changes in the underlying composition of schools (Gewertz, 2021), meaning that the score changes reported might be the result of a changing sample rather than the impact of COVID alone. We begin to explore the extent to which school composition changed over the sample period by comparing descriptive statistics estimated before (SY2017-18 and SY2018-19) and after (SY 2020-21) the pandemic began. Results are reported in Table A4.

Table A4: Descriptive Statistics Before and After the Pandemic (Reading Sample)

VARIABLE NAME	GRADES 1-2		GRADES 3-5		GRADES 6-8	
	Pre SY2020	Post SY2020	Pre SY2020	Post SY2020	Pre SY2020	Post SY2020
Pct white	40.06	40.52	40.61	40.86	41.99	41.25
Pct Black	21.94	19.38	22.06	19.40	21.27	18.71
Pct Asian	4.78	5.65	4.57	5.52	5.13	5.53
Pct Hispanic*	27.09	27.92	26.72	27.77	26.08	28.43
Pct Non-white	59.31	58.70	58.65	58.37	57.21	57.88
Pct FRL	62.56	60.96	62.59	61.04	60.58	61.35
Share Title I Status	0.68	0.66	0.68	0.65	0.58	0.55
Median Household Income	64,746	67,390	64,517	67,231	65,605	66,289

While there are some differences in the demographic and socioeconomic characteristics of the sample schools observed in the pre- and post-COVID years, the differences are small and would underestimate our findings. The largest differences (such as the decrease in the percent of non-white students in some grade groups, decrease percent of FRL-eligible students in the early and upper elementary grades, and increase in median household income in all grade groups) are typically correlated with higher achievement, meaning that our post-COVID averages are likely a conservative estimate of the effects of the pandemic on student learning .

Lastly, in Table A5 we report descriptive statistics for the reading and math samples, pooled across grades 1 through 8, and compare them to those measured at the national level to gauge the extent to which the samples used in this study are representative of the schools in the country as a whole. We limit our comparison to measures of racial/ethnic composition and

* Here, we use the term "Hispanic" to be consistent with the terminology used in the NCES data.

income. Overall, the schools represented in the sample of i-Ready Diagnostic users are more likely to enroll non-white students compared to schools nationwide – approximately 57 percent vs. 50 percent, respectively. Schools in the sample of i-Ready Diagnostic users are also more likely to serve students who are eligible for free or reduced-price lunch, although they are less likely to be identified as Title I schools.

Table A5: Descriptive Statistics by Subject and NCES Samples (Grades 1-8)

VARIABLE NAME	NCES	Reading Sample	Math Sample
Pct white	49.26	42.22	43.70
Pct Black	14.62	17.94	17.64
Pct Asian	4.22	5.37	5.04
Pct Hispanic*	25.57	28.86	27.99
Pct Non-white	50.43	57.48	55.88
Pct FRL	48.52	58.36	58.01
Share Title I Status	0.79	0.62	0.62

Methodology

We first conducted a descriptive analysis of reading and math achievement trends using cross-sectional datasets of students in first through eighth grade from SY2016-17 to SY2021-2022. For each grade, we generated a line graph of actual scores over time against the score ranges deemed to represent "on-grade" performance using Curriculum Associate's designation of placement levels in SY2019 (Curriculum Associates, 2019).[6] Together, these descriptive analyses allow for an accurate interpretation of the observed changes in scores relative to grade-level benchmarks that make appropriate distinctions across grades and subjects.

We quantified changes in actual achievement within grade-levels and across time, with a particular focus on pre- vs. post-COVID years in the sample. These differences were described in raw scale score points and in "weeks-of-learning" equivalents. We translated the differences in scores into estimated equivalent weeks of learning to quantify the changes in meaningful terms. For purposes of this brief series, we calculated the estimated weeks of learning associated with each point difference by first identifying the pre-pandemic grade- and subject-specific annual growth targets set by Curriculum Associates for SY2017.[7] For each grade, we divided the expected growth by a 30-week school year to obtain an estimate of

* Here, we use the term "Hispanic" to be consistent with the terminology used in the NCES data.
[6] We selected a pre-COVID designation of placement levels to refer to achievement expectations in a "typical" year.
[7] We selected annual growth targets for a pre-COVID year in order to measure achievement expectations under "typical" circumstances. Since SY2017, Curriculum Associates has updated their growth targets to vary based on students' baseline achievement.

how much a student would have to improve their scores to “stay on track”. Using this rough estimate, we converted the point difference in scores (e.g., between 2021 and 2019) into the equivalent number of weeks.[8]

Our last set of descriptive analyses compared reading and math score trends across school types defined by the share of students of color and Title I status. Each bar graph depicts the raw scale score gap across schools, thus indicating the difference in achievement between schools with high and low proportions of students of color, as well as non-Title I and Title I schools. Raw scale score gaps were also translated into weeks of learning equivalents for ease of interpretation.

A note on the use of weeks of learning

There are limitations to this type of estimate. For example, past scholars have noted that a model like this assumes learning is linear and is experienced the same way across all grade levels (Baird & Pane, 2019). This can be particularly problematic if grades and subjects are pooled to generate the weeks of learning (Kraft, 2020). Furthermore, this estimate has the risk of being misleading since the calculation reflects (though it doesn’t account for) factors inside and outside of the school, including developmental variables (Kraft, 2020).

For our purposes, we have opted to use this translation despite its limitations. Using weeks of learning as a metric helps start the important conversation between practitioners, policymakers, and researchers to inform the actions that must be taken to best support the country’s youth. Other scholars have used similar methods and have articulated their usefulness. We attempt to ameliorate some of the concerns discussed above by using growth targets that are specific to each respective grade and subject, without pooling grades for analyzing the trends across time. In the absence of proper control groups, we also use benchmarks set before the onset of the pandemic that reflect what “typical” growth could have occurred had it not been for the disruptions associated with the pandemic. In using weeks of learning, we can begin to concretize the severity of the impacts the pandemic had on student learning.

[8] For example, the learning growth target in math for a second-grade student, before the pandemic, was set to be 27 points. Dividing the expected learning growth target (27 points) by a 30-week school year, we estimate that a student would improve their math score by 0.9 points each week. Taking the difference in average second-grade math scores between 2021 and 2019 (419 – 429.1 = -10.1 points) and dividing that by the 0.9 average growth per week, we find that this -10.1 point difference in math scores is associated with the equivalent of about 11 weeks of learning.

References

Baird, M. D., & Pane, J. F. (2019). Translating Standardized Effects of Education Programs Into More Interpretable Metrics. Educational Researcher, 48(4), 217–228.

Curriculum Associates. (2019). I-Ready Assessments Technical Manual: March 2018.

Curriculum Associates. (2020). National Data Quantifies Impact of COVID Learning Loss: Raises Questions About At-Home Testing. First Look. Retrieved from: www.curriculumassociates.com/-/media/mainsite/files/i-ready/caimpact-of-covid-learning-loss-fall-2020.pdf

Curriculum Associates. (2022). i-Ready Diagnostic: Correlation Study with the State of Texas Assessments of Academic Readiness (STAAR). Retrieved from: https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-diagnostic-assessments-linking-study-overview-tx-2020.pdf

Dee, T. S., Huffaker, E., Phillips, C., & Sagara, E. (2021). The revealed preferences for school reopening: Evidence from public-school disenrollment. American Educational Research Journal, 00028312221140029.

Gewertz, C. (2021, October 21). State test results are in. Are they useless? Education Week. https://www.edweek.org/teaching-learning/state-test-results-are-in-are-they-useless/2021/10

Goldhaber, D., Kane, T. J., McEachin, A., Morton, E., Patterson, T., & Staiger, D. O. (2022). The consequences of remote and hybrid instruction during the pandemic (No. w30010). National Bureau of Economic Research.

Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. Educational Researcher, 49(4), 241–253.

Kuhfeld, M., Lewis, K., Meyer, P., & Tarasawa, B. 2020a. Comparability analysis of remote and in-person MAP Growth testing in fall 2020. NWEA. Retrieved from: www.nwea.org/research/publication/comparabilityanalysis-of-remote-and-in-person-map-growth-testing-in-fall-2020/

Musaddiq, T., Stange, K., Bacher-Hicks, A., & Goodman, J. (2022). The pandemic's effect on demand for public schools, homeschooling, and private schools. Journal of Public Economics, 212, 104710.