

Understanding changes in academic achievement in Pittsburgh Public Schools during the COVID-19 pandemic: Methodology memo

Whitney Kozakowski, Brian Gill, Patrick Lavalley, Alyson Burnett,
and Jonathan Ladinsky

August 2021



This memo was funded by the U.S. Department of Education’s Institute of Education Sciences (IES) under contract ED-IES-17-C-0006 by Regional Educational Laboratory Mid-Atlantic administered by Mathematica. The content does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

NWEA MAP test scores. The NWEA MAP test score files include test scores for math and reading in grade levels K-12. NWEA MAP tests are offered in fall, winter, and spring in PPS. At the time the data were collected in early spring 2021, we received fall, winter, and spring test scores for 2019/20 and fall and winter scores for 2020/21.

Because the NWEA MAP tests are vertically aligned, students across grades receive scores on the same scale. We use these scale scores for some analyses to identify how much students have grown in terms of scale score points over the last year. We also calculate standardized scale scores, standardizing within subject, grade level, and year relative to the national means and standard deviations reported by NWEA from a norming study conducted before the pandemic (Thum & Kuhfeld, 2020). Standardizing scores in each year relative to the same pre-pandemic benchmark allows us to examine changes in student achievement relative to pre-pandemic national norms.

We also use the standardized scores to create indicator variables to assign students to quartiles of the national distribution using the fall 2019/20 math and reading scores. These indicators identify student subgroups based on baseline test scores.

Table 1. Student characteristics and outcomes used in the study

Characteristic	Description
Demographic and behavior variables	
School year	Academic year (either 2019/20 or 2020/21)
Grade level (fall)	Which grade level (K-12) the student was at the time of NWEA fall MAP testing
Grade level (winter)	Which grade level (K-12) the student was at the time of NWEA winter MAP testing
Grade level (spring)	Which grade level (K-12) the student was at the time of NWEA spring MAP testing
Female	Whether a student was female
Race/ethnicity	Whether a student was <ul style="list-style-type: none"> • Black • White
Economically disadvantaged	Whether student directly certified for the national school lunch program
IEP (Individualized Education Program)	Whether student received special education services
Chronic Absentee	Whether a student was absent for more than 10 percent of the first semester (quarter 1 and quarter 2)
Grades	
Indicator for failing any courses/subjects	Indicator for failing any graded course/subject (received an E or F) in the first semester
Grade point average (GPA)	Average grade in graded courses/subjects that students took in the first semester. Courses that were pass/fail or ungraded were not included in GPA.
NWEA MAP tests	
Math scale score fall	NWEA MAP scale score for math in the fall
Math standardized (std.) scale score fall	NWEA MAP scale score for math in the fall, standardized within grade level relative to national norms
Reading scale score fall	NWEA MAP scale score for reading in the fall
Reading std. scale score fall	NWEA MAP scale score for reading in the fall, standardized within grade level relative to national norms
Math scale score winter	NWEA MAP scale score for math in the winter
Math standardized (std.) scale score winter	NWEA MAP scale score for math in the winter, standardized within grade level relative to national norms
Reading scale score winter	NWEA MAP scale score for reading in the winter
Reading std. scale score winter	NWEA MAP scale score for reading in the winter, standardized within grade level relative to national norms
Math scale score spring	NWEA MAP scale score for math in the spring
Math standardized (std.) scale score spring	NWEA MAP scale score for math in the spring, standardized within grade level relative to national norms
Reading scale score spring	NWEA MAP scale score for reading in the spring

Characteristic	Description
Reading std. scale score spring	NWEA MAP scale score for reading in the spring, standardized within grade level relative to national norms
Quartile 1 of std. math scale score fall	Indicator that standardized math scale score in fall is below the 25th percentile
Quartile 2 of std. math scale score fall	Indicator that standardized math scale score in fall is in the 25th percentile and below the median
Quartile 3 of std. math scale score fall	Indicator that standardized math scale score in fall is in the 50th percentile and below the 75th percentile
Quartile 4 of std. math scale score fall	Indicator that standardized math scale score in fall is in the 75th percentile and above
Quartile 1 of std. reading scale score fall	Indicator that standardized reading scale score in fall is below the 25th percentile
Quartile 2 of std. reading scale score fall	Indicator that standardized reading scale score in fall is in the 25th percentile and below the median
Quartile 3 of std. reading scale score fall	Indicator that standardized reading scale score in fall is in the 50th percentile and below the 75th percentile
Quartile 4 of std. reading scale score fall	Indicator that standardized reading scale score in fall is in the 75th percentile and above

Table 2. Descriptive statistics

Characteristic	2019/20	2020/21
Demographic and behavior variables (Grades K-12)		
Female	0.49	0.49
Black	0.52	0.52
White	0.32	0.31
Economically disadvantaged	0.70	0.70
IEP (Individualized Education Program)	0.22	0.22
Chronic Absentee	0.25	0.25
Grades (Grades 1-12)		
Indicator for failing any courses/subjects	0.11	0.18
Grade point average (GPA)	2.96	2.75
Number of courses failed (First semester)	0.29	0.57
NWEA MAP tests (Grades 2-8)		
Math scale score fall	200.2	201.6
Math standardized (std.) scale score fall	-0.25	-0.19
Reading scale score fall	196.1	197.8
Reading std. scale score fall	-0.24	-0.17
Math scale score winter	204.6	203.9
Math standardized (std.) scale score winter	-0.34	-0.40
Reading scale score winter	198.3	199.0
Reading std. scale score winter	-0.42	-0.39
Math scale score spring	218.4	—
Math standardized (std.) scale score spring	-0.18	—
Reading scale score spring	208.2	—
Reading std. scale score spring	-0.31	—
Quartile 1 of std. math scale score fall	0.33	0.32
Quartile 2 of std. math scale score fall	0.26	0.26
Quartile 3 of std. math scale score fall	0.21	0.21
Quartile 4 of std. math scale score fall	0.20	0.21
Quartile 1 of std. reading scale score fall	0.32	0.33
Quartile 2 of std. reading scale score fall	0.25	0.21

Characteristic	2019/20	2020/21
Quartile 3 of std. reading scale score fall	0.22	0.21
Quartile 4 of std. reading scale score fall	0.21	0.25

Notes: In the first section, the sample includes all students enrolled in PPS in 2019/20 (n= 21,819) or 2020/21 (n=20,630). In the second section, the sample includes all students enrolled in PPS in 2019/20 in grades 1-12 who have grades in the first semester (n=19,390) or the equivalent for 2020/21 (n=18,716). In the third section, the sample includes all students enrolled in PPS in 2019/20 in grades 2-8 who took the fall, winter, or spring tests, respectively, or the equivalent for 2020/21.

Sample

The sample sizes differed by analysis (table 3).

Table 3. Sample sizes by research question and analysis sample for the analyses presented in the slides

Research question (RQ)	Analysis sample	Sample size
RQ A		
RQ A. Proportion of students taking the NWEA MAP tests	Students enrolled in grade levels K–12 during the testing window for each assessment window (fall, winter, spring)	Fall 2019/20: 21,819 Winter 2019/20: 21,394 Spring 2019/20: 22,445 Fall 2020/21: 20,630 Winter 2020/21: 20,529
RQ A. Changes in the composition of students taking the test (grades 2–8)	All students enrolled in grade levels 2–8 during the testing window for each assessment window (fall, winter, spring)	Fall 2019/20: 11,727 Winter 2019/20: 11,522 Spring 2019/20: 12,073 Fall 2020/21: 11,229 Winter 2020/21: 11,163
RQ A. Proportion of PPS test-takers in each national quartile of the fall 2019/20 reading test distribution who take the fall reading test in 2019/20 and 2020/21	All students enrolled in grade levels 2–7 who took the fall 2019/20 reading test	Took test in fall 2019/20: 8,486 Took test also in fall 2020/21: 7,328
RQ A. Proportion of students with grades	All students enrolled in grade levels K–12 during the fall testing window for 2019/20 and 2020/21	2019/20: 21,819 2020/21: 20,630
RQ B		
RQ B. Change in individual student math scores (winter-to-winter)	Students who were enrolled in grade levels 2–7 in 2019/20 who took the winter math test in 2019/20 and 2020/21	7,517
RQ B. Change in individual student math scores (winter-to-winter, imputed)	Students who were enrolled in grade levels 2–7 in 2019/20 who took the winter math test in 2019/20 and who had grade and absence data in 2020/21.	8,420
RQ B. Change in individual student reading scores (winter-to-winter)	Students who were enrolled in grade levels 2–7 in 2019/20 who took the winter reading test in 2019/20 and 2020/21.	7,366
RQ B. Change in individual student reading scores (winter-to-winter, imputed)	Students who were enrolled in grade levels 2–7 in 2019/20 who took the winter reading test in 2019/20 and who had grade and absence data in 2020/21.	8,338
RQ B. Change in individual student math scores compared to NWEA study during the same period	Students who were enrolled in grade levels 3–7 in 2019/20 who took the winter math test in 2019/20 and fall math test in 2020/21	6,423
RQ B. Change in individual student reading scores compared to NWEA study during the same period of time	Students who were enrolled in grade levels 3–7 in 2019/20 who took the winter reading test in 2019/20 and fall reading test in 2020/21	6,194
RQ B. Change in grade outcomes (GPA, whether a student failed a course)	Students who were enrolled in grade levels 1–12 in 2019/20 and 2020/21 with first semester grades.	2019/20: 19,390 2020/21: 18,716
RQ B. Distribution of first-semester grades	All grades received in instructional courses	High school:

Research question (RQ)	Analysis sample	Sample size
in 2019/20 vs. 2020/21	in grade levels 6–12	47,741 grades in 2019/20 46,786 grades in 2020/21 Middle school: 34,641 grades in 2019/20 33,622 grades in 2020/21
RQ B. Change in proportion of students chronically absent	Students who were enrolled in grade levels K–12 in 2019/20 and 2020/21 with absence data	2019/20: 21,806 2020/21: 20,628
RQ B. Average Days Absent by Number of Courses Failed	Students who were enrolled in grade levels 1–12 in 2019/20 and 2020/21 with grades and absence records	19,390 students in 2019/20 18,715 students in 2020/21

Source: Authors' samples based on administrative data and online learning application data provided by Pittsburgh Public Schools, 2019/20 and 2020/21.

Analysis methods

This study is focused on understanding how academic achievement, as measured by test scores and grades, changed during the pandemic. One challenge in this analysis is that the pandemic may have disrupted the number of students tested in the 2020/21 school year or the number of grades submitted. As a result, taking the average student performance during the baseline period (before school closures in 2019/20) and comparing it to 2020/21 may be misleading.

To understand the scope of this problem, we first address research question A.

First, we describe changes in the proportion of students enrolled in each grade level in PPS during the testing window for each test (fall, winter, and spring of 2019/20 and fall and winter of 2020/21) who take NWEA math and reading tests.

Second, to understand whether the sample of students taking the tests are representative of the enrolled students, we calculate the proportion of students who have a characteristic (e.g., who are female) in the enrolled student population and the tested student population. We then calculate the difference between these two and determine if the standardized difference exceeds 0.05 standard deviations. For this analysis, we focus on students in grade levels 2–8, as these are the tested grade levels that we use for the analysis of the change in test scores. We examine differences in the proportion of students who are female, Black, White, economically disadvantaged, or have an individualized education program (IEP). (We do not include English learners as a group because they represent less than 5 percent of students in PPS in 2020/21. Similarly, we do not include other non-Black minority groups because 83 percent of students enrolled in 2020/21 are either Black or White and no other race/ethnicity groups exceed 10 percent of the enrolled population).

Third, we examine changes in the demographics of students who took the NWEA MAP tests in 2020/21 compared to 2019/20 to understand if students who took the test in both periods had higher or lower test scores in 2019/20. To do so, we use a sample of students in grade levels 2–7 who took the test in fall 2019/20 and assign them to quartiles based on national norms. We then report the proportion of students who took the test for each quartile in fall 2019/20 and those who also took the test in fall 2020/21.

Next, we address research question B. From research question A, even for grade levels 2–8 which have the highest rates of test participation, we still observe a small decline in the proportion of students taking the test and some shifts in the characteristics of students taking the test, although these changes did not exceed 0.05 standard deviations. To address these challenges, we conduct a longitudinal analysis comparing individual students' performance to their own performance in a prior period. For example, we compare a student's score in winter 2020/21 to a period in 2019/20, specifically fall 2020/21 or winter 2020/21. (We focus on winter 2019/20 to winter 2020/21, except for an analysis comparing the changes in PPS against another study (Kuhfeld et al., 2020a) which

reported differences from winter 2019/20 to fall 2020/21). This has the benefit of holding constant the sample with test scores in 2019/20 and 2020/21. However, one drawback of this approach is that we cannot examine the changes in test scores for students who do not take the test in 2020/21.

To address this drawback, we also conduct a sensitivity check in which we impute scores for those with scores winter 2019/20 who do not have them in winter 2020/21. To do so, for students within grade levels 2–7 in 2019/20, we regress the Winter 2020/21 score for a given subject on the winter 2019/20 score for that subject, first semester 2020/21 GPA, the number of course failures in the first semester of 2020/21, and the number of absences in the first semester 2020/21. We also include indicator variables for the following demographic characteristics: gender, race/ethnicity, economically disadvantaged, IEP status, and English learner status. To improve model fit, we include second and third order polynomials for the winter 2019/20 score, GPA, number of courses failed, and absences. We run regressions separately for students in each grade level. We then impute the predicted winter 2020/21 score for those who are missing winter 2020/21 scores. In total, 90.5 percent of students with winter 2019/20 reading scores and 90.4 percent of students with winter 2019/20 math scores were enrolled in winter 2020/21. Among these students, 12 percent (915 students in math, 985 in reading) who took the winter 2019/20 math test did not take the winter 2020/21 test. We impute scores for all but 14 of these students who were missing course grades. We did not attempt to impute scores for students no longer enrolled in PPS. In total, we have real or imputed scores for 90.4 percent of those with winter math scores in 2019/20 and 90.4 percent of those with winter reading scores in 2019/20 (Students with actual scores in Winter 2020/21 constitute 80.1 percent of those with scores in the preceding winter, and 89.2 percent of those who had scores in the preceding winter *and* were enrolled in PPS in winter 2020/21).

Imputing scores allows us to include almost all students still enrolled in the district during the winter 2020/21 testing window. Additionally, one advantage of using variables based on grades and absences in 2020/21 for the imputation, as opposed to just prior test score, is that we are able to use information about students from the 2020/21 school year to inform the imputation. However, all imputation has error and may over- or under-estimate the test scores for those who did not take the test. In this case, imputed scores may over-estimate true scores, given that many students who were failing courses in the first semester of 2020/21 did not take the MAP test in 2020/21 (table 4).

Table 4. Proportion of students in grade levels 2–7 in 2019/20 who have a winter 2020 math score, by the number of courses failed during the first semester of 2020/21

Number of courses failed in first semester of fall 2020	Proportion with a winter 2020 math score	Frequency
0	0.89	7,802
1	0.72	522
2	0.64	213
3	0.49	155
4	0.39	110
5	0.30	63
6	0.38	39
7	0.40	5
8 or more	0.29	9
Total		8,918

Source: Sample includes all students in grade levels 2–7 in 2019/20. Authors’ samples based on administrative data provided by Pittsburgh Public Schools, 2020/21.

To interpret test score findings, we use a threshold of 0.1 standard deviations to identify meaningful changes in standardized test scores. For scale score changes, we use a threshold in scale score points that is equivalent to 0.1 standard deviations in the pre-pandemic national distribution (For example, for grade 2 in math, this would be 1.3

scale score points). We use these thresholds when discussing how individual students' scores changed from 2019/20 to 2020/21 and when discussing average differences in the changes students experienced in different groups (e.g., male versus female students). Because the analyses are not intended to generalize to a larger population, we do not conduct tests of the statistical significance of differences.

In contrast to test scores, the proportion of students with outcomes based on grades (specifically, whether they fail a course and GPA) is high in all grades but kindergarten and does not change from 2019/20 to 2020/21 (see slide 13 in the attached deck). We therefore conduct a cross-sectional analysis for grade-based outcomes because we are not concerned about the composition of the sample changing in ways that might bias the results from 2019/20 to 2020/21. Specifically, we compare successive cohorts in the same grade levels or subgroups, calculating the difference in outcomes based on first-semester grades from 2019/20 and 2020/21. The cross-sectional analysis is preferable to a longitudinal analysis because there are increases or declines in course failures that typically occur at some grade level transitions (for example, students moving from grade 8 to grade 9 often experience an increase in course failures). A longitudinal analysis would conflate the effects of the pandemic with these natural changes that occur across some grade-level transitions, while the cross-sectional analysis avoids this problem by holding grade level constant.

To interpret findings when discussing changes in the proportion of students failing at least one course, we use a threshold of 5 percentage points to identify meaningful changes. When interpreting changes in GPAs, we use a threshold of 0.1 GPA points.

Limitations

The test score analysis has a few limitations. First, the analysis sample was limited to students in grade levels 2–7 in 2019/20 who took the test in both testing periods. Students who took the test in 2019/20 but not in 2020/21 would not be included in the main analysis. It is possible students who did not take the test in 2020/21 would have different learning gains than those who did take the test. Also, the analysis sample was restricted to grade levels 2–7 because a lower proportion of students took the NWEA MAP tests in earlier grade levels (kindergarten and grade 1) or late grade levels (grades 8–12) in both 2019/20 and 2020/21. It is possible the findings could differ in these grades. Second, the tests were administered remotely in the 2020/21 school year for both fall and winter. Test scores in remote environments were found by NWEA to be reliable in grade levels 3–8, but should be used with caution in earlier grades (Kuhfeld et al., 2020b). Third, MAP tests were first offered in 2019/20 in Pittsburgh. As a result, part of the change in scores from 2019/20 to 2020/21 may be due to students and teachers becoming more familiar with the test. This may have helped mitigate some of the declines relative to pre-pandemic national norms that PPS students experienced. Fourth, for the sensitivity check that uses imputation, imputation is not perfect and may under- or over-estimate the scores for students who had scores in 2019/20 but did not take the test in 2020/21. In particular, it is possible imputed scores were overly optimistic given that many students who were failing courses in the first semester of 2020/21 did not take the MAP test in 2020/21.

The analysis of grades also faces a limitation. Criteria for failing a course may have shifted during the pandemic. To the extent that teachers lowered grading standards, the change in course failure rates we calculate would understate what the change would have been had the failure criteria stayed constant.

References

- Kuhfeld, M., Tarasawa, B., Johnson, A., Ruzek, E., & Lewis, K. (2020a). Learning during COVID-19: Initial findings on students' reading and math achievement and growth. NWEA. <https://www.nwea.org/research/publication/learning-during-covid-19-initial-findings-on-students-reading-and-math-achievement-and-growth/>
- Kuhfeld, M., Lewis, K., Meyer, P., & Tarasawa, B. (2020b). *Comparability analysis of remote and in-person MAP Growth testing in fall 2020*. NWEA. <https://www.nwea.org/research/publication/comparability-analysis-of-remote-and-in-person-map-growth-testing-in-fall-2020/>

Thum, Y. M., & Kuhfeld, M. (2020). NWEA 2020 MAP Growth Achievement Status and Growth Norms Tables for Students and Schools. NWEA. <https://teach.mapnwea.org/impl/NormsTables.pdf>