Michigan's 2020-21 and 2021-22 Benchmark Assessments

OCTOBER 2022

Education Policy Innovation Collaborative
COLLEGE OF EDUCATION | MICHIGAN STATE UNIVERSITY

236 Erickson Hall, 620 Farm Lane, East Lansing, MI 48824 | (517) 884-0377 | www.EPICedpolicy.org

DISCLAIMER

The Education Policy Innovation Collaborative (EPIC) at Michigan State University is an independent, non-partisan research center that operates as the strategic research partner to the Michigan Department of Education (MDE) and the Center for Educational Performance and Information (CEPI). EPIC conducts original research using a variety of methods that include advanced statistical modeling, representative surveys, interviews, and case study approaches.

This research used data structured and maintained by the MERI-Michigan Education Data Center (MEDC). MEDC data are modified for analysis purposes using rules governed by MEDC and are not identical to those data collected and maintained by the MDE and/or CEPI. Results, information, and opinions solely represent the author(s)

TABLE OF CONTENTS

Executive Summary

BACKGROUND AND PURPOSE

In order to monitor students' progress toward learning goals in the wake of the COVID-19 pandemic, the Michigan legislature mandated benchmark assessment testing for all K-8 students in both fall and spring semesters of the 2020-21, 2021-22, and 2022-23 school years (2020 PA 149, 2021 PA 48). To help provide context to student outcomes during this period, the Education Policy Innovation Collaborative (EPIC) is releasing a series of reports in partnership with the Michigan Department of Education (MDE), the Center for Educational Performance and Information (CEPI), the Michigan Data Hub (MDH), and the Michigan Education Data Center (MEDC). This is the fourth report in the series, all of which will be delivered to the governor and the Senate and House standing committees responsible for education legislation in the Michigan legislature to provide insight into Michigan students' progress toward learning goals during the COVID-19 pandemic.

The first and second reports found that a substantial proportion of students in 2020-21 experienced slower rates of achievement growth on benchmark assessments relative to a typical school year—especially Black, Latino, economically disadvantaged, and special education students. The third report showed that, although the majority of Michigan students demonstrated at least some achievement growth from fall 2020 to fall 2021, they were less likely than students pre-pandemic to achieve a full year of growth. Students with access to some in-person instruction in 2020-21 were less affected than those whose districts offered only fully remote instruction, and achievement gaps between these two groups of students grew during the 2020-21 school year before improving slightly over summer 2021.

RESEARCH QUESTIONS AND STUDY OVERVIEW

In this report, we use K-8 math and reading benchmark assessment data from the 2020-21 and 2021-22 school years to investigate each of the following questions:

1. **How did achievement trajectories of Michigan students throughout the 2020-21 and 2021-22 school years compare to national or state trends from pre-pandemic?** To better understand how much the COVID-19

pandemic has affected Michigan students, we examine how average scores on state-mandated benchmark assessments for K-8 students have progressed during the 2020-21 and 2021-22 school years. We then compare students' trajectories to national and Michigan-specific norms from pre-pandemic.

2. **How did student achievement growth in 2020-21 and 2021-22 compare to typical year-to-year growth before the COVID-19 pandemic?** We compare students' growth on fall and spring benchmark assessments from 2020-21 and 2021-22 to national and state norms that each assessment provider established before the COVID-19 pandemic for students in the same grade level and subject with similar initial achievement scores.

3. **How did achievement growth differ across subgroups of students?** We compare patterns in student achievement growth across demographic subgroups as well as subgroups of students whose districts offered different modes of instruction (i.e., in-person, hybrid, or remote) in 2020-21 and students who received different modes of instruction in 2021-22.

Our analyses include benchmark assessment results from approximately 730,000 of Michigan's 935,000 K-8 students in 725 of Michigan's 848 school districts. While these analyses help to deepen our understanding of how Michigan public school students progressed and learned between fall 2020 and spring 2022, they are based on imperfect and incomplete data. For instance, prior research has shown that the COVID-19 pandemic has had a greater negative effect on achievement and achievement growth for students who are Black, Latino, economically disadvantaged, or English learners, and these same student populations are underrepresented in our analysis.

Michigan's benchmark assessment legislation allows districts to choose an appropriate assessment from one of four MDE-approved providers, and thousands of students participated in assessments from each of the four. NWEA's MAP Growth assessment was Michigan's most frequently used benchmark assessment; more than 600 districts participated in an NWEA MAP Growth assessment in spring 2022. Sixty-seven districts administered Curriculum Associates' i-Ready Diagnostic assessments, 75 administered Renaissance Learning's Star 360 assessments, and 28 administered one of DRC's assessments. These assessments are all designed in slightly different ways, cover slightly different content, and tend to appeal to different types of districts. However, as we discuss in detail in later in this report, we find several common themes in the results from different assessments.

KEY FINDINGS

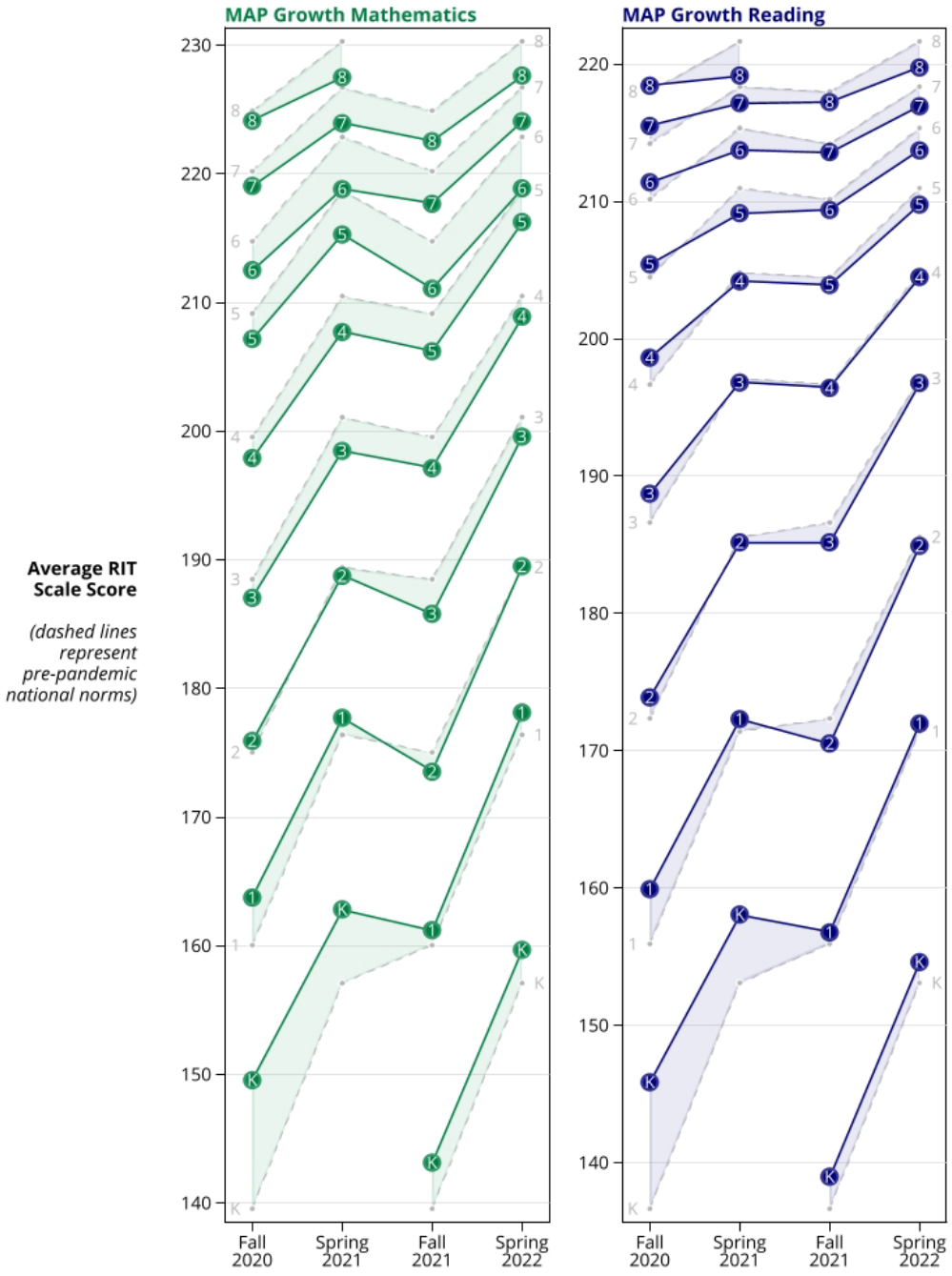
On Average, Students in 2021-22 Started the School Year Behind and Ended the School Year Ahead of Students Who Were in the Same Grade Levels in 2020-21

Even though students in fall 2021 often started the school year behind students from the prior cohort, they experienced greater achievement growth during the 2021-22 school year relative to pre-pandemic national norms. While the 2020-21 cohort generally saw decreases in relative performance between the fall and spring testing periods, the 2021-22 cohort maintained or improved their percentile ranks and ultimately surpassed 2020-21 achievement levels in most grade levels by spring 2022.

Students Experienced More Growth in 2021-22, but Often Not Enough to Counteract the Effects of Unfinished Learning in 2020-21

Figure 1 shows trends in average math and reading scores for Michigan students who took the MAP Growth math and reading assessments in fall 2020, spring 2021, fall 2021, and spring 2022. Average math and reading scores increased between fall 2021 and spring 2022, with elementary student growth generally outpacing typical changes established by the pre-pandemic norms. As a result, spring 2022 average math and reading scores for students enrolled in kindergarten in fall 2020 remained above pre-pandemic norms, while students in the 1st- through 4th-grade cohorts generally drew closer to their respective pre-pandemic norms. Average math and reading score growth among students in the 5th- through 7th-grade cohorts trailed typical changes established by the pre-pandemic norms and students in these grade levels fell further behind. Overall, fewer Michigan students scored above pre-pandemic national averages in 2021-22 than in 2020-21.

Figure 1. Trends in Average Scale Scores, NWEA MAP Growth

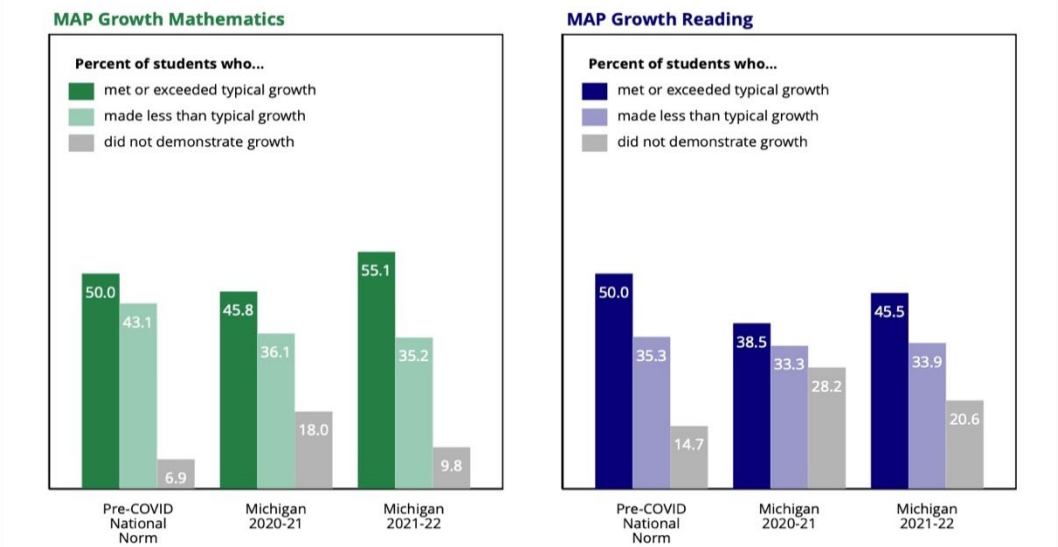


Note: These averages include only students with benchmark assessment scores for every possible testing period. The comparison points in the figure represent the 50th percentile of NWEA's conditional growth distribution. The y-axis scales range from the kindergarten fall norm to the 8th grade spring norm; these values are slightly different across subjects. RIT stands for Rasch unit scale.

Many Students Did Not Demonstrate Achievement Growth in 2021-22 and Most of These Students are Far Below Grade-Level Proficiency

Across subjects, grade levels, and assessment providers, the percentages of students who met or exceeded a typical year's growth increased, the percentages who did not demonstrate any growth at all decreased, and students who made less than a typical year's growth made more progress, on average, toward their growth targets. However, while student growth improved, on average relative to 2020-21, there were still far more students who did not demonstrate any growth at all in 2021-22 than in the pre-pandemic national norming samples. Figure 2 shows that, on the MAP Growth assessments, about 10% and 21% of students did not demonstrate growth in math and reading, respectively, in 2021-22. While these percentages are far lower than in 2020-21, they are still higher than prior to the COVID-19 pandemic, when 7% and 15% of students did not demonstrate growth in math and reading. Most of these students scored at a level assessment providers equate to "not proficient" (i.e., the lowest proficiency level) on the M-STEP, suggesting that the many of the lowest-performing students are those who made no progress on benchmark assessments during the 2021-22 school year.

Figure 2. Percent of Students Achieving Typical Growth, NWEA MAP Growth



Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. We base "typical growth" on NWEA's conditional growth distribution for each grade level, subject, and initial achievement level (Thum & Kuhfeld, 2020).

Of Students Who Made Less Than Typical Growth, Average Learning Gains Were Greater in 2021-22 Than in 2020-21

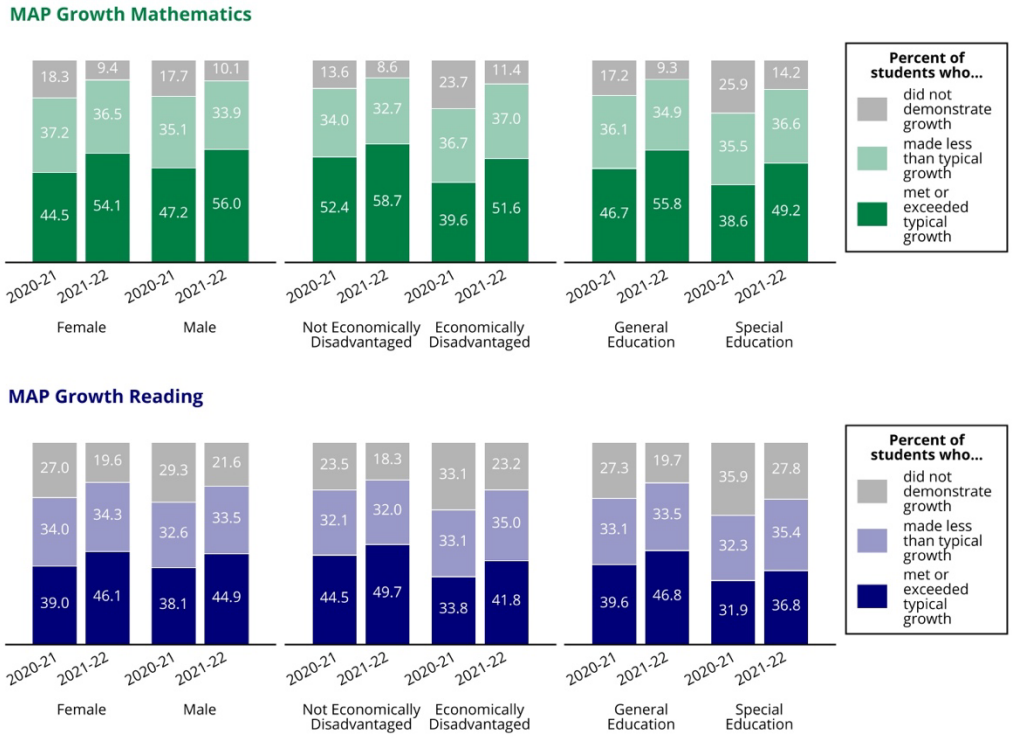
In both 2020-21 and 2021-22, the average "partial growth" student achieved more than 50% of typical yearly growth in math and reading, meaning that they learned more than half of a typical year's worth of tested content. Average achievement for these students improved over time, with students achieving a higher percentage of typical math and reading growth in 2021-22 compared to the previous school year. For instance, students on average learned 62% of the year's math content tested on the NWEA MAP Growth assessment in 2021-22, relative to 59% the year prior. In reading, students on average learned 59% of the year's content relative to 56% in 2020-21.

This change is driven by a shift in the distributions of partial growth that is common across both subjects. We see that more students made less than 50% of typical growth in each subject during the 2020-21 school year whereas more students made more than 50% of typical growth in 2021-22.

While Economically Disadvantaged, Special Education, Black, and Latino Students Remained Less Likely to Achieve a Typical Year's Growth, Disparities Decreased Over Time

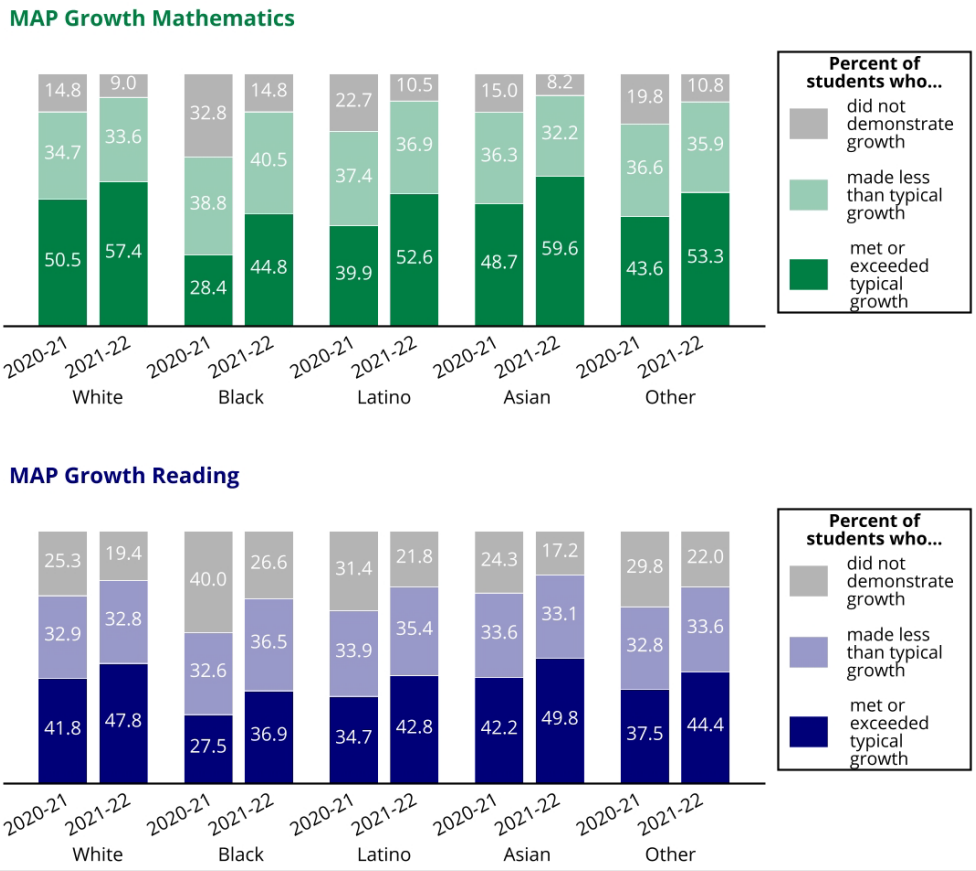
Compared to their peers, economically disadvantaged students and students with disabilities were less likely to achieve a typical year's growth in either subject in both 2020-21 and 2021-22. However, Figure 3 shows that growth rates for both economically disadvantaged students and students with disabilities improved at a faster rate compared to their counterparts, and the disparities between these groups decreased from 2020-21 to 2021-22. Similarly, Figure 4 shows that Black and Latino students were less likely to achieve typical growth in either subject or school year than White or Asian students. However, growth rates for Black and Latino students improved more than those of White and Asian students, and as a result, the disparities shrunk considerably by spring 2022.

Figure 3. Yearly Achievement Growth by Subgroup, NWEA MAP Growth



Note: These percentages only include students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. We base "typical growth" on NWEA's conditional growth distribution for each grade level, subject, and initial achievement level (Thum & Kuhfeld, 2020).

Figure 4. Yearly Achievement Growth by Race/Ethnicity, NWEA MAP Growth



Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. We base "typical growth" on NWEA's conditional growth distribution for each grade level, subject, and initial achievement level (Thum & Kuhfeld, 2020).

Students with Less Access to In-Person Instruction in 2020-21 Experienced Less Achievement Growth but Gaps Shrunk Substantially Once Most Districts Returned to In-Person Learning in 2021-22

Students whose districts offered in-person instruction for only part of the 2020-21 school year made clear improvements in both math and reading once in-person instruction was more readily available in 2021-22. Students in districts that did not offer in-person instruction at any time during the 2020-21 school year were the least likely to achieve a typical year's growth and the most likely to demonstrate no growth

in either year, though performance gaps shrunk substantially once most districts returned to in-person learning in 2021-22. Improvements in growth outcomes between 2020-21 to 2021-22 were consistently larger for students who received in-person instruction in 2021-22.

SUMMARY

This report shows that, following more than two years of unprecedented disruptions to learning and schooling, student achievement trends in Michigan show early signs of progress but still have a long way to go. Michigan students, on average, started the 2021-22 school year below students who completed the same grade the year before. However, 2021-22 student learning over the course of the year outpaced students in the cohort before them, and ultimately surpassed the 2020-21 cohort's achievement levels by the end of the year. Even so, there are substantial numbers of Michigan students who are falling further behind rather than catching up.

Results from our previous report showed that substantial portions of students did not demonstrate typical growth in math and reading in 2020-21. Outcomes for the most recent school year are not as bleak. Across all grades, math and reading growth improved as students were more likely to demonstrate growth, more likely to make substantial progress toward yearly growth targets, and more likely to achieve a full year's growth in 2021-22 than in 2020-21. For many students, these improvements were not enough to offset the effects of unfinished learning during the COVID-19 pandemic. In particular, students with less access to in-person instruction in 2020-21 experienced less achievement growth than those with more access, though these gaps began to diminish after most districts returned to in-person learning in 2021-22. Similarly, demographic gaps in achievement growth shrunk considerably between 2020-21 and 2021-22.

All these results must be placed in the context of the imperfect data available to analyze student learning growth during the COVID-19 pandemic. Not only were participation rates lower than for a typical end-of-year summative assessment, but the resulting analytic samples are not entirely reflective of Michigan's larger student population. In addition, concerns about differences in early elementary students' at-home testing environments in fall 2020 make it difficult to assess growth and achievement trajectories for young cohorts of students. Moreover, the data on instructional modalities offered by school districts in the 2020-21 school year reflect only what the districts themselves offered not how individual students learned. Previous work in Michigan and elsewhere has shown that districts retained local autonomy over the choice of modality offerings, and parents and families were almost always able to make their own choices to learn remotely if they chose to do so.

Nonetheless, the results presented herein provide important information for policymakers, educators, and stakeholders as we continue to grapple with the academic effects of the COVID-19 pandemic on Michigan's students.

Section One: Introduction

On August 20, 2020, Michigan Governor Gretchen Whitmer signed a series of three "Return to Learn" bills intended to provide districts with flexibility to adapt their programs as necessary to safely provide instruction during the COVID-19 pandemic (2020 PA 147, 148, 149). For the 2020-21 school year only, the state legislature waived many instructional requirements, including minimum number of days and hours, what learning activities count toward the attendance, and enrollment calculations that determine state aid allocations. Along with this increased flexibility, the Return to Learn legislation outlined a new set of requirements for the 2020-21 school year to ensure that districts continued to adequately meet students' needs without the same instructional requirements in place.

As a condition for receiving state aid for the year, the legislation required each district to develop an extended COVID-19 learning plan describing how it would deliver instruction and establishing educational goals for the 2020-21 school year. These educational goals were to include increased student achievement or growth as measured using one or more benchmark assessments, overall and for all subgroups of students. Districts were required to select and administer appropriate benchmark assessments to all K-8 students at the beginning and end of the school year to determine whether students made meaningful progress toward mastery of state standards in reading and mathematics.

The Return to Learn legislation allowed districts to choose one of four state-approved benchmark assessments or one or more benchmark assessments that contain progress monitoring and enhanced diagnostics in reading or math. Alternately or in addition, districts were allowed to use a locally developed benchmark assessment that met the same requirements. While the legislation prohibited the use of these data for accountability purposes, districts that elected to use an approved provider's benchmark assessment were required to compile and report their results through the Michigan Datahub network for use in a statewide aggregate report for the governor and the House and Senate standing committees responsible for education legislation in the Michigan legislature.

To continue tracking student learning and progress toward educational goals during the COVID-19 pandemic, the Michigan legislature again passed legislation in summer 2021 that provided districts with the same flexibility to choose among benchmark assessments and required districts to administer those assessments to all K-8 students in both fall and spring semesters during the 2021-22 and 2022-23 school years (2021 PA 48).

This report is the fourth in a series that the Education Policy Innovation Collaborative (EPIC) at Michigan State University has provided to the Michigan Department of Education (MDE), the governor, and the House and Senate standing committees responsible for education legislation to give insight into Michigan students' progress toward learning goals in the wake of the COVID-19 pandemic. EPIC prepared this report in collaboration with MDE, the Center for Educational Performance and Information (CEPI), the Michigan Data Hub (MDH), and the Michigan Education Data Center (MEDC) at the University of Michigan as a summary of the student academic growth across the 2020-21 and 2021-22 school years.

ACHIEVEMENT IN OTHER STATES BETWEEN FALL 2020 AND SPRING 2022

There is mounting evidence that students across the country missed important opportunities to learn over the past two school years. Results from spring 2022 standardized testing across the country, however, show that reading achievement in many states is beginning to return to pre-pandemic levels while math achievement is still "catching up."

For example, in Tennessee, slightly more than a third of elementary, middle, and high school students scored proficient on the spring 2022 ELA standardized assessment. The scores for each grade span all matched or exceeded pre-pandemic achievement levels, and the elementary and high school scores both reflect the highest proficiency levels in the past five years. Math proficiency levels have yet to recover, though proficiency gains across all grade levels closed 30 to 50% of the learning gaps created by school closures at the beginning of the COVID-19 pandemic (Tennessee Department of Education, 2022). The state education agencies in Florida, Idaho, Indiana, Ohio, and Texas have all reported similar results that show students have made progress academically during the 2021-22 school year but many students in each state still trail pre-pandemic achievement levels (Appleton, 2022; Greater Fort Lauderdale Allliance, 2022; Kogan, 2022; Texas Education Agency, 2022; Tennessee Department of Education, 2022; Idaho State Department of Education, 2022).

A study summarizing aggregate national achievement among students who completed an NWEA assessment shows that the national recovery from the COVID-19 pandemic has been more tepid (Kuhfeld & Lewis, 2022). Overall, learning rates between 2020-21 and 2021-22 generally mirror pre-pandemic achievement trends, and in some cases achievement growth has exceeded those from a typical school year where a subset of students in 3rd to 8th grade were able to make up as much as a quarter to a third of the unfinished learning experienced throughout school closures and remote instruction over the past two school years. However, even if academic improvements continue at the rates seen during the 2021-22 school year, it may be years before students experience a full recovery; Kuhfeld and Lewis estimate that grade-specific achievement gaps created by the COVID-19 pandemic may not fully close for three to five years in elementary school grades while middle school gaps may need at least five years to return to pre-pandemic achievement levels.

PURPOSE OF THIS REPORT

The first two reports in this series, released in August of 2021 and January of 2022, found that a substantial proportion of students experienced slower rates of achievement growth on benchmark assessments compared to a typical school year—especially Black, Latino, economically disadvantaged, and special education students. The third report (April, 2022) documented that, although the majority of Michigan students demonstrated at least some achievement growth from fall 2020 to fall 2021, only about 40% reached their growth targets over the same time period. Further, students with access to some in-person instruction in 2020-21 were less affected than those whose districts offered only fully remote instruction, and achievement gaps between these two groups of students grew during the 2020-21 school year before improving slightly over summer 2021.

This fourth report extends our assessment of student progress toward learning goals through the end of the 2021-22 school year. Specifically, in this analysis, we examine achievement trajectories and growth over the last four semesters—from fall 2020 to spring 2022—and assess differences in performance across subgroups of students with different demographic characteristics and who participated in different modes of instruction (e.g., fully in-person, fully remote, or hybrid instruction) in 2021-22.

In what follows, we first discuss the data and methods we use in this report. Section Three provides results from our analyses. We conclude in Section Four with a brief discussion of the implications of our findings for Michigan K-12 education as we start the 2022-23 school year.

Section Two: Data and Methods

Each year, millions of K-12 students across the country participate in benchmark assessments. Benchmark assessments are designed to help educators and administrators track students' progress toward grade-level standards and learning goals, and to provide feedback to help drive future instruction.

Under Michigan's benchmark assessment legislation, districts must administer either a benchmark assessment from the MDE-approved provider list, an assessment that provides adequate progress monitoring, or a local benchmark assessment to all K-8 students in the fall and spring of each school year. This requirement first went into effect in the 2020-21 school year and will continue through the end of 2022-23. Districts that choose an assessment from one of the four approved providers must provide aggregate data regarding the results of these assessments through MDH. The MDH is designed to collect student-level data, and districts were encouraged to submit student-level data rather than aggregating the data themselves. Doing so allows MEDC and EPIC to complete all necessary aggregations in a consistent manner across districts, while still ensuring that state agencies maintain access only to aggregate data as stipulated in Michigan's benchmark assessment legislation (2020 PA 149 and 2021 PA 48).

In this section, we describe the indicators of academic performance from the benchmark assessment data and the analytic samples we will use in this report. For a full description of the unique characteristics of each MDE-approved benchmark assessment, please see the first report in this series.

INDICATORS OF ACADEMIC PERFORMANCE ON BENCHMARK ASSESSMENTS

To meet the reporting requirements in Michigan's benchmark assessment legislation, we include analyses summarizing student achievement and growth, as measured by an approved benchmark assessment, between fall 2020 and spring 2022. In what

follows, we provide details about the benchmark assessment data that districts submitted to the MDH and explain how we use these data to derive indicators of average achievement and growth during and across the two full years of COVID-19 pandemic-affected schooling, 2020-21 and 2021-22.

Comparisons to Pre-Pandemic Achievement

For a general understanding of how Michigan students' benchmark assessment scores compare to national averages from before the COVID-19 pandemic, we first calculate the percent of students in each testing period who scored above the pre-pandemic national average for their grade level. To put these results into the context of Michigan students' past performance and Michigan-specific proficiency standards, we also use information from each benchmark assessment provider to translate students' scores into approximate M-STEP proficiency levels (not proficient, partially proficient, proficient, or advanced). We then compare these M-STEP proficiency equivalencies to the actual M-STEP proficiency levels of similar students in 2018-19.[1]

Trends in Average Scale Scores

The next set of analyses examines trends in average scale across four testing periods: fall 2020, spring 2021, fall 2021, and spring 2022. The MAP Growth, i-Ready, Star 360, and Smarter Balanced ICA benchmark assessments are scored on vertical scales that are consistent within vendors across all grade levels, allowing us to compare scores for the same group of students on the same assessment across two different school years. However, because each benchmark assessment has its own unique scale and scale scores are not comparable across assessments (e.g., MAP Growth scores range from 100 to 350 whereas i-Ready scores range from 0 to 800), we present cohort-specific trends in average scale scores separately for each provider.

As comparison points help us interpret these overall trends, we plot the Michigan trend lines over the four testing periods alongside norms that each assessment provider established before the COVID-19 pandemic. While we use pre-pandemic national or Michigan-specific medians as comparison points for all benchmark assessments, not all providers calculate or present this information in the exact same ways. For the MAP Growth assessments, we use NWEA's conditional growth distributions to identify comparison points based on the 50th percentiles of initial fall scores and fall-to-spring growth for two consecutive grade levels (Thum & Kuhfeld, 2020). For Curriculum Associates and Renaissance Learning, we use the fall and spring distributions of scale scores from the norming sample to identify fall and spring medians as pre-pandemic comparison points for the i-Ready and Star 360 assessments, respectively (Curriculum Associates, 2020b; Renaissance Learning, 2021b, c, d). Finally, for the Smarter Balanced ICAs, we compare fall scores to the 50th

percentile for the norming sample from the prior grade level and compare spring scores to the 50th percentile of the norming sample for the current grade level (Smarter Balanced Validity Research, 2020). Unlike the distributions for Curriculum Associates and Renaissance Learning, the Smarter Balanced ICA distributions are only available as end-of-year percentiles and not presented separately for the fall and spring semesters. Hence, we set the fall norms equal to the spring norms for the prior grade level (e.g., the fall norm for 4th grade is the same as the spring norm for 3rd grade). This allows us to compare growth throughout the school year; however, we cannot account for any "summer slide" for students who completed these tests.

Percent of Students Achieving Typical Yearly Growth

Although we can compare average scale scores across grades, it is important to note that the "typical" amount of test score growth over the course of a school year often differs by grade level, subject, *and* initial achievement level. To account for all of these differences, we compare changes in students' scale scores to "typical growth," based on pre-determined growth norms for each assessment provider, subject, grade level, and initial (fall 2020 or fall 2021) score range.

The growth norms for each assessment are defined in slightly different ways and have slightly different meanings. For students who completed MAP Growth assessments, we use as a growth norm the 50th percentile of the fall-to-spring conditional growth distribution for students with the same initial percentile rank in the same grade level and subject (Thum & Kuhfeld, 2020). For Curriculum Associates, we use typical growth targets from the i-Ready assessment growth models which indicate the median growth of students in the same grade level with the same initial placement levels (nationwide pre-pandemic; Curriculum Associates, 2020b). For Star 360 and Smarter Balanced ICA assessments, we use pre-pandemic scale score distributions to identify "typical growth" as the change in scale scores necessary for a student to maintain the same percentile rank in the spring as they did in the fall (Renaissance Learning, 2020b, c, d; Smarter Balanced Validity Research, 2020). These measures represent the increase in scale scores necessary for a Star 360 or Smarter Balanced ICA student who scored, for example, in the 25th percentile in fall 2020 to also score in the 25th percentile on their spring 2021 benchmark assessment.

While these benchmarks help us gain a better understanding of academic growth among Michigan students during the COVID-19 pandemic, it is important to note that we are using summary tables released by each assessment provider to assign growth norms to groups of students; each assessment provider uses sophisticated student-level models to derive growth measures and we are unable to perfectly replicate those measures from just the summary tables and the aggregate district-level data made available under the Return to Learn legislation. For example, most assessment

providers account for the number of instructional days a student received between two testing occasions, based on the test dates relative to the district's instructional calendar, in their growth calculations. For our aggregate, statewide analyses, we cannot account for the exact amount of instructional time between each student's annual fall and spring assessments and, accordingly, we assign growth norms as though the timing were the same for all students.

To assess students' actual growth relative to "typical growth," we first calculate the difference between each student's spring and fall scale scores from the same school year, then compare this fall-to-spring change to the appropriate growth norm (i.e., the *typical scale score increase* based on the assessment provider, grade level, subject, and the student's initial achievement level). Before aggregating the data to the district level, we group students into three categories that describe their fall-to-spring growth for each school year: Students who did not demonstrate any growth at all (i.e., their scale scores remained the same or decreased from fall to spring); students who achieved partial growth (i.e., their scale scores increased from fall to spring, but the increase was less than the typical growth for their grade, subject, and initial achievement level); and students who met or exceeded their growth targets (i.e., their scale scores increased by an amount equal to or greater than the typical growth for their grade, subject, and initial achievement level). For students who made partial progress toward, but did not reach, typical growth, we calculate the proportion of typical yearly growth that they achieved.

DATA AGGREGATION AND ANALYSIS

General Assessment Data Exclusions

Before aggregating the student-level benchmark assessment data provided through the MDH, we restricted the sample to exclude: 1) districts that were not required to report data under Michigan's benchmark assessment legislation (i.e., districts that did not use products from an MDE-approved assessment provider and districts that opened after the official fall 2021 count date or closed before the official spring 2022 student count date); 2) students who are not in grades K-8; 3) results from assessments in subject areas other than math and reading/ELA or that cover only a narrow sub-topic within math or reading/ELA (i.e., the Smarter Balanced Interim Assessment Blocks, or IABs, which do not cover as broad a range of topics as the ICAs); and 4) results from assessments that are not normed for the grade level of the assessed student (i.e., results from Star Early Literacy assessments for students above grade 3 and results from Star Math assessments for students in kindergarten).

Demographic and Modality Data

We merged the benchmark assessment data with student characteristics from the Michigan Student Data System (MSDS) fall 2020 and fall 2021 General Collections for the purpose of identifying student subgroups based on their gender, race/ethnicity, economically disadvantaged status, and special education status. We excluded students from subgroup breakdowns if they were missing the necessary demographic data to determine whether they belonged to the particular subgroup of interest.

We also merged these data with instructional modality data collected from districts during the 2020-21 and 2021-22 school years. The modality data from each year are collected differently and provide different information. In particular, the 2020-21 modality data capture information about which modes of instruction (fully in-person, hybrid, or remote) *a district offered* throughout the year, while the 2021-22 data capture information about which mode(s) of instruction *a specific student received* throughout the year.

2020-21 District Modality Data

The 2020-21 data summarizes districts' submissions to the Reconfirmed COVID-19 Learning Plan Monthly Questionnaire administered through MDE's Grant Electronic Monitoring System/Michigan Administrative Review System (GEMS-MARS) application. Each month, districts were asked to indicate whether they planned to instruct any of their students in a fully in-person (student receive 100% of their instruction in person), fully remote (students receive 100% of their instruction remotely), or hybrid (students attend school in person for part of the week and participate in remote instruction for part of the week) format.

2021-22 Student Modality Data

For the 2021-22 school year, districts provided modality data directly through the MDH. These new data describe the mode of instruction that each Michigan student received throughout the 2021-22 school year, noting changes in students' mode of instruction *that were longer than a month* and the dates when those changes occurred. Given the emphasis to educate students in person as much as possible during the 2021-22 school year, very few students received hybrid or remote instruction for any substantial amount of time during the year. For example, if a student received in-person instruction for most of the school year but was also educated remotely for at least a month in the middle of the year, we would see three unique modality "spells" for this student along with the start and end dates for each unique period. For students who were educated in a single modality for the whole school year, we would observe only one modality spell and the start and end dates for this spell correspond with the beginning and end of the 2021-22 school year.

Aggregate Data File Construction

To construct the final aggregate data files used for the analysis, we calculated the average scale scores and the percentages of students who did not demonstrate growth, achieved less than "typical growth," and met or exceeded "typical growth" in each school year across all students in the same subgroup and grade level who completed an assessment from the same provider. We calculated each of these aggregate measures both by district and for the state as a whole. We then combined the resulting aggregate datasets with data from individual districts that prepared their own aggregate data files in lieu of submitting student level data through the MDH. We completed this process separately for three types of analytic samples (described later in this section) to create aggregate measures appropriate for examining student achievement in one specific testing period, growth across a single school year (2020-21 or 2021-22), and longitudinal trends for a consistent group of students across the fall 2020, spring 2021, fall 2021, and spring 2022 testing periods. The results we present in this report are aggregated to the state level. To prevent identification of any individual students from very small subgroups, we suppress results for any cells that represent fewer than ten students.

District Participation

Under Michigan's benchmark assessment legislation, school districts serving K-8 students are expected to submit benchmark assessment data in some form. For this analysis, CEPI identified districts of interest as those that served students in at least one K-8 grade level and were open as of the official fall student count date for the 2021-22 school year (October 6th, 2021) and remained open as of the official spring student count date (February 9th, 2022).

In total, 730 of Michigan's 849 school districts provided some form of spring 2022 benchmark assessment data through the MDH. Of these, 714 provided student-level data, 14 provided aggregate files that they prepared themselves, and two provided both student-level and aggregated data. We omitted five of these districts from our analyses because all the assessment results they provided were from time periods, grade levels, or subject areas that are not within the scope of this report. The remaining 725 districts are represented in our analyses (709 that provided student-level data, 14 that provided aggregate data, and two that provided both). This includes 609 districts using NWEA's MAP Growth, 67 using Curriculum Associates' i-Ready assessments, 75 using Renaissance Learning's Star 360 assessments, and 28 using DRC's ICAs and MDE's K-2s. Fifty-two of these districts administered assessments from two different providers and one used assessments from three providers. These 725 districts teach 818,454 K-8 students, or 88% of the population of K-8 students in Michigan.

The Return to Learn legislation specifies a few options for districts as alternatives to the four approved benchmark assessment providers. The remaining 119 districts that did not provide any data through the MDH indicated through GEMS/MARS (2020 PA 149) that they selected an alternate vendor or locally developed assessment, did not plan to submit any benchmark assessment data, did not provide the necessary authorization for MEDC and EPIC to access their data in the MDH, or provided the authorization but did not have any student benchmark assessment data in the MDH by the deadline for us to include them in the report.

Analysis Samples

While the full sample includes data from all students with valid test scores for a given testing period, we impose sample restrictions for our longitudinal analyses to ensure that comparisons of aggregate measures over time reflect changes in student performance as opposed to changes in the populations of students tested. We have two types of restricted analytic samples.

1. **The 2020-21 and 2021-22 single-year school year growth samples** include students with valid test scores in the fall and spring of 2020-21 or fall and spring of 2021-22 that were completed in the same subject, grade level, and district (i.e., the 2020-21 growth sample includes students with valid test scores in fall 2020 and spring 2021, while the 2021-22 growth sample includes students with valid test scores in fall 2021 and spring 2022).
2. **The two-year growth sample** includes students with valid test scores in all four semesters between fall 2020 and spring 2022 that were completed in the same subject and same district, in the same grade level in both the fall and spring of 2020-21, and the subsequent grade level in both the fall and spring of 2021-22.

Table 2.1 shows the total number of districts and students for whom we received spring 2022 student-level data through the MDH and the subsets of these students who we can and cannot include in the restricted samples for our longitudinal analyses. The figures in the top panel represent the exclusions for the 2021-22 school year growth sample. Though not shown here, we use similar exclusions to identify a 2020-21 school year growth sample for comparisons of fall-to-spring growth for each of the two school years. Figures in bottom panel shows the exclusions for the two-year growth sample.

Table 2.1. Restricted Analytic Samples and Reasons for Exclusions

Exclusions	Districts		Students	
	N	%	*N*	%

2021-22 School Year Growth Sample and Exclusion Reasons				
Spring 2022 sample	**711**	**100**	**679,104**	**100**
Missing fall 2021 data	-6	-1	-26,579	-4
Different district in fall 2021 than in spring 2022	-0	-0	-5,598	-1
Different test in fall 2021 than in spring 2022	-0	-0	-406	-0
2021-22 school year growth sample	**705**	**99**	**646,521**	**95**
Two-Year Growth Sample and Exclusion Reasons				
2021-22 school year growth sample	**705**	**99**	**646,521**	**95**
New kindergarten cohort, not tested in 2020-21	-0	-0	-64,852	-10
Missing fall 2020 and/or spring 2021 data	-120	-17	-147,057	-22
Different district in 2020-21 than in 2021-22	-0	-0	-25,434	-4
Different test in 2020-21 than in 2021-22	-11	-2	-12,703	-2
Two-year growth sample	**574**	**81**	**396,475**	**58**

Note: The counts and percentages in this table do not include data from districts that prepared their own aggregate datasets. The district and student percentage for each exclusion may not add to 100 due to rounding.

The full sample for the spring 2022 testing period includes data from 679,104 students in 711 districts. For analyses of 2021-22 growth, we exclude a total of six districts and 32,583 students. The most common reason for exclusion of students from our 2021-22 school year growth sample was simply that districts did not provide fall 2021 data for a particular student, or in some cases, for any students at all. The remaining 646,521 students and 705 districts in the 2021-22 school year growth sample participated in comparable benchmark assessments, in the same district, in the same grade levels for the entire 2021-22 school year.

The two-year growth sample is the most restrictive sample in that it only includes the 396,475 students (from 574 districts) who participated in comparable benchmark assessments in fall 2020, spring 2021, fall 2021, *and* spring 2022 in the same district, in the same grade level in both the fall and spring of 2020-21, and in the next consecutive grade level in both the fall and spring of 2021-22. About 61% of students and 81% of districts in the 2021-22 school year growth sample were also in the two-year growth sample. The new cohort of students who began kindergarten in 2021-22, and therefore did not participate in benchmark testing the previous year, accounts for nearly 65,000 of the students we exclude from the two-year growth sample. Another major factor driving the high exclusion rate is that more than 100 districts that chose to use a locally developed benchmark assessment in 2020-21 switched to an assessment from one of the four MDE-approved providers the following year, as the benchmark assessment legislation for 2021-22 allocated new funding for districts to

implement these assessments (2021 PA 48). We include these districts in our spring 2022 and 2021-22 school year analyses but cannot measure growth over a two-year period, as they have only one year of benchmark assessment data.

When possible, we include data from the 16 districts that prepared their own aggregate files. These aggregate files include benchmark assessment data for another 49,095 Michigan students, meaning that the combined dataset we constructed from both the student-level and district-provided aggregate data represents 728,199 (or about 77%) of all K-8 students in Michigan. All 16 of the districts that completed their own aggregation provided data that meets the criteria for the spring 2022 and 2021-22 school year growth samples, and 15 of the 16 provided data that meets the criteria for the two-year growth sample. Within these districts, the percentages of students who are represented in the 2021-22 school year and two-year growth restricted samples are nearly identical to the percentages among districts that provided student-level data. Appendix 2 provides information about which districts have data used in each of our analysis samples.

Sample Characteristics and Representativeness

Table 2.2 provides summary statistics of K-8 students in Michigan and by assessment provider. In general, students who took the NWEA MAP Growth assessment in spring 2022 are relatively similar to the statewide population of K-8 students. This is not surprising given that this group represents the majority of Michigan's K-8 students. Students who took the i-Ready assessments are less representative; these students were substantially more likely to be Black, less likely to be White, and somewhat more likely to be Asian compared to the full K-8 population. This is largely driven by Detroit Public Schools Community District, which is the largest school district in Michigan and accounts for more than one-fifth of all students who took an i-Ready assessment despite there being 67 districts that used i-Ready. Compared to the rest of the state, students who participated in the Star 360 assessments are overwhelmingly White and considerably less likely to be eligible for English learner services. Students who participated in DRC assessments (MDE and ICAs) are the least likely to be economically disadvantaged or eligible for special education or English learner services. Students who did not participate in any of these benchmark assessments in spring 2022 were more likely to be eligible for special education or English learner services but were otherwise similar to the statewide population.

Table 2.2. Summary Statistics of K-8 Students in All Michigan Districts and by Assessment Provider						
Demographics (%)	**All MI**	**MAP Growth**	**i-Ready**	**Star 360**	**MDE/ICA**	**Not tested**
Female	48.6	47.8	48.1	47.7	48.5	48.0
Asian	3.6	2.5	4.9	1.4	0.2	5.0
Black	18.4	15.6	34.0	7.7	18.6	17.9
Latino	8.8	8.7	9.0	9.2	3.6	8.4
White	63.1	65.1	46.9	73.5	73.7	62.2
Economically disadvantaged	55.7	52.7	59.1	54.0	52.9	56.7
Special education	14.1	12.8	11.7	13.5	9.8	16.6
English learner	7.7	5.4	9.9	4.1	1.4	10.8
N students	946,987	505,838	125,951	58,989	5,242	258,488
% of MI K-8 students	100.0	53.4	13.3	6.2	0.6	27.2

Note: The "All MI" column includes the full population of K-8 students across Michigan. Each vendor-specific column includes all students who took a MAP Growth, i-Ready, Star 360, or K-2/ICA assessment in spring 2022, respectively. The "not tested" column includes K-8 students who did not participate in a benchmark assessment from any of these providers in spring 2022.

Research exploring trends in academic achievement over the past two years, including EPIC's third report using Michigan benchmark data, makes clear that the effects of the COVID-19 pandemic on students varied across student populations and the COVID-19 pandemic has had a greater and more negative effect on the achievement and achievement growth of economically disadvantaged, Black, and Latino students, as well as English learners (e.g., Kilbride, Hopkins, Strunk, & Imberman, 2021; Kuhfeld & Lewis, 2022; Kogan & Lavertu, 2021; NCES, 2022; Pier et al., 2021; Sass & Goldring, 2021). Given that these specific student populations are underrepresented in the analytic samples for some of the benchmark assessment providers, it is likely that our overall results overstate any academic growth observed throughout the 2020-21 and 2021-22 school years.

Table 2.3 presents grade-specific enrollment counts and percentages of enrolled students who are represented in each of our analytic samples. The denominator for each inclusion rate is the aggregate enrollment count across all districts offering a particular benchmark assessment for a particular grade level (e.g., a district may use MAP Growth for some grade levels and a locally developed assessment for others). Since grade-specific enrollment counts and inclusion rates were relatively consistent across our reading and math samples, we provide figures for the percentage of students with valid test scores in at least one subject area. The percentages in this table do not include students from districts that submitted their own aggregate data.

These districts reported math and reading outcomes separately but did not indicate how many students participated in benchmark testing for both subjects.

As seen in the table, inclusion rates vary across grade levels and analytic samples. Across all three samples, inclusion rates among elementary school students increase with grade level while rates among middle school students decrease for each grade level. Additionally, inclusion rates across all grade levels are consistently higher among students who participated in the spring 2022 assessments since the group includes any student with at least one valid test score in spring 2022. The largest differences in inclusion rates between the spring 2022, 2021-22 fall-to-spring, and two-year growth samples are for the students who were in kindergarten or 1st grade in fall 2020. This is likely due to transitions between tests (e.g., students who take the Star Early Literacy assessment in kindergarten then switch to the Star Math and Reading assessments in first or second grade) or districts that use assessments from different providers for their lower and upper elementary and middle school students.

Table 2.3. Percent of Enrolled Students Included in Analytic Samples, All Four Assessment Providers

2020-21 Grade	2021-22 Grade	2021-22 Enrollment	Inclusion Rates by Analytic Sample		
			Spring 2022	2021-22 School Year Growth	Two-Year Growth
---	K	83,997	84.0	76.1	0.0
K	1st	74,559	90.6	88.0	49.9
1st	2nd	79,030	91.1	88.6	56.9
2nd	3rd	79,441	92.3	90.1	62.4
3rd	4th	79,840	92.8	90.6	65.7
4th	5th	81,091	93.3	91.0	66.0
5th	6th	82,196	91.9	89.3	63.3
6th	7th	83,716	91.4	88.5	62.6
7th	8th	86,132	90.3	87.0	61.8

Note: The "Enrollment" column represent the total number of students from a specific K-8 grade level enrolled in districts that offered an NWEA MAP Growth, Curriculum Associates i-Ready, Renaissance Learning Star 360, or MDE K-2/Smarter Balanced ICA benchmark assessment during the spring 2022 semester and provided student-level data to the MDH. The "2021-22 School Year Growth," and "Two-Year Growth" columns represent the percentage of students from each grade with valid mathematics or reading benchmark assessment scores in fall 2021 and spring 2022; or fall 2020, spring 2021, fall 2021, and spring 2022, respectively.

Analyses Post-Aggregation

Subgroup Comparisons

In addition to assessing all students' actual growth on each benchmark assessment relative to their "typical growth," we also investigate differences across student demographics and the instructional modality students received throughout the 2021-22 school year. Specifically, we compare achievement and growth on benchmark assessments across race/ethnicity, economically disadvantaged status, and special education status. We also examine differences across subgroups of students whose *districts offered* different modes of instruction in 2020-21 as well as subgroups of students who *received* different modes of instruction in 2021-22.

Table 2.4 summarizes the instructional modalities that districts offered throughout 2020-21. The top panel shows both the percent of *districts* that offered each mode of instruction for the majority of the school year and the percent of *students* who attended a district that offered each mode of instruction for the majority of the school year. Percentages for each of the three modalities add to more than 100 because districts could, and typically did, offer more than one mode of instruction at the same time. Nearly all districts offered remote instruction for the majority of the year, while about two-thirds offered in-person instruction and one-third offered hybrid instruction. Discrepancies between these percentages and the percentages of students in districts that offered each modality suggest that the districts that offered in-person instruction tended to be smaller than the state average. Modality offerings among MAP Growth districts generally mirrored the statewide offerings, while i-Ready districts were less likely to offer in-person or hybrid instruction, Star 360 districts were substantially more likely to offer in-person and hybrid instruction, and districts that used the Smarter Balanced ICA and MDE assessments were more likely to offer in-person instruction but less likely to offer hybrid instruction.

The bottom panel of Table 2.4 shows the number of months that districts offered each mode of instruction, both as an average across districts and as an average across students. The highest possible number of months is 9, as districts reported their 2020-21 modality offerings each month from September 2020 to May 2021. On average, students had the option to learn in person for between half and two-thirds of the school year, while hybrid options were available for about a third of the year, and remote instruction was offered throughout the entire year. Students in districts that used the Star 360, Smarter Balanced ICA, and MDE K-2 assessments tended to have more access to in-person instruction, while students in i-Ready districts had the least access.

Table 2.4. Districts' Instructional Modality Offerings in 2020-21					
Mode of instruction offered for majority of 2020-21	**All**	**MAP Growth**	**i-Ready**	**Star 360**	**MDE/ ICA**
Fully in-person					
Percent of districts	64.6	63.5	57.4	80.0	74.1
Percent of students	59.9	57.0	64.5	73.6	79.4
Hybrid					
Percent of districts	32.5	31.9	26.2	54.0	18.5
Percent of students	35.7	38.0	15.7	61.9	16.3
Fully remote					
Percent of districts	95.4	94.8	98.4	96.0	100
Percent of students	97.6	97.0	99.2	99.7	100
Number of months mode of instruction was offered in 2020-21	**All**	**MAP Growth**	**i-Ready**	**Star 360**	**MDE/ ICA**
Fully in-person					
Average across districts	5.8	5.7	5.1	7.2	6.9
Average across students	5.4	5.2	5.3	6.4	7.3
Hybrid					
Average across districts	2.9	2.8	2.7	4.3	1.7
Average across students	3.3	3.4	2.5	4.8	1.4
Fully remote					
Average across districts	8.4	8.4	8.7	8.6	8.7
Average across students	8.6	8.6	8.8	8.6	8.9

Note: Districts that offered more than one mode of instruction for the majority of the year are included in calculations for all modes of instruction that apply. The numbers in this table do not include data from districts that prepared their own aggregate data sets.

Table 2.5 shows that nearly all students (about 98%) returned to in-person learning in 2021-22. About 2% received fully remote instruction, and almost none received hybrid instruction or switched between modalities (excluding short-term changes due to a local COVID-19 outbreak, for instance). This pattern is consistent across vendors and across students' access to in-person instruction in the previous school year. However, we note that these percentages are based on the 87% of students whose districts reported student-level modality data for the 2021-22 school year; we cannot determine how the remaining 13% of students received their instruction. Given how few students reportedly received instruction in hybrid or remote formats, we limit our comparisons to just two subgroups: students who received in-person instruction all year and students who received any other mode of instruction (hybrid, remote, or a combination of modalities). In total, 15,035 students from 170 different districts received remote or hybrid instruction in 2021-22. Two-thirds of these students

attended traditional public and charter schools that provide face-to-face instruction, while the other third attended charter schools that have always operated virtually.

Table 2.5. Percent of Students Participating in Each Modality in 2021-22; Overall, by Assessment, and by Access to In-Person Instruction in the Prior School Year

	In-Person	Hybrid	Remote	Multiple	No Data
All students in spring 2022 sample	97.5	0.0	2.2	0.4	13.0
By assessment					
MAP Growth	97.7	0.0	1.8	0.5	2.6
i-Ready	97.9	0.0	2.1	0.1	30.5
Star 360	94.6	0.1	5.3	0.1	9.5
Smarter Balanced ICA/MDE K-2	99.6	0.0	0.5	0.0	6.8
By access to fully in-person instruction in 2020-21					
Never offered	97.2	0.1	2.2	0.5	11.9
Offered for less than half of year	97.6	0.0	1.7	0.7	7.1
Offered for at least half of year	98.8	0.0	0.9	0.4	21.3
Offered all year	99.0	0.0	0.9	0.1	9.0

Note: The percentages in this table do not include data from districts that prepared their own aggregate data sets. Students whose districts did not provide mode of instruction data for 2021-22 are not included in calculations for the first four columns.

Regression Analyses

In addition to simple comparisons of average scale scores and achievement growth for all students and across subgroups, we use the aggregated benchmark assessment data in multiple regression models to estimate the relationships between the average achievement in a district over time while controlling for other district characteristics. Multiple regression is a statistical technique used to predict an outcome variable using two or more explanatory variables. This technique allows us to estimate the unique relationship between academic achievement in two consecutive semesters and show how this relationship changed over time, when all else is equal between districts.

We use separate regression models for each assessment provider, where the achievement outcomes in each model represent the average benchmark assessment score for a particular district, grade level, and subject at a particular point in time. We standardize scores relative to the means and standard deviations of student scores from the pre-pandemic norming samples for each assessment to interpret the standardized scores in terms of how a district's average achievement compares to the average achievement of students across the country before the COVID-19 pandemic.

For example, a standardized score of negative one indicates that a district's average score is one standard deviation below the national pre-pandemic average.

We estimate the following baseline model:

$$Y_{dgsvt} = \beta_0 + \beta_1 S21_t + \beta_2 F21_t + \beta_3 S22_t + \boldsymbol{\beta_4 DCHAR_{dgt}} + \delta_g + \epsilon_t$$

where Y_{dgsvt} is the average standardized test score of students in district *d*, grade *g,* completing subject test *s* from assessment provider *v*, in semester *t*. *S21, F21, and S22* are binary indicators identifying the semester associated with the outcome of interest, Y_{dgsvt}. The coefficients on these indicators, shown here as β_1, β_2, and β_3, describe the change in average standardized test score attributable to each semester, spring 2021, fall 2021, and spring 2022, respectively, relative to fall 2020.

We control for a set of district characteristics, $\boldsymbol{DCHAR_{dgt}}$, including the proportions of students in each district-grade who are female, of different races/ethnicities, economically disadvantaged, eligible for special education services, and English learners. We mean-center these characteristics in each school year so that a value of zero represents the state average, allowing us to interpret coefficients for each time period as the predicted achievement for an average district. δ_g is a grade fixed effect which controls for differences in standardized tests scores that are unique to a particular grade level, thus enabling us to compare across grades in the same model.

After estimating the baseline model, we include additional explanatory variables describing the instructional modalities offered to students during the 2020-21 school year to see how achievement varied across students with access to differed modes of instruction. We examine the effects of *each additional month* in which a district offered in-person or hybrid instruction. For each explanatory variable in our models, we estimate a regression coefficient which tells us how districts' average achievement scores would be expected to change if the explanatory variable were to increase by one unit. For example, a regression coefficient of 0.05 for the number of months a district offered in-person instruction would indicate that every additional month of in-person instruction is associated with an increase in district-average achievement of five-hundredths (or 5%) of a standard deviation relative to students in a particular pre-pandemic norming sample.

SUMMARY

The analyses included in this report are based on data representing 85% of Michigan districts (725 of the 848 total districts) and 77% of students in the state (728,199 of the 934,987 total students). However, those who are represented in our analyses may not

be reflective of those who are not included. While it is important to keep in mind the limitations of the data when interpreting results, the report nonetheless helps deepen our understanding of how Michigan public school students progressed academically between the fall 2020 and spring 2022 semesters. The analyses presented in Section Three continue to expand on the descriptive results presented in our previous reports, providing a more refined estimate of academic growth by incorporating another semester of assessment data and comparing academic trajectories of Michigan students to pre-pandemic trajectories of students from across the country.

Section Three: Results

In this section, we summarize outcomes for Michigan students on benchmark assessments that districts administered between fall 2020 and spring 2022. We compare Michigan achievement to pre-pandemic national averages and Michigan specific proficiency rates. We also show changes in average scale scores throughout the 2020-21 and 2021-22 school years using both unadjusted averages and multiple regression estimates which control for other district and grade-specific student characteristics. Finally, we assess how Michigan students' test score growth in 2020-21 and 2021-22 compares to pre-determined growth norms that each assessment provider established before the COVID-19 pandemic, as well as differences across student characteristics and instructional modalities (e.g., fully in-person, fully remote, hybrid instruction).

Before interpreting results from these analyses, it is important to remember that each benchmark assessment has its own unique scale and scale scores are not comparable across assessments. We therefore analyze data from each assessment provider separately. However, we compare general patterns and some standardized metrics across the different assessments to identify commonalities and points of divergence in the results.

COMPARISON BETWEEN COVID-19 PANDEMIC AND PRE-PANDEMIC ACHIEVEMENT

For a general understanding of how Michigan students' benchmark assessment scores compare to national averages from before the COVID-19 pandemic, we calculate the average percentage of students across districts in each COVID-19 pandemic testing period who scored above the pre-pandemic national average for their grade level.[2] We also use information from each benchmark assessment provider to translate students' benchmark assessment scores into approximate M-STEP proficiency levels (not proficient, partially proficient, proficient, or advanced). We then compare these M-STEP proficiency equivalencies to the actual M-STEP proficiency levels of students in the same districts in 2018-19. This allows us to gauge how well Michigan students might perform on the state's summative assessment and compare these estimates to the actual M-STEP proficiency levels of students in the same districts in 2018-19.[3]

HOW TO INTERPRET FIGURES 3.1.1 TO 3.1.5

Figure 3.1.1 shows the percentage of Michigan students across all grade levels who scored above the pre-pandemic national average for their grade level on the MAP Growth, i-Ready, Star 360, and Smarter Balanced ICA assessments each semester between fall 2020 and spring 2022. These percentages are based on all Michigan K-8 students who participated in one of these assessments in a particular semester.

Green and blue bars represent the percentages for math and reading, respectively. The **lighter shades** of each color represent results from the 2020-21 school year, while the **darker shades** represent 2021-22. If a bar is above the **grey dashed line**, this means that the majority of Michigan students scored above the pre-pandemic national average. If a bar is below this line, the majority of Michigan students scored below the pre-pandemic national average.

Figures 3.1.2 to 3.1.5 show the percentages of Michigan students across 3rd to 7th grade who are classified into each of the four M-STEP proficiency levels, based on definitions from each assessment provider linking scores on the benchmark assessments to an equivalent M-STEP proficiency category. To understand how these proficiency rates compare to similar students' performance before the COVID-19 pandemic, we also show the actual proficiency rates from the 2018-19 M-STEP, among students in the same districts that use a particular benchmark assessment. For example, the first set of stacked bars in Figure 3.1.2 shows the percentages of students who scored in the not proficient, partially proficient, proficient, and advanced categories on the 2018-19 M-STEP, within the group of districts that used the MAP Growth assessments in 2020-21 or 2021-22.

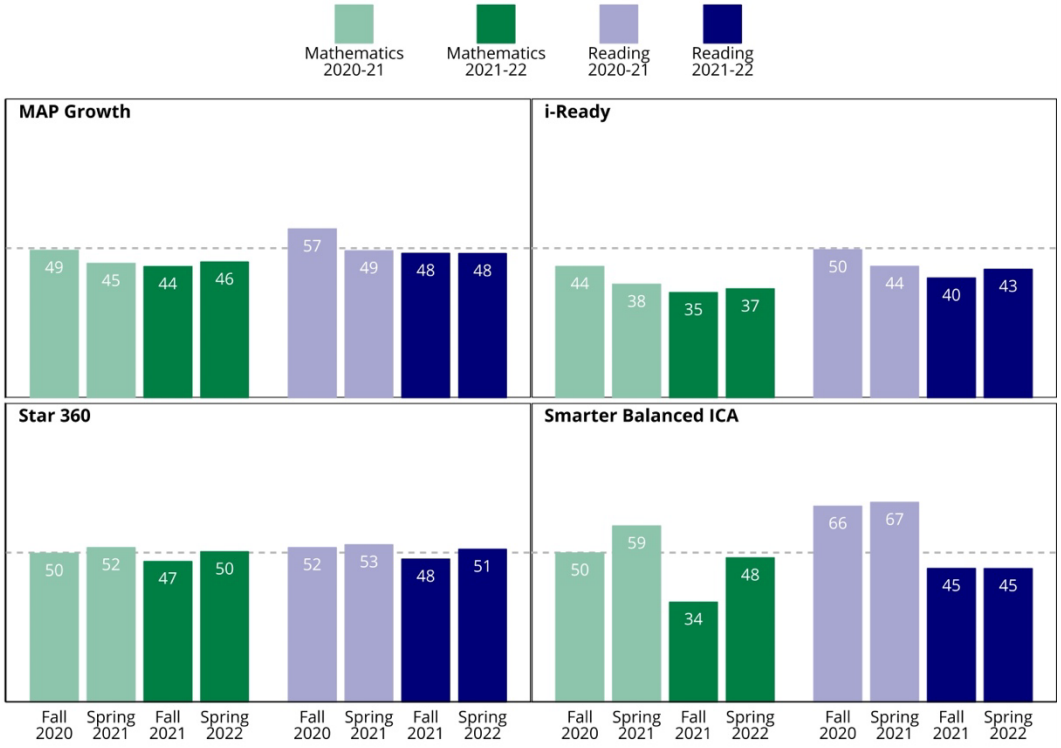
Fewer Michigan Students Scored Above Pre-Pandemic National Averages in Math and Reading in 2021-22 Than in 2020-21

The percentage of Michigan students scoring above national averages decreased during the COVID-19 pandemic. In fall 2020, about half of all Michigan students scored above national averages in math on the MAP Growth, Star 360, and Smarter Balanced ICA assessments, and more than half scored above national averages in reading. Among those who took the i-Ready assessment, about 44% scored above the national average in math, and half did so in reading.

These percentages generally declined over the next two semesters, with some slight increases in spring 2022. While this pattern is fairly consistent for the MAP Growth and

i-Ready assessments, we see fewer changes for the Star 360 assessments, with the percentage of students scoring above national averages hovering close to 50 in all four semesters. Conversely, the percentage of students scoring above national means on a Smarter Balanced ICA assessment changed dramatically between the 2020-21 and 2021-22 school years. This appears to be driven mostly by differences in the groups of students who took a Smarter Balanced ICA assessment each semester; although there were relatively few changes in the districts that used these assessments from semester to semester, districts switching between the Smarter Balanced ICAs and IABs (which cover more narrow sets of subtopics than the ICAs) for some grade levels in some semesters makes it difficult to compare results from the ICAs.

Figure 3.1.1. Percent of Students Who Scored Above the Pre-Pandemic National Average for Their Grade Level



Note: These percentages include all students with a valid benchmark assessment score in any testing period. Grade-specific national averages represent average achievement by students in each assessment provider's national norming sample.

In Most Districts, Estimated Proficiency Rates Changed Little from Spring 2021 to Spring 2022

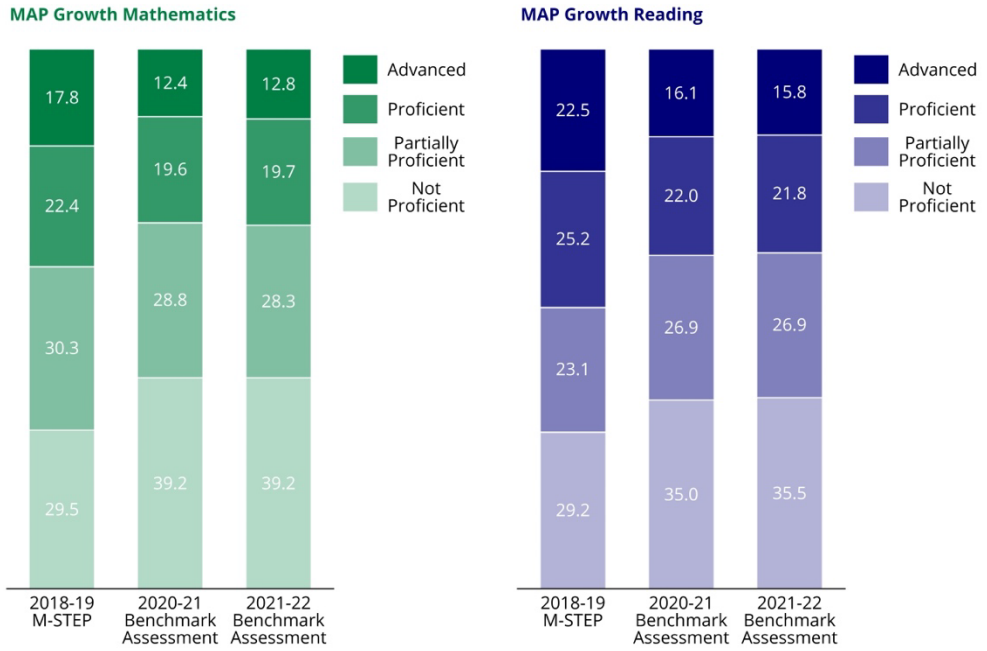
Figures 3.1.2 through 3.1.5 show the distribution of proficiency levels for Michigan 3rd- through 7th-grade students, based on their spring 2021 and spring 2022 benchmark assessment scores. As mentioned earlier, we translate student scores on each benchmark assessment into approximate M-STEP proficiency levels to gauge how well Michigan students might perform on the state's summative assessment and compare these estimates to the actual M-STEP proficiency levels of students in the same districts in 2018-19.

Figures 3.1.2 through 3.1.4 show that estimated M-STEP proficiency rates for students in MAP Growth, i-Ready, and Star 360 districts generally trailed those for similar students who completed an M-STEP assessment during the 2018-19 school year. Specifically, across all three assessment providers and both subjects, more students were classified as not proficient and fewer were classified as proficient or advanced based on their spring 2021 or spring 2022 benchmark assessment scores, compared to students in the same districts who took the M-STEP in 2018-19. However, the percentages of students in each of these proficiency levels did not change much between 2020-21 and 2021-22.

As was the case with percentages of students scoring above pre-pandemic national norms, differences in the population of students who took the Smarter Balanced ICAs from semester to semester make it difficult to compare their results over time. In Figure 3.1.5, it is clear that comparisons between proficiency rates on the 2018-19 M-STEP and estimates from the 2021-22 Smarter Balanced ICAs follow a similar pattern as the other three vendors, while the results from 2020-21 do not. As a result, we primarily focus on the other three vendors in the longitudinal analyses that follow.

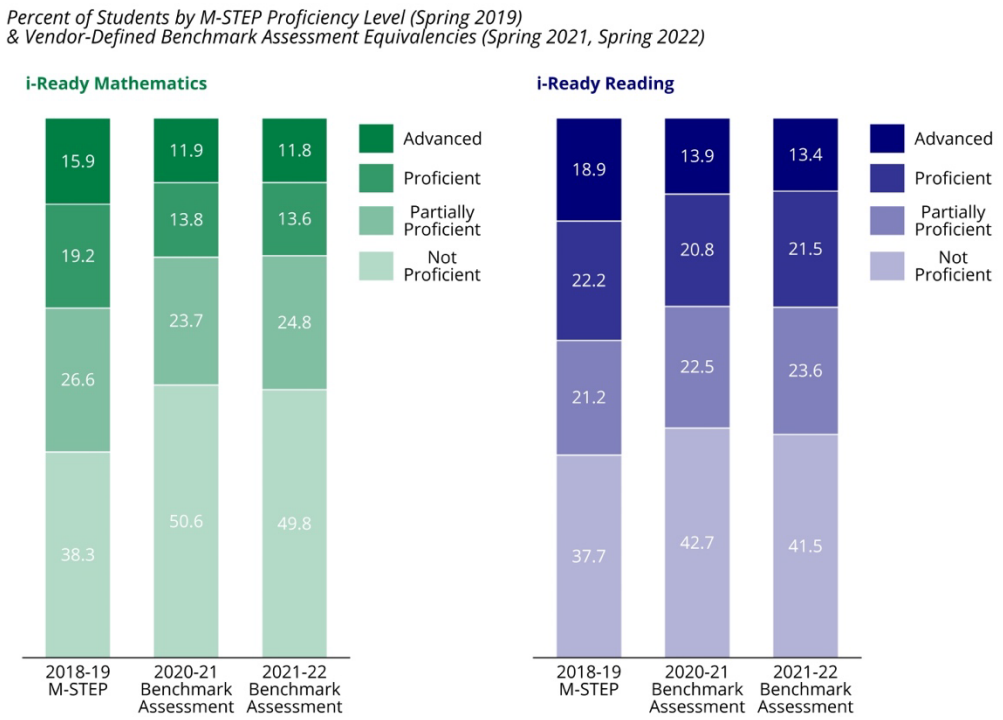
Figure 3.1.2. M-STEP Proficiency Levels and Vendor-Defined M-STEP Equivalencies, NWEA MAP Growth

Percent of Students by M-STEP Proficiency Level (Spring 2019)
& Vendor-Defined Benchmark Assessment Equivalencies (Spring 2021, Spring 2022)



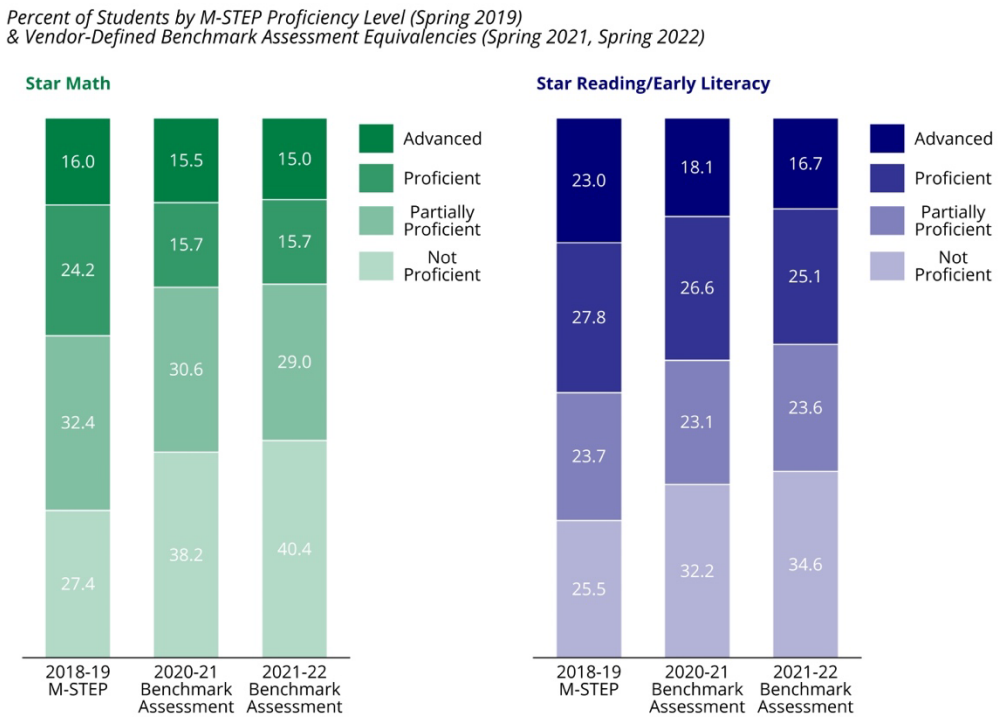
Note: These percentages include 3rd- through 7th-grade students with a valid benchmark assessment score in spring 2020 or spring 2021. Benchmark assessment scores are converted to an estimated M-STEP proficiency category based on a linking study from NWEA. Proficiency rates from the 2018-19 M-STEP include all students in districts that use the MAP Growth assessments.

Figure 3.1.3. M-STEP Proficiency Levels and Vendor-Defined M-STEP Equivalencies, Curriculum Associates i-Ready



Note: These percentages include 3rd- through 7th-grade students with a valid benchmark assessment score in spring 2020 or spring 2021. Benchmark assessment scores are converted to an estimated M-STEP proficiency category based on a linking study from Curriculum Associates. Proficiency rates from the 2018-19 M-STEP include all students in districts that use the i-Ready assessments.

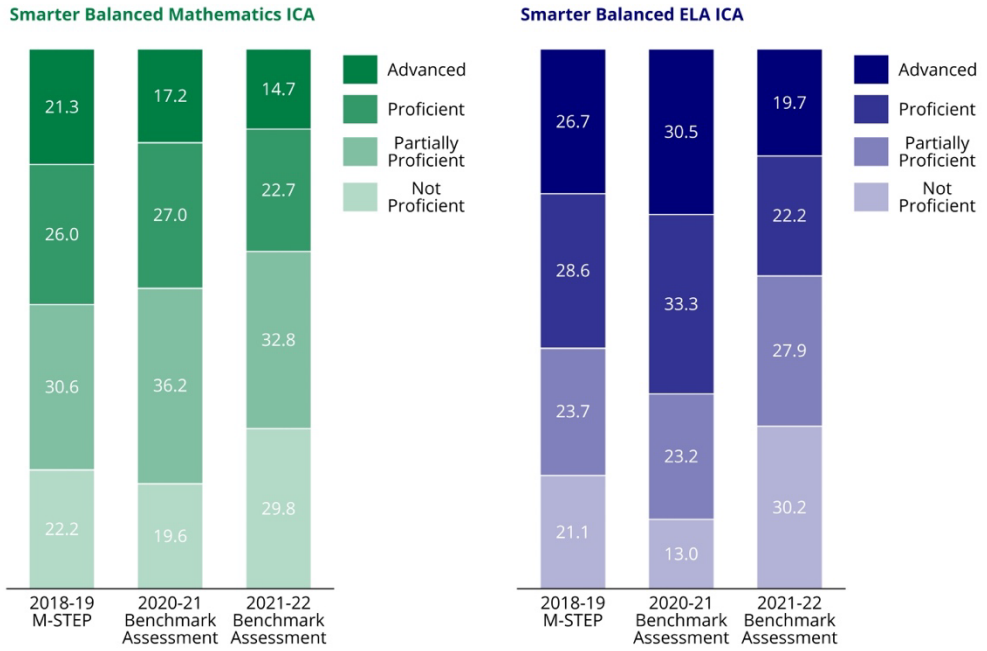
Figure 3.1.4. M-STEP Proficiency Levels and Vendor-Defined M-STEP Equivalencies, Renaissance Learning Star 360



Note: These percentages include 3rd- through 7th-grade students with a valid benchmark assessment score in spring 2020 or spring 2021. Benchmark assessment scores are converted to an estimated M-STEP proficiency category based on a linking study from Renaissance Learning. Proficiency rates from the 2018-19 M-STEP include all students in districts that use the Star 360 assessments.

Figure 3.1.5. M-STEP Proficiency Levels and Vendor-Defined M-STEP Equivalencies, Smarter Balanced ICA

Percent of Students by M-STEP Proficiency Level (Spring 2019)
& Vendor-Defined Benchmark Assessment Equivalencies (Spring 2021, Spring 2022)



Note: These percentages include 3rd- through 7th-grade students with a valid benchmark assessment score in spring 2020 or spring 2021. Benchmark assessment scores are converted to an estimated M-STEP proficiency category using information about the M-STEP and Smarter Balanced scales from each assessment's technical manual. Proficiency rates from the 2018-19 M-STEP include all students in districts that use the Smarter Balanced ICA assessments.

TRENDS IN AVERAGE ACHIEVEMENT DURING THE COVID-19 PANDEMIC

In this section, we show changes in unadjusted and adjusted average scale scores between fall 2020 and spring 2022. We estimate these trends using the two-year growth sample (i.e., students who were tested in all four semesters), as well as the 2021-22 kindergarten and 2020-21 8th-grade cohorts of students who were tested in both semesters of their respective school year but were not eligible or required to participate in testing across both years. We examine trends separately for each cohort of students and compare their trends to pre-pandemic norms from each assessment provider to help place them in the context of how students in the same grade levels performed in previous years.

Average Scale Score Trajectories

Typical-year growth trajectories should show increases in average scale scores over time as students receive more instruction and progress academically, along with slight decreases between each spring and subsequent fall semester as students experience "summer slide" (e.g., see McEachin & Atteberry, 2017; Quinn et al., 2016; von Hippel & Hamrock, 2019). However, the 2020-21 and 2021-22 school years have not been "typical." Prior research—including our own—shows that, on average, student learning between fall 2020 and fall 2021 occurred at a slower rate than would have been expected in a typical year (Amplify Education, 2021; Dorn et al., 2020; Kilbride, Hopkins, & Strunk, 2021; Kogan & Lavertu, 2021; Sass & Goldring, 2021). The average scale score trend results in this report will help us understand whether Michigan student achievement fell further below pre-pandemic norms or rebounded during the 2021-22 school year.

As a reminder, the pre-pandemic comparison points are unique to each assessment provider. We use NWEA's conditional growth distributions to identify comparison points for the MAP Growth assessments, using the median semester-to-semester growth for students with initial fall scores at the 50th percentile (Thum & Kuhfeld, 2020). Given the differences in the characteristics of students from i-Ready districts relative to the state average (see Section Two of this report), we use the distributions of scale scores from Michigan districts that completed the i-Ready assessments in 2018-19, as these comparison points better reflect the population of students studied in this report than the national distributions from Curriculum Associates' norming sample. Finally, we use Renaissance Learning's fall and spring distributions of scale scores from their norming samples to provide fall and spring medians as

pre-pandemic comparison points. Appendix Tables A1 and A2 provide the exact values depicted in each of the following figures.

HOW TO INTERPRET FIGURES 3.2.1 TO 3.2.3

These figures show trends in average math and reading scale scores for Michigan students who took the MAP Growth, i-Ready, or Star 360 assessments.

Green and blue points represent average math and reading scale scores, respectively, for Michigan students in each testing period. All points that represent the same cohort of students are connected by **solid lines**, showing the achievement trajectory for the cohort over time. **Labels on each point** indicate the grade level of students in the cohort during each testing period.

Dashed gray lines represent pre-pandemic comparison points for each cohort of students. **The shaded area between pairs of solid and dashed lines** depicts the difference between each cohort's average score and the pre-pandemic norm for students in the same grade level. When a solid line is above the corresponding dashed gray line, average scale scores for that cohort were above the national pre-pandemic norm. Conversely, when a solid line is below the corresponding gray line, average scale scores for that cohort were less than the norm.

The **y-axis scales** in each figure extend approximately from the kindergarten fall norm for each assessment to the 8th grade spring norm. Although the exact numbers differ slightly between subjects and differ greatly across vendors, the total distance from the bottom to the top of each y-axis always represents the range of grade-level norms from the beginning of kindergarten to the end of 8th grade.

These figures reiterate several key points discussed in our previous report. First, students in early elementary grade levels experienced an "at-home testing advantage," likely due to differences in students' testing environments and the amount of assistance they received on their assessments from their care providers at home. As a result, fall 2020 average scale scores for kindergarten and 1st-grade students in NWEA and i-Ready districts are well above pre-pandemic math and reading norms, but these scores may not accurately reflect students' true achievement levels. Second, fall 2020 average scale scores were noticeably different across students who completed different assessments. This may be due to differences in the characteristics of districts that took each assessment (e.g., i-Ready districts have higher proportions of economically disadvantaged and Black students and are more often located in urban

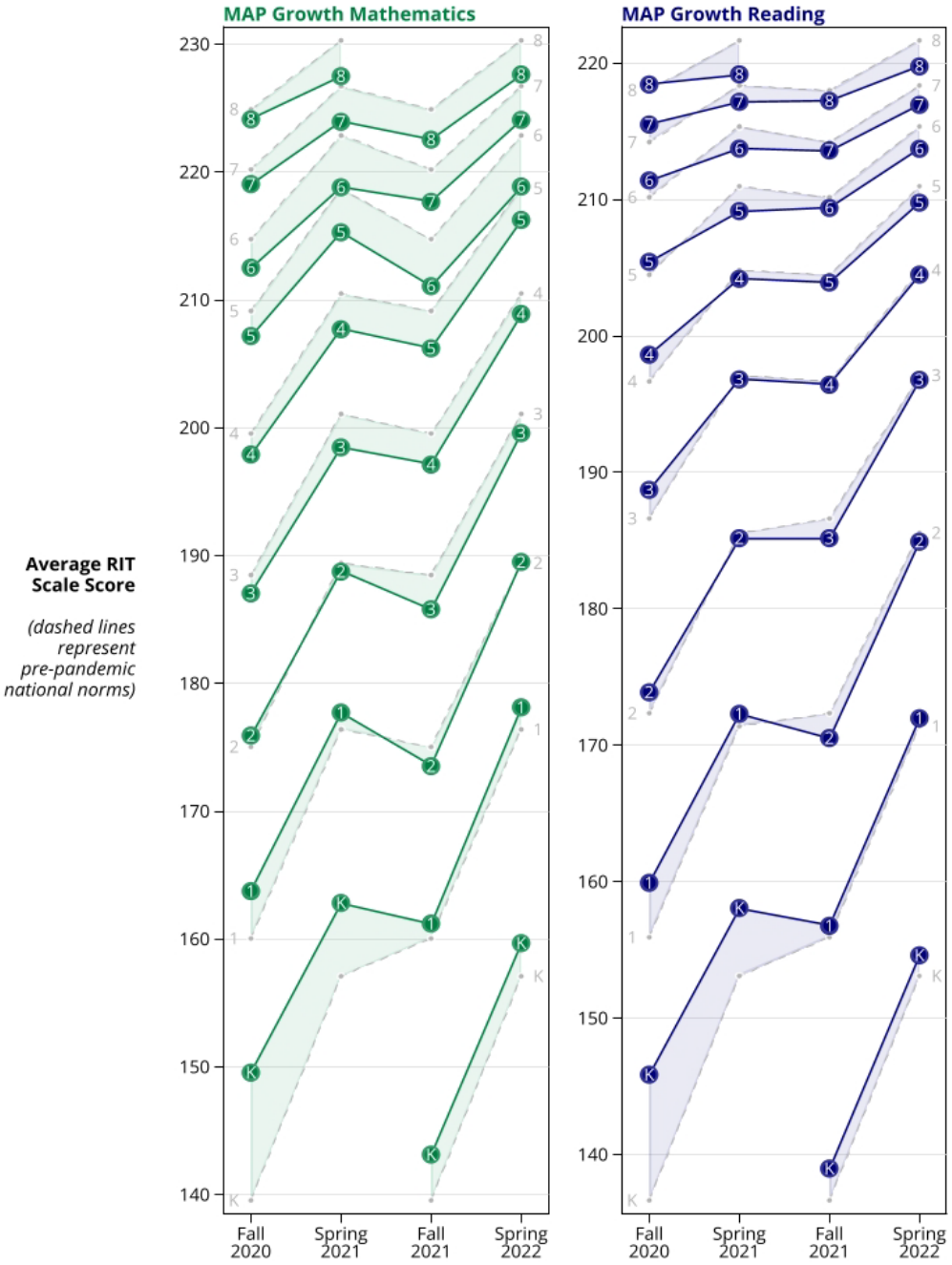
centers whereas Star 360 districts have largely White populations of students and are smaller, on average) or they may be attributable to the use of different modalities in the 2020-21 school year. Third, by fall 2021, average math and reading scale scores across most grade levels were either at or below pre-pandemic norms. Finally, average math and reading scale scores across grade levels generally increased over time. However, by comparing the slopes of the solid lines to the slopes of the grey dashed lines, we can see that the increases realized by students at most grade levels in the 2020-21 school year typically trailed pre-pandemic norms.

On Average, Elementary Students' Growth During the 2021-22 School Year Outpaced Pre-Pandemic National Norms, While Middle School Students' Growth Did Not

Similar to the 2020-21 school year, average math and reading scale scores increased between fall 2021 and spring 2022. In some cases, these increases surpassed those of similar students before the COVID-19 pandemic, resulting in average scores becoming closer to (or even exceeding) pre-pandemic norms. Average scale score growth among elementary school students in NWEA districts generally outpaced changes in the pre-pandemic norms. As a result, spring 2022 average math and reading scale scores for students in the fall 2020 kindergarten cohort remained above pre-pandemic norms, while students in the 1st- through 4th-grade cohorts drew closer to their respective pre-pandemic norms (except in 4th-grade reading). Similarly, almost all the Star 360 math, reading, and early literacy scale scores changes for elementary grade levels in 2021-22 exceeded changes in the national norm. Students in these grade levels ended the 2021-22 school year at or above national norms.

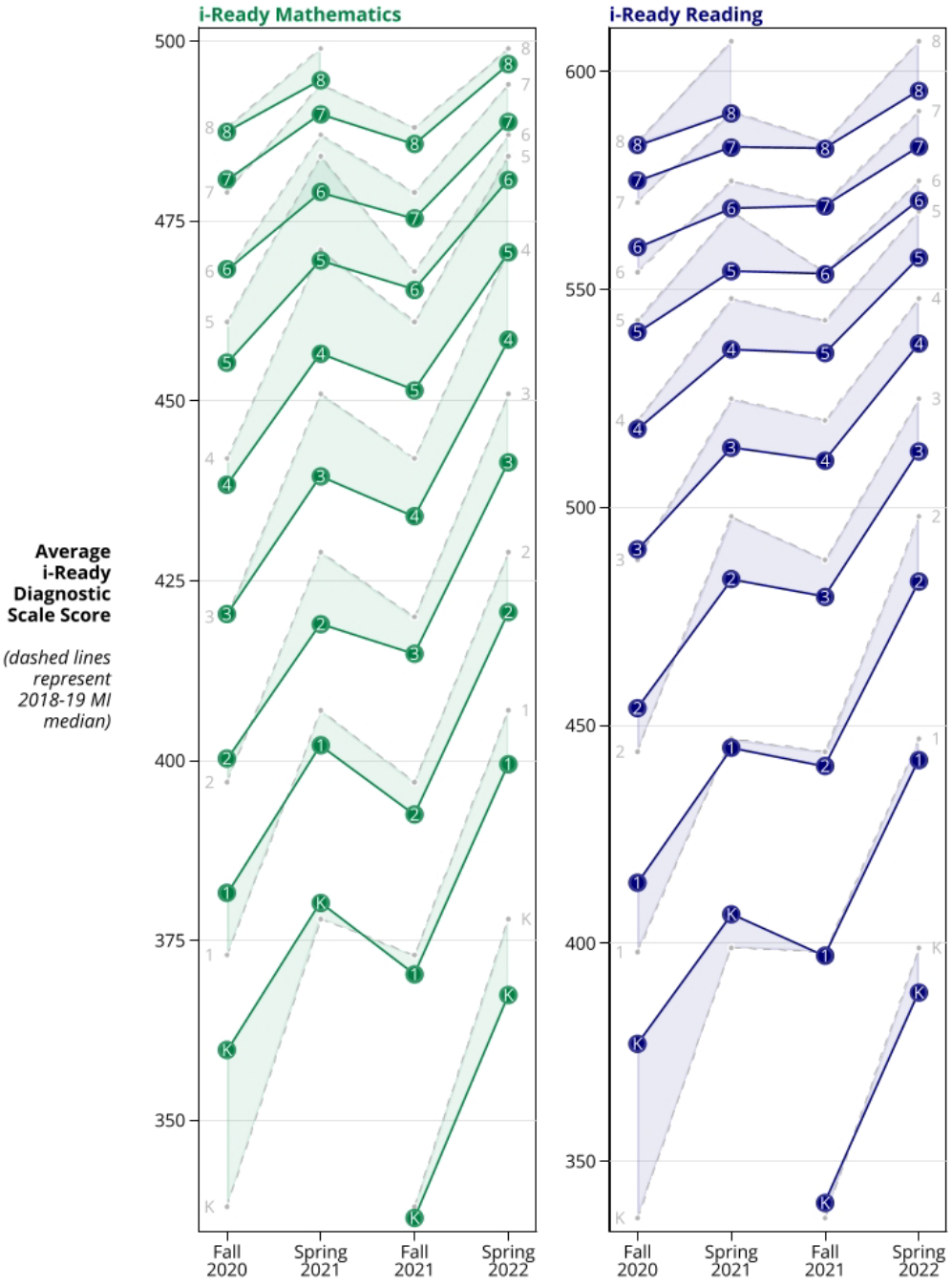
In other cases, scale scores increased at a slower rate than before the COVID-19 pandemic, and average scores fell further below pre-pandemic norms. This was particularly the case for middle school cohorts in NWEA and Star 360 districts. Average math and reading scale score growth among students in the 5th- through 7th-grade cohorts in NWEA districts trailed changes in the pre-pandemic norms and students in these grade levels fell further behind. In Star 360 districts, average math and reading scale score changes for students in the 5th-, 6th-, and 7th-grade cohorts were below national norms and disparities for these students increased during the 2021-22 school year. The fall-to-spring increases realized by i-Ready students were consistently less than pre-pandemic norms across grade levels. Hence, by spring 2022, Michigan i-Ready students in nearly all grade levels scored further below the norm than they did at the beginning of the year.

Figure 3.2.1 Trends in Average Scale Scores, NWEA MAP Growth



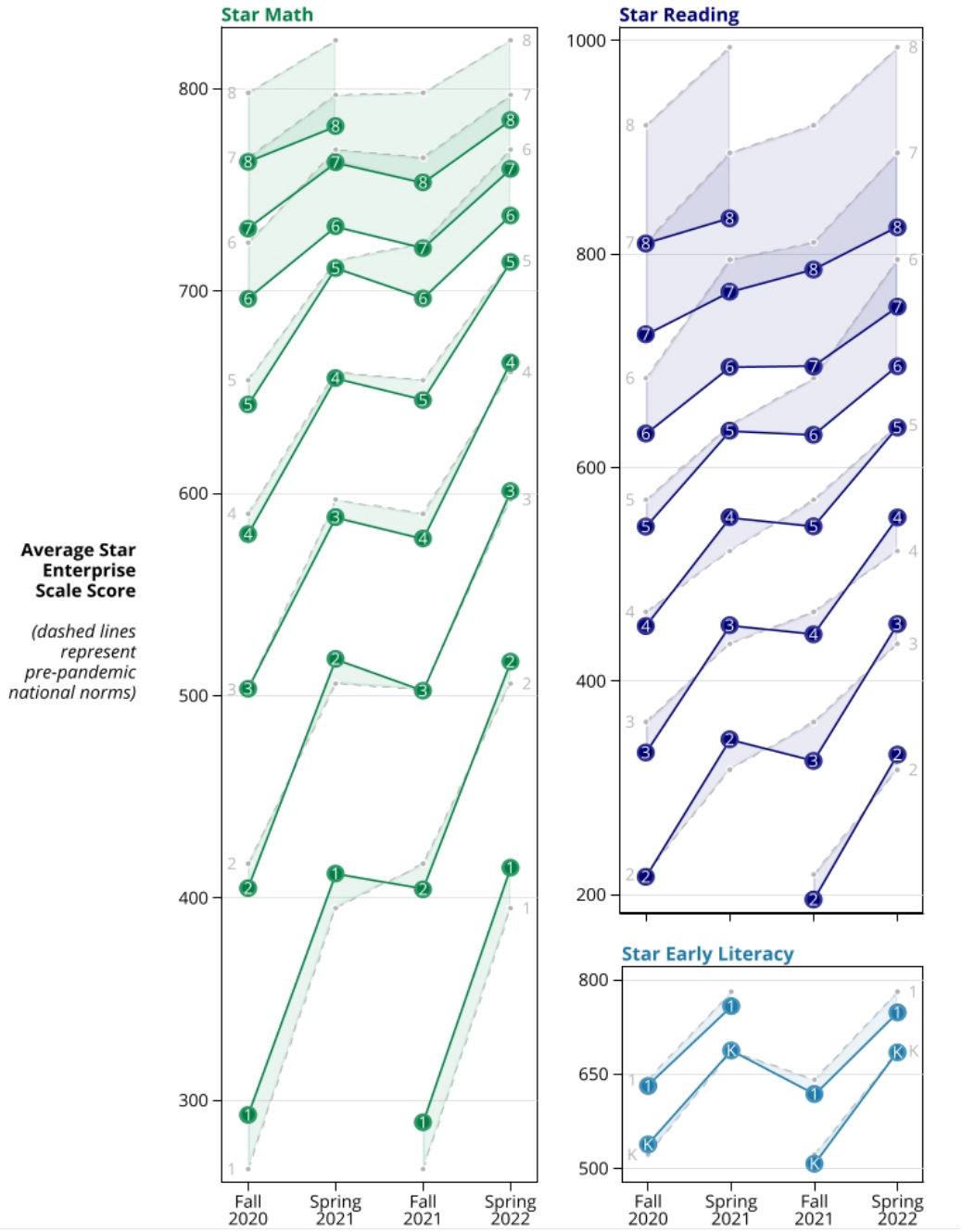
Note: These averages include only students with benchmark assessment scores for every possible testing period. The comparison points in the figure represent the 50th percentile of NWEA's conditional growth distribution. The y-axis scales range from the kindergarten fall norm to the 8th grade spring norm; these values are slightly different across subjects. RIT stands for Rasch unit scale.

Figure 3.2.2 Trends in Average Scale Scores, Curriculum Associates' i-Ready



Note: These averages include only students with benchmark assessment scores for every possible testing period. The comparison points in the figure represent median scores for Michigan students in 2018-19. The y-axis scales range from the kindergarten fall norm to the 8th grade spring norm; these values are slightly different across subjects.

Figure 3.2.3 Trends in Average Scale Scores, Renaissance Learning's Star 360



Note: These averages include only students with benchmark assessment scores for every possible testing period. The comparison points in the figure represent median scores for Renaissance Learning's norming sample. The y-axis scales range from the kindergarten fall norm to the 8th grade spring norm; these values are slightly different across subjects.

Regression-Adjusted Achievement Trends

While Figures 3.2.1 through 3.2.3 help us describe how Michigan students' performance compares to other students pre-pandemic, these unadjusted averages do not account for differences in the demographic characteristics of students in each district. Thus, we also use multiple regression to examine how Michigan districts' average scores differ from national pre-pandemic norms *after* adjusting for differences in districts' demographic composition. Prior research examining student outcomes over the past two school years has shown that students from different backgrounds or those who need specialized instruction performed differently during the COVID-19 pandemic. The regression estimates we show attempt to separate the confounding effect of these unique experiences from the temporal changes in average scale scores experienced by the average Michigan student.

In these regressions, we standardize benchmark assessment scores using the means and standard deviations from the pre-pandemic norming samples for each grade, subject, and testing period (fall or spring). This allows us to interpret the results in terms of how far a district's average achievement is from the pre-pandemic national average, relative to the standard deviation for a particular grade level. This is helpful because the standard deviations of scale scores for lower grade levels are typically much smaller than the standard deviations for upper grades. While the unadjusted average scale score trajectories in Figures 3.2.1 through 3.2.3 look substantially different in the lower versus upper grades, once we put these scores in the context of how much variation is typical for those grade levels, those differences are not as drastic.

Results from our regression analysis show how much Michigan districts' average scores differed from pre-pandemic norms, controlling for demographic characteristics of the students in each grade level within each district (i.e., gender, race/ethnicity, economically disadvantaged status, special education status, and English learner status). The regression coefficients associated with each semester- and grade-specific variable in our models describe the disparity between Michigan student achievement for a particular grade and time point relative to students in each pre-pandemic norming sample: Estimates below zero (or percentile ranks below 50) show that Michigan students were performing below average compared to students in the pre-pandemic norming sample while positive estimates (or percentile ranks above 50) show that Michigan students outperformed the average student in a particular norming sample (see Appendix Table 3 for our complete regression output).

HOW TO INTERPRET FIGURES 3.2.4 TO 3.2.5

These figures show adjusted trends in average math and reading achievement for each cohort of students, standardized relative to pre-pandemic national norms for each grade, subject, and testing period. To ease interpretation of these values, we convert each estimate to a **percentile rank** to show how Michigan student achievement compares to the distribution of test scores among students in each assessment provider's norming sample. **A percentile rank of 50 indicates that Michigan students scored at the pre-pandemic national average.**

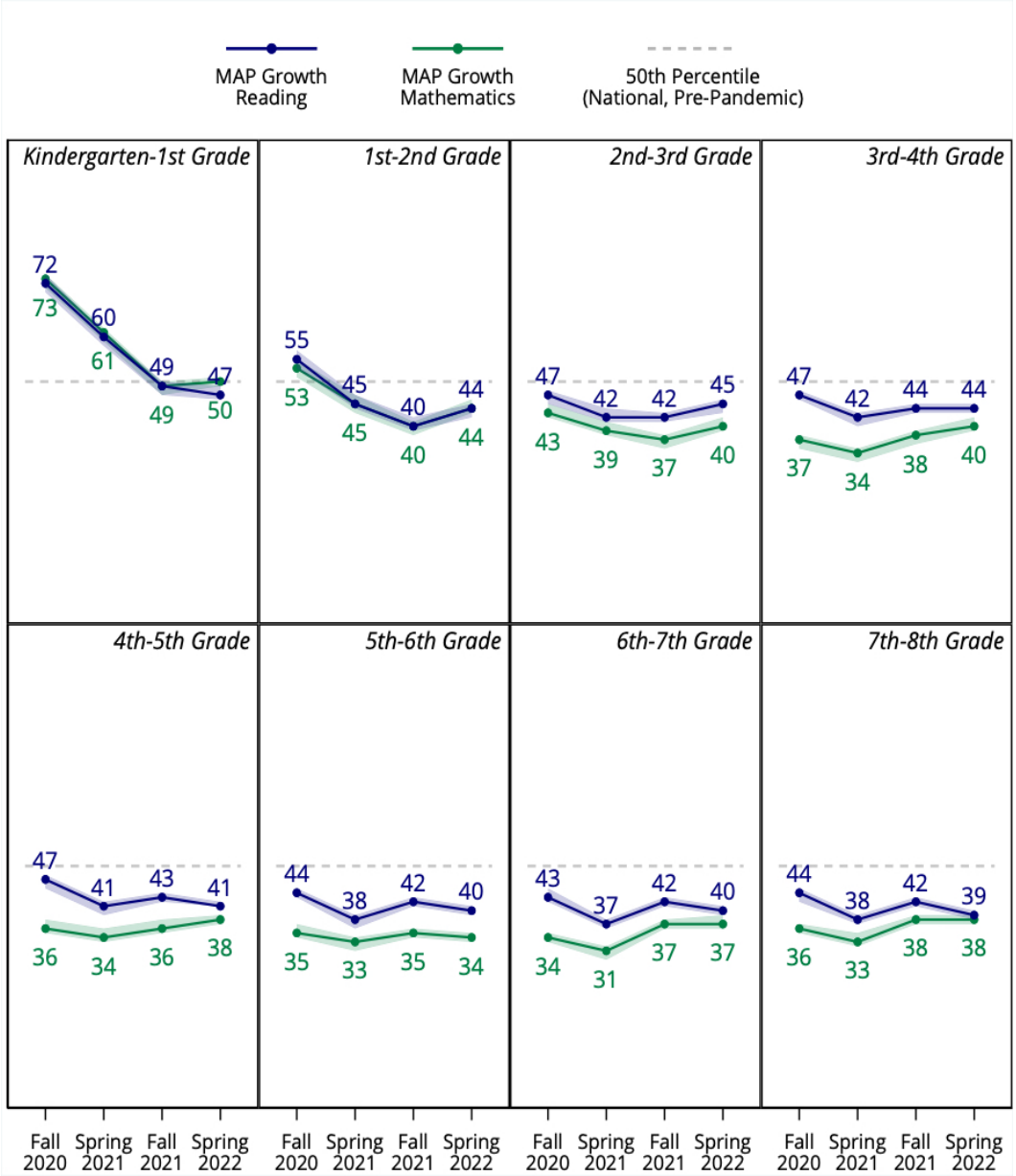
Unlike the scale score trajectory figures in the previous subsection, we would not necessarily expect standardized achievement trends to increase over time as students continue to learn. Rather, if students learn at a rate consistent with pre-pandemic norms, we would see a **flat trend line**, indicating that average scores for Michigan students maintained their percentile rank each testing period. If students learned at a slower rate than those in the norming sample, this would not be sufficient to maintain the same percentile rank, and we would therefore see a **decreasing trend**. If students experienced a decrease in relative achievement during the COVID-19 pandemic**, they would need to learn at a *faster* rate than students in the norming sample to effectively "catch up"** to where they were relative to other students before the COVID-19 pandemic.

The **shaded areas above and below each trend line** show the 95% confidence interval for each percentile rank estimate. This represents the range of values that the "true" percentile rank for Michigan students is likely to fall within, given that we can only estimate based on an incomplete sample of students and not the full population. If the shaded area overlaps with the grey dashed line, this means that the estimate is not significantly different from the pre-pandemic national average.

Average Achievement Decreased in the 2020-21 School Year Relative to Pre-Pandemic Norms

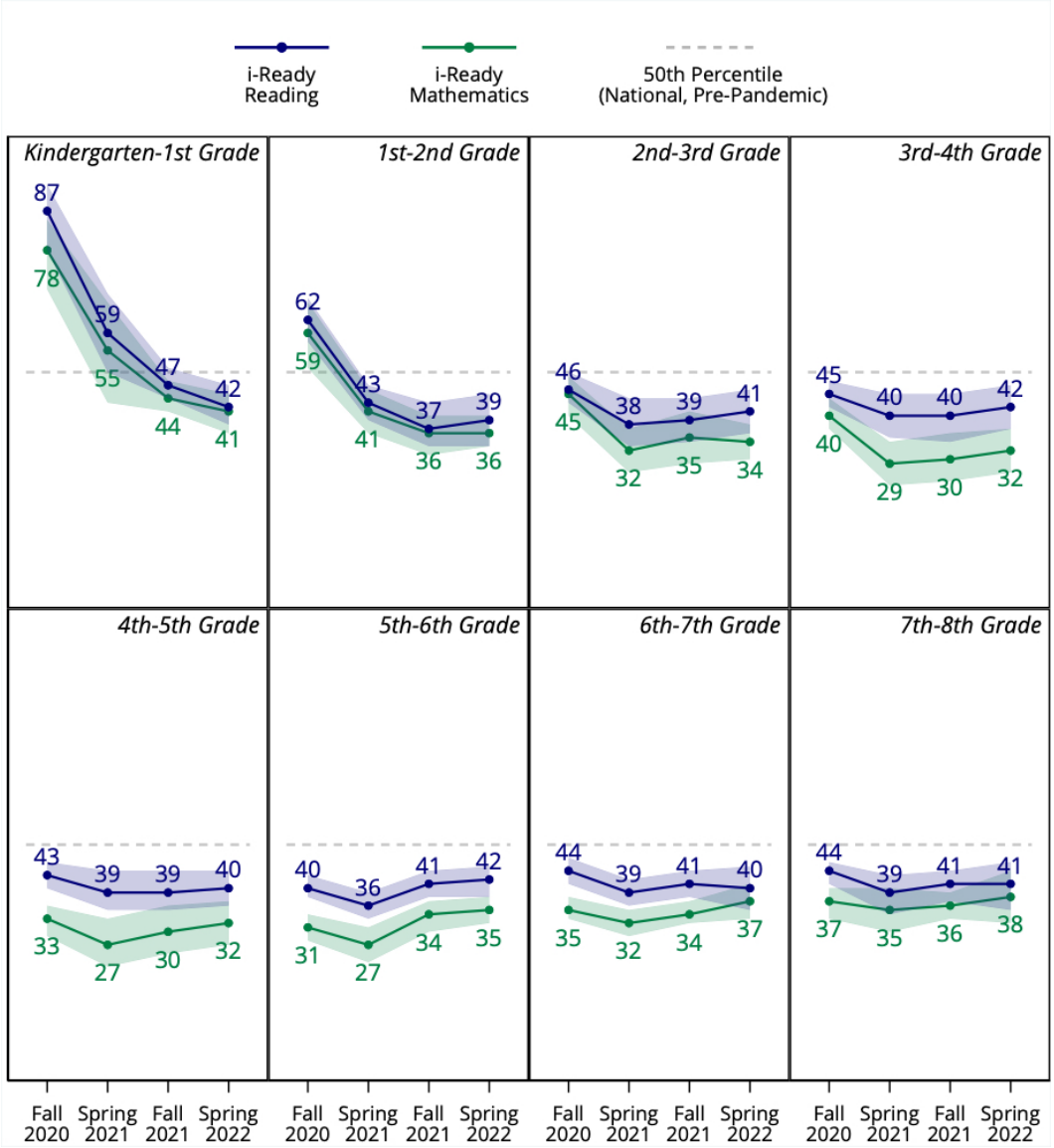
Across all grades, subjects, and assessment providers, average achievement for Michigan students decreased between fall 2020 and spring 2021 relative to pre-pandemic norms. These decreases were generally larger and more consistent in reading than in math, particularly in upper elementary and middle school grades. We again note that average scores for students in kindergarten and 1st grade in fall 2020 registered well above the 50th percentile of each respective pre-pandemic norming sample, which is likely related to the "at-home testing advantage" and not a realistic representation of their true performance at that time. Students in other grades scored below the 50th percentile of each pre-pandemic norming sample in fall 2020.

Figure 3.2.4 Regression Adjusted Percentile Ranks, NWEA MAP Growth



Note: These regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to NWEA's pre-pandemic national norms, and estimates are converted to percentile ranks.

Figure 3.2.5 Regression Adjusted Percentile Ranks, Curriculum Associates' i-Ready



Note: These regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to Curriculum Associates' pre-pandemic national norms, and estimates are converted to percentile ranks.

Average Achievement Growth in the 2021-22 School Year Varied by Grade Level

We find significant differences in achievement trends across grade levels during the 2021-22 school year. Between fall 2021 and spring 2022, students in early elementary grade levels generally exhibited some growth in both math and reading relative to pre-pandemic norms (except the cohort of students who were in kindergarten in 2020-21 and 1st grade in 2021-22). Upper elementary students experienced slight changes in math and reading achievement relative to the pre-pandemic norming sample, though these changes were inconsistent across subjects and assessment providers. Among students in districts that administered an NWEA assessment, 3rd-, 4th-, and 5th-grade reading achievement typically declined relative to the pre-pandemic norms, while similar achievement levels for students in i-Ready districts increased by one to two percentiles. Math achievement for later elementary school students in both NWEA and i-Ready districts increased marginally during the 2021-22 school year.

While the previous two figures summarize regression-adjusted achievement trends within *the same cohorts of students over time*, Figure 3.2.6 and Figure 3.2.7 compare changes in math and reading outcomes relative to the pre-pandemic norming samples for students *in the same grade across the 2020-21 and 2021-22 school years*. This comparison is important to consider in order to understand how students across the Michigan public school system were affected in a given year relative to students in the same grade level in a different year.

HOW TO INTERPRET FIGURES 3.2.6 AND 3.2.7

In these figures, the **green and blue bars** show the average performance on math and reading benchmark assessments, respectively, in the fall and spring of each full COVID-19 pandemic school year (2020-21 and 2021-22). These are converted into percentile ranks, which indicate where in the percentile distribution from 1 to 100th students scored, on average.

The **lighter shaded bars** depict percentile ranks for students who were in a particular grade level in 2020-21 and the **darker bars** show percentile ranks for students who were in the same grade in 2021-22. If a dark blue or green bar is lower than the lighter blue or green bar for the same grade level and testing period (fall or spring), this means that students who were in that grade in 2021-22 scored lower, on average, than students who were in the same grade in 2020-21. If the dark blue or green bar is higher than the light blue or green bar, this means that students who were in that grade in 2021-22 scored higher than students who were in the same grade in 2020-21.

Each **grey dashed line** marks the 50th percentile of the pre-pandemic norming sample for a particular assessment. If a bar is above the grey dashed line, this means that the average score for the group of students represented in that bar is above the 50th percentile for the pre-pandemic national norming sample. If a bar is below the grey dashed line, this means that the average score is below the 50th percentile.

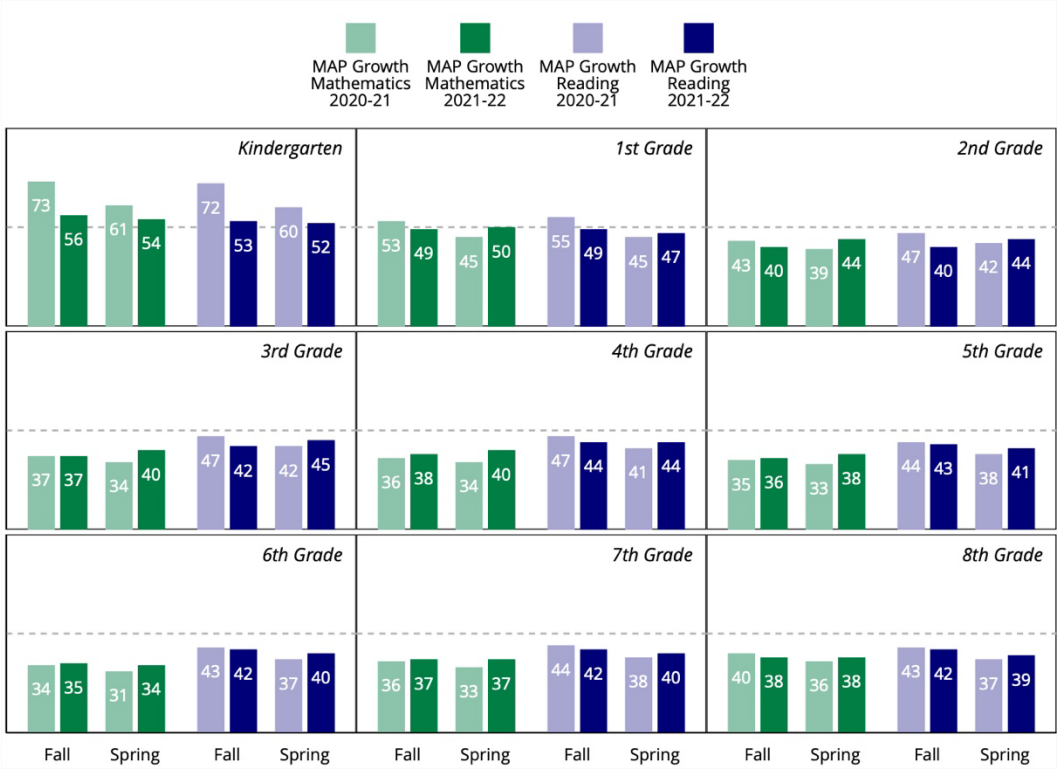
Students Generally Started the 2021-22 School Year with Lower Scores than Students Who Were in the Same Grade the Year Before

These figures show that, with few exceptions, average fall scores for students in 2021-22 were close to or below those of students who were in the same grade in 2020-21. Kindergarten and 1st -grade students started the 2021-22 school year with particularly low fall scores compared to students who were in the same grade the year before. While this is likely driven at least in part by the "at-home testing advantage" for the 2020-21 cohort, we see the same pattern—although to a lesser extent—for students in grade levels that were not affected in the same ways by remote testing in fall 2020. In fact, average fall scores for *all* grade levels on the MAP Growth reading, i-Ready reading, and i-Ready math assessments were lower in 2021-22 than they were in 2020-21. This likely reflects the substantial interruptions students experienced during the 2020-21 COVID-19 pandemic school year, which caused the majority of Michigan students to perform at lower levels in fall 2021 than students in the same grades had performed the year before.

Students Improved Relative to Pre-Pandemic Norms in 2021-22, Ending the Year Ahead of Students Who Were in the Same Grade in 2020-21

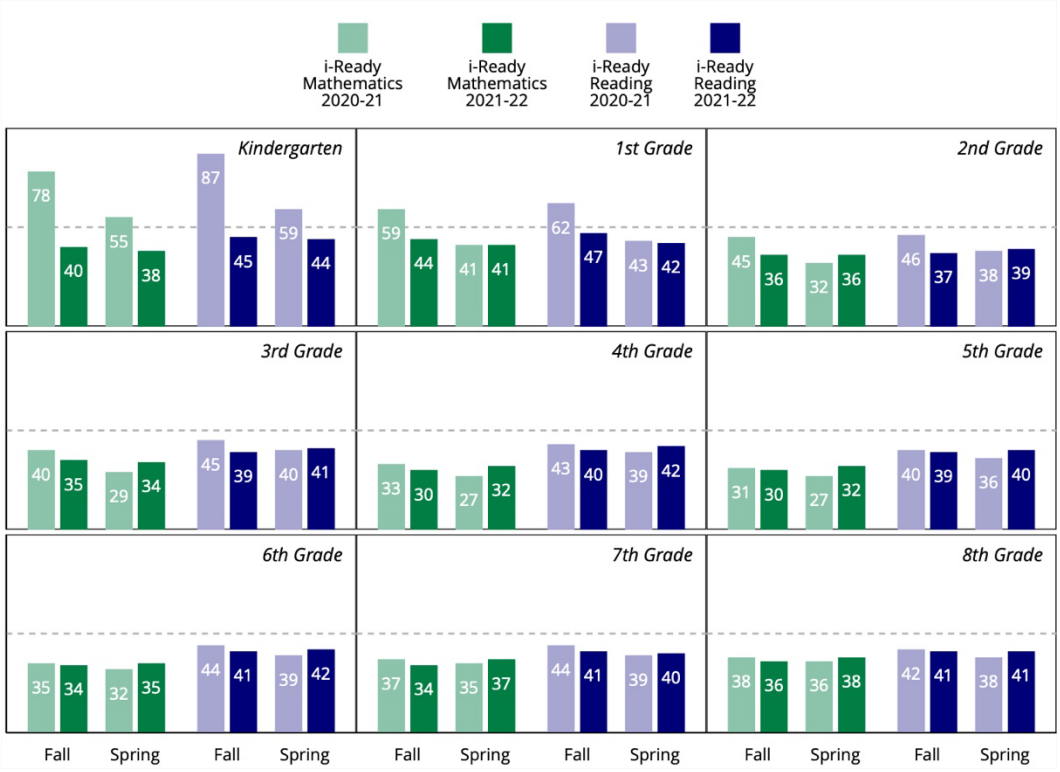
Despite often starting behind students from the prior cohort, students experienced greater achievement growth over the course of the 2021-22 school year, relative to pre-pandemic national norms. While the 2020-21 cohort generally saw decreases in relative performance between the fall and spring testing periods, the 2021-22 cohort maintained or improved their percentile ranks and ultimately surpassed 2020-21 achievement levels in nearly all grade levels by spring 2022. The only exceptions to this were students in kindergarten and 1st grade, who in some cases ended the 2021-22 school year with the same or lower percentile ranks as kindergarten and 1st grade students in 2020-21. However, again, it is difficult to interpret achievement trends for these grade levels, as fall scores for the 2020-21 cohort were likely inflated due to "at-home testing advantage."

Figure 3.2.6. Regression Adjusted Percentile Ranks by Grade Cohort, NWEA MAP Growth



Note: These regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to NWEA's pre-pandemic national norms, and estimates are converted to a percentile rank.

Figure 3.2.7. Regression Adjusted Percentile Ranks by Grade Cohort, Curriculum Associates' i-Ready



Note: These regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to Curriculum Associates' pre-pandemic national norms, and estimates are converted to a percentile rank.

PERCENTAGE OF STUDENTS ACHIEVING TYPICAL YEARLY GROWTH

To understand the proportion of Michigan students that achieved "typical" growth over the course of each school year, where "typical" is defined by assessment providers' expectations of growth in a non-pandemic school year, we examine the distributions of Michigan students across three categories of growth outcomes:

1. **met or exceeded typical growth** (i.e., the increases to their scale score met or surpassed the growth norm for students in their grade level with similar prior achievement scores);

2. **made less than typical growth** (i.e., their scale scores increased by less than the pre-pandemic growth norm); or

3. **did not demonstrate growth** (i.e., their scale scores either did not change or decreased from fall to spring).

We show these percentages for all students in the 2020-21 and 2021-22 school year growth samples and provide additional results disaggregated by grade, initial fall 2020 or fall 2021 achievement level, student demographic characteristics (i.e., gender, race/ethnicity, economically disadvantaged and special education status), and mode of instruction. For students who made partial progress toward—but did not reach—the appropriate pre-pandemic growth norm, we show the full distribution of partial growth in terms of the percentage (i.e., between 1% and 99%) of a typical year's growth that students achieved each year. Finally, we examine end-of-year achievement levels in terms of M-STEP proficiency equivalencies (established by the assessment providers) among students in each of the three growth categories to better understand which students are represented in each category.

Due to low subgroup counts among students who completed a Star 360 assessment, we present the overall percentages of students in each typical growth category for students who completed an NWEA, i-Ready, or Star 360 assessment but only show subgroup-specific results for students in districts that administered an NWEA or i-Ready assessment.

Overall Results

Figures 3.3.1 through 3.3.3 show the percentages of Michigan students who met or exceeded typical growth, made less than typical growth, or did not demonstrate growth on an NWEA, i-Ready, or Star 360 assessment administered during the 2020-21 and 2021-22 school years. Additionally, Figure 3.3.1 shows the percentages of students from NWEA's pre-pandemic norming sample that fall into each of these growth categories to compare Michigan's outcomes to students nationally.[4]

We note that the terminology of "typical growth" or "growth targets" used in this report differs in meaning from the way practitioners use similar terms, such as "growth goals," in the classroom. In classroom contexts, teachers likely set student growth goals or targets that represent what the teacher expects an individual student to achieve in a given period. This is different from the targets for "typical growth" that we use in this report, which indicate *the median growth that students with similar prior scores achieved before the COVID-19 pandemic*.

HOW TO INTERPRET FIGURES 3.3.1 TO 3.3.3 AND FIGURES 3.3.8 TO 3.3.26

Within these figures, the **dark green and blue bars** represent the percentages of students who met or exceeded typical growth in math and reading, respectively, in a given year. The **lighter shades of green and blue** represent students who achieved less than typical growth. The **grey bars** represent students who did not demonstrate any growth at all.

For the MAP Growth and i-Ready assessments, a student's "typical growth" is the median growth across students with similar baseline scores from the pre-pandemic national norming samples for each assessment. For the Star assessments, it is the amount of growth necessary for students to maintain their fall percentile rank. Thus, **in a typical year, we would expect to see about 50% of students meet or exceed their "typical growth."**

Our previous report showed that in 2020-21, Michigan students were less likely to reach or exceed typical growth and more likely to demonstrate no growth at all, compared to students in the pre-pandemic norming sample. If the same pattern continued in 2021-22, this would mean that students fell *further behind*, on average. If growth outcomes returned to the pre-pandemic norm in 2021-22, this would mean that students remained as far behind as they were the previous year. **Michigan students would need to surpass pre-pandemic growth rates to "catch up" to national norms.** In other words, more than 50% of students would need to meet or exceed typical growth in the 2021-22 school year to "catch up" after the slower than typical growth experienced in 2020-21.

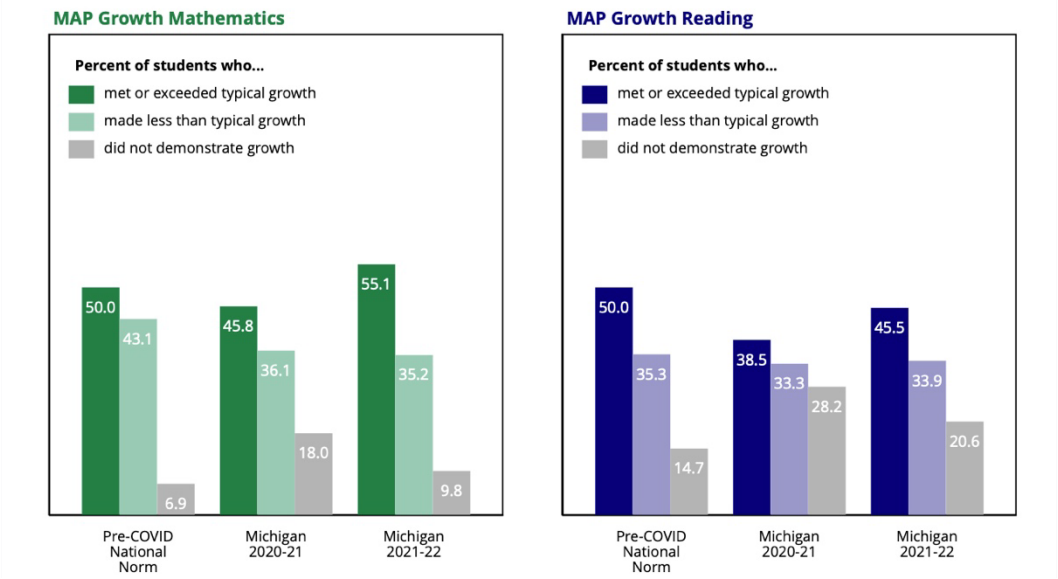
In 2021-22 More Students Met or Exceeded Typical Growth and Fewer Demonstrated No Growth Than in the Year Prior

As discussed in our previous report, Michigan students who took NWEA's MAP Growth assessments during the 2020-21 school year were less likely to meet or exceed typical growth and more than twice as likely to demonstrate no growth than would have been expected before the COVID-19 pandemic. Results for the 2021-22 school year, shown in Figure 3.3.1, are not quite as bleak. Greater percentages of students met or exceeded typical growth in 2021-22 than in 2020-21, increasing from 46% to 55% (now five percentage points above the pre-pandemic norm) in math and from 39% to 46% in reading (closer to but still below the pre-pandemic norm). In addition, the percentages of students who did not demonstrate growth in 2021-22 were far lower than in the previous year, decreasing from 18% to 10% in math and from 28% to 21% in reading. However, these percentages are still substantially higher than would have

been expected before the COVID-19 pandemic (about 7% and 15% in math and reading, respectively).

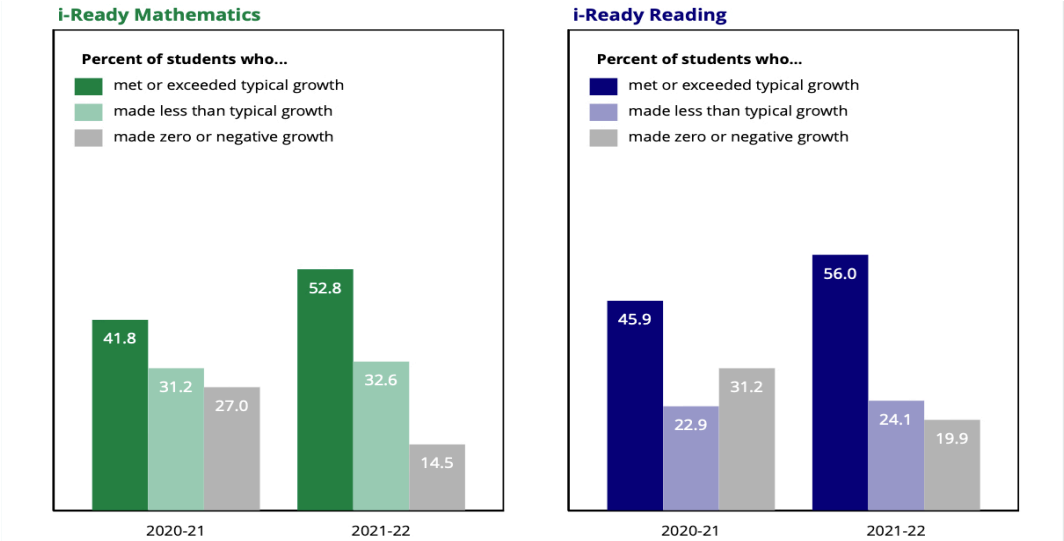
Figure 3.3.2 and Figure 3.3.3 show similar patterns for students who took the i-Ready and Star 360 assessments, respectively. However, the increases in the share of students meeting or exceeding typical growth and decreases in the share of students demonstrating no growth are consistently more pronounced for students who took the i-Ready assessments than for other assessments, particularly relative to those who took the Star 360 assessments. A possible explanation for this—which we will explore in greater detail later in this section—is that i-Ready districts, on average, remained fully remote for much of the 2020-21 school year, while Star 360 districts were primarily in-person. Thus, we may expect to see more pronounced changes in i-Ready results after these districts returned to in-person instruction in 2021-22, compared to Star 360 districts where the mode of instruction was more consistent across the two years.

Figure 3.3.1. Yearly Growth in 2020-21 and 2021-22, MAP Growth



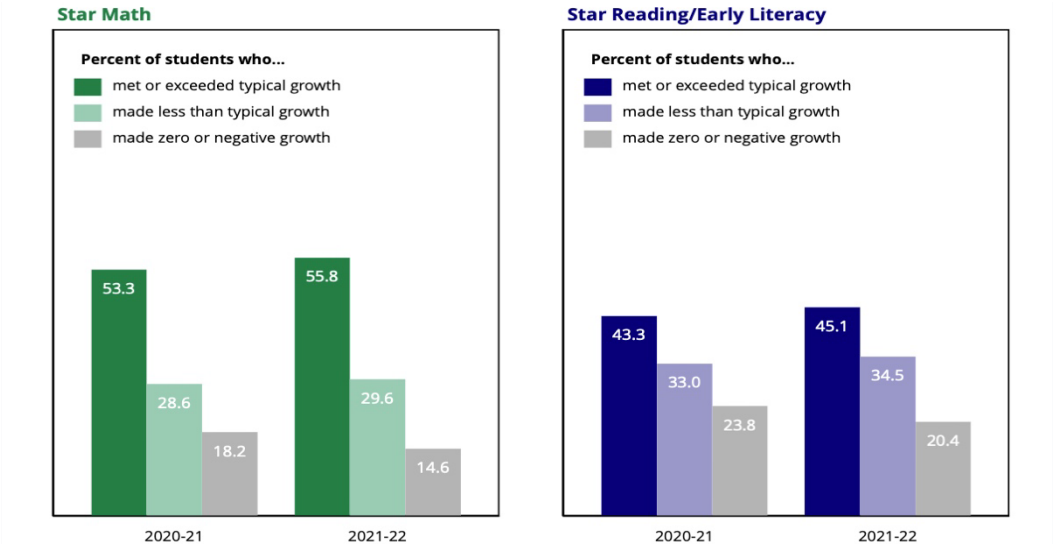
Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. We determine growth expectations using the 50th percentile of NWEA's fall-to-spring conditional growth distribution for each grade level, subject, and students' initial achievement level (Thum & Kuhfeld, 2020).

Figure 3.3.2. Yearly Growth in 2020-21 and 2021-22, i-Ready



Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. Typical growth targets come from the i-Ready assessment growth models which indicate median math and reading growth for students at each grade level and initial achievement level (Curriculum Associates, 2020).

Figure 3.3.3. Yearly Growth in 2020-21 and 2021-22, Star 360



Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. We determine growth expectations for students at each grade level, subject, and students' initial achievement level using the change in scale scores necessary for a student to maintain the same percentile rank as they did in the previous year (Renaissance Learning, 2021b, c, d).

Distribution of Partial Growth

Despite these sometimes sizable shifts from 2020-21 to 2021-22 in the proportions of students meeting or exceeding typical growth and the proportions of students not demonstrating any growth, we see relatively little change in the proportions of students making less than typical growth. In both years, about one-quarter to one-third of all Michigan students (depending on the subject and assessment provider) demonstrated growth but did not reach the typical amount of growth for students in their grade with similar initial scores. To better understand the group of students represented by the light green and blue bars in Figures 3.3.1 to 3.3.3, we look more closely at *how much* of a typical year's growth they achieved.

We focus just on the MAP Growth and i-Ready assessments for this analysis, as the number of students who participated in each of these assessments is large enough for us to identify and compare subgroups who achieved various proportions of a typical year's growth.

HOW TO INTERPRET FIGURES 3.3.4 TO 3.3.7

These figures show the distribution of growth for students who achieved less than typical growth (i.e., those students represented by the light-colored bars in Figures 3.3.1 through 3.3.3). These students learned somewhere between 1 and 99% of what was typical in a year prior to the COVID-19 pandemic. The **x-axis** represents the percentage of a typical year's growth achieved and the **y-axis** represents the percentage of students (among those who achieved less than typical growth) whose growth fell within each range.

The **green and blue bars** represent the distributions of math and reading growth outcomes, respectively, for students who achieved less than typical growth in 2021-22. The **grey bars** represent the distributions for students who achieved less than typical growth in 2020-21. The **dashed lines** show the average percentages of yearly growth for students in each of the two years.

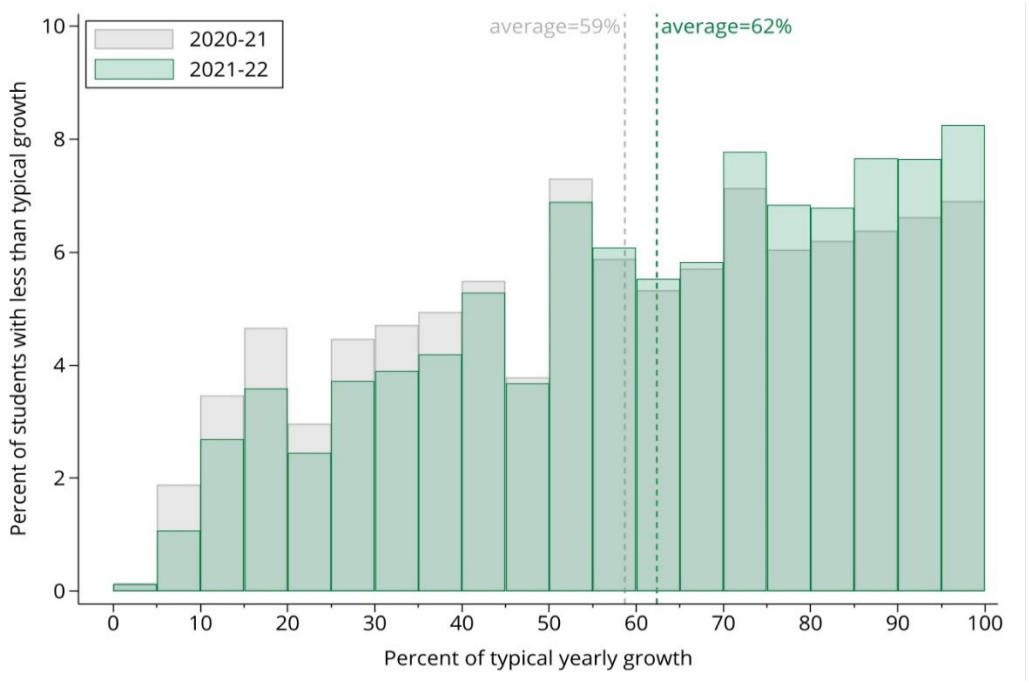
Any differences in the colored relative to the grey distributions imply a shift in the percentage of typical yearly growth realized by students in each respective school year.

Of Students Who Made Less Than Typical Growth, Average Learning Gains Were Greater in 2021-22 Than in 2020-21

Overall, we find similar achievement trends across students who made less than typical growth on the MAP Growth and i-Ready assessments. In both 2020-21 and 2021-22, the average partial growth student achieved more than 50% of typical yearly growth in math and reading, meaning that they learned more than half of a typical year's worth of tested content. Average achievement for these students improved over time, with students achieving a higher percentage of typical math and reading growth in 2021-22 compared to the previous school year. For instance, students on average learned 62% of the year's math content tested on the NWEA MAP Growth assessment in 2021-22, relative to 59% the year prior. In reading, students on average learned 59% of the year's content relative to 56% the in 2020-21.

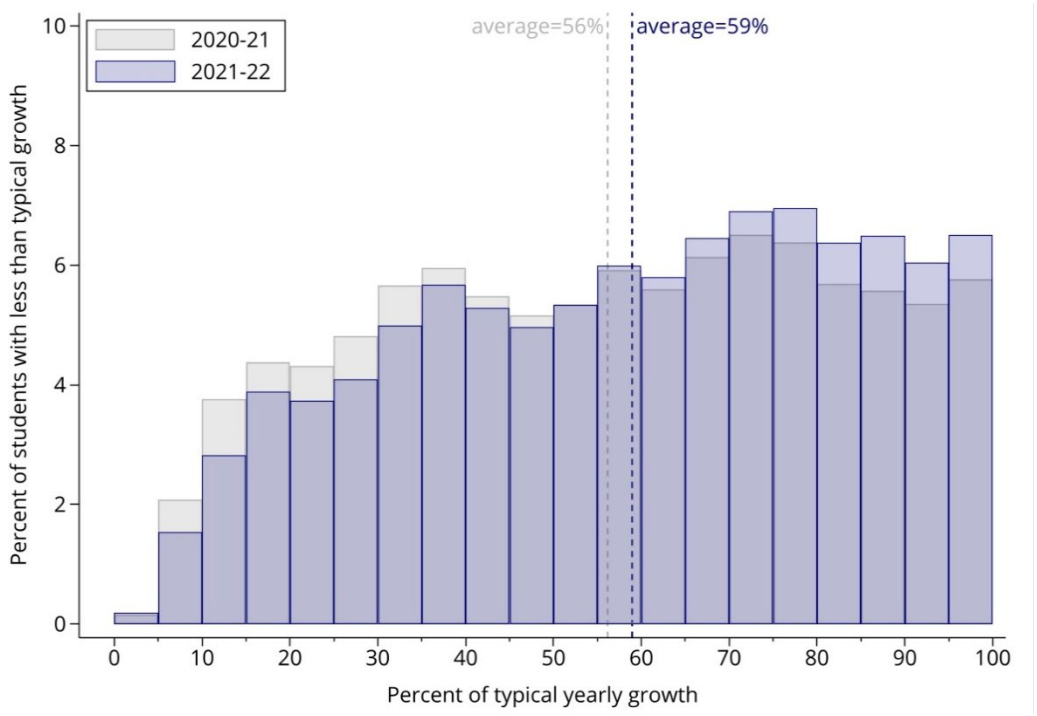
This change is driven by a shift in the distributions of partial growth that is common across both subjects and assessment providers. By comparing the relative heights of the grey and colored bars in each figure, we see that more students made less than 50% of typical growth in each subject during the 2020-21 school year whereas more students made more than 50% of typical growth in 2021-22.

Figure 3.3.4. Distribution of Partial Growth, MAP Growth Math



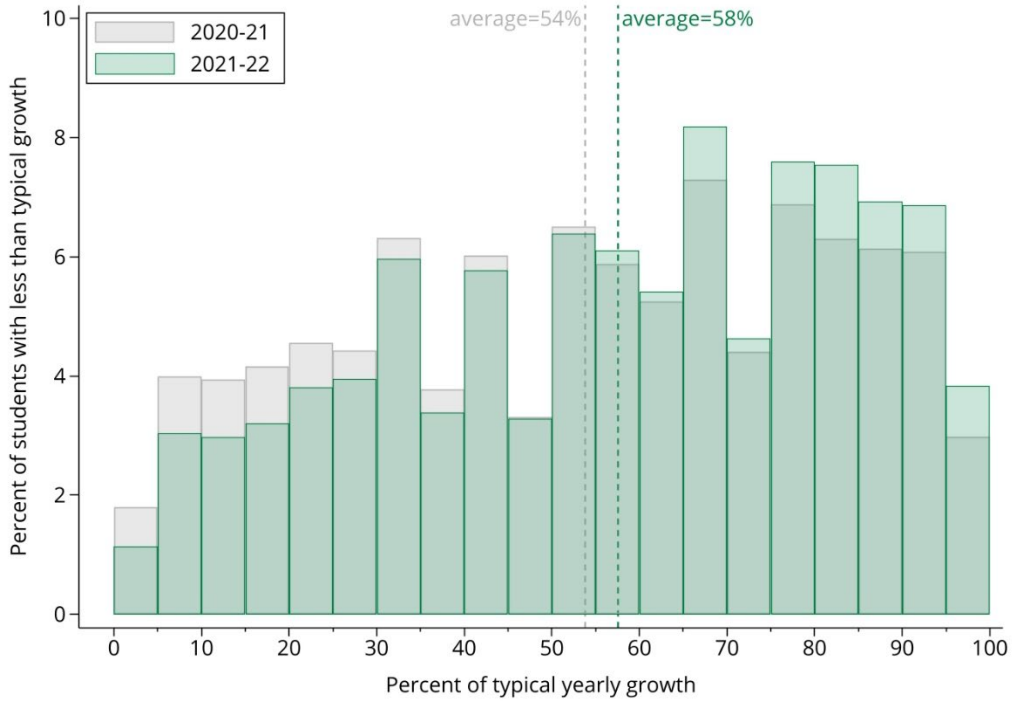
Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22 who made less than typical growth in that school year. We determine growth expectations using the 50th percentile of NWEA's fall-to-spring conditional growth distribution for each grade level, subject, and students' initial achievement level (Thum & Kuhfeld, 2020).

Figure 3.3.5. Distribution of Partial Growth, MAP Growth Reading



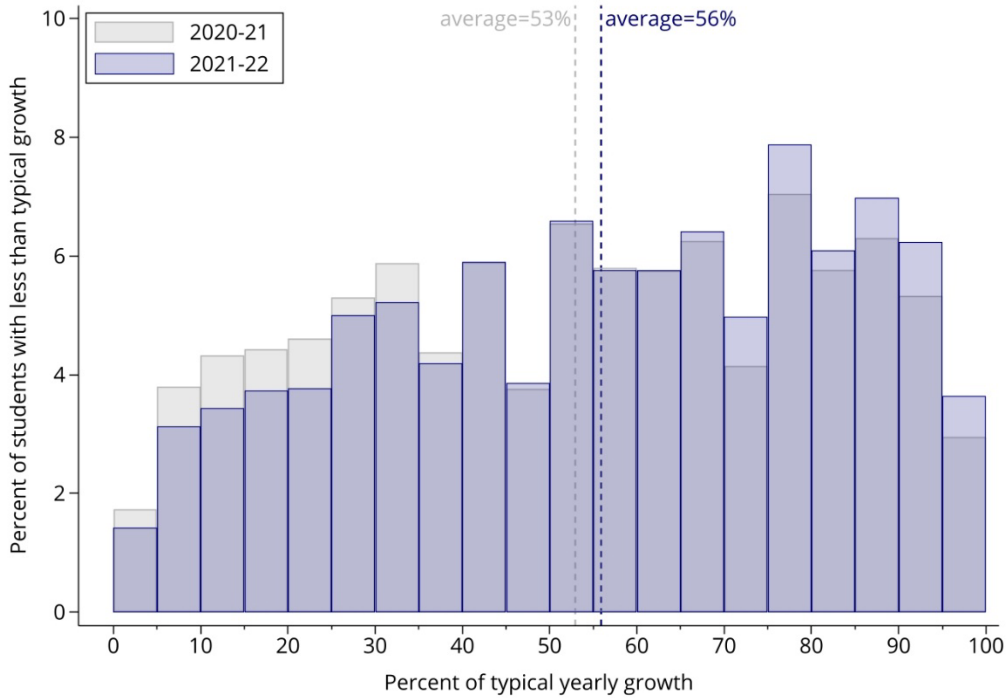
Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22 who made less than typical growth in that school year. We determine growth expectations using the 50th percentile of NWEA's fall-to-spring conditional growth distribution for each grade level, subject, and students' initial achievement level (Thum & Kuhfeld, 2020).

Figure 3.3.6. Distribution of Partial Growth, i-Ready Math



Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22 who made less than typical growth in that school year. Typical growth targets come from the i-Ready assessment growth models which indicate median math growth for students at each grade level and initial achievement level (Curriculum Associates, 2020b).

Figure 3.3.7. Distribution of Partial Growth, i-Ready Reading



Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22 who made less than typical growth in that school year. Typical growth targets come from the i-Ready assessment growth models which indicate median reading growth for students at each grade level and initial achievement level (Curriculum Associates, 2020b).

Yearly Growth by Grade Level

While the above analyses examine the proportion of Michigan students *overall* who met or exceeded typical growth, made less than typical growth, or demonstrated no growth during each of the full COVID-19 pandemic school years (2020-21 and 2021-22), Figures 3.3.8 and 3.3.9 show these percentages separately for students in each grade level. Because the "at-home testing advantage" for early elementary students in 2020-21 makes it difficult to interpret growth for students in kindergarten through 2nd grade, we also examine growth patterns by combinations of grade levels and initial fall performance levels (Figures 3.3.10 through 3.3.13) in each year.

After Experiencing Greater Growth in 2021-22, Upper Elementary and Middle School Students are Beginning to "Catch Up" to Pre-Pandemic Norms, Particularly in Math

Results from the previous section showed that, on average, students were more likely to meet or exceeded typical growth and less likely to not demonstrate any growth in 2021-22 compared to the previous school year. Figure 3.3.8 and Figure 3.3.9 show that

this holds within each individual grade level. Moreover, in most grade levels, the percentages meeting or exceeding typical growth in math in 2021-22 were not only higher than in 2020-21, but often above the pre-pandemic norm of 50%. This was also the case in reading for students taking the i-Ready reading assessment. This suggests that these groups of students are likely making progress toward "catching up" to national norms established prior to the COVID-19 pandemic.

Although Lower Elementary Students No Longer Had an "At-Home Testing Advantage" in 2021-22, They Still Made Less than a Typical Year's Achievement Growth

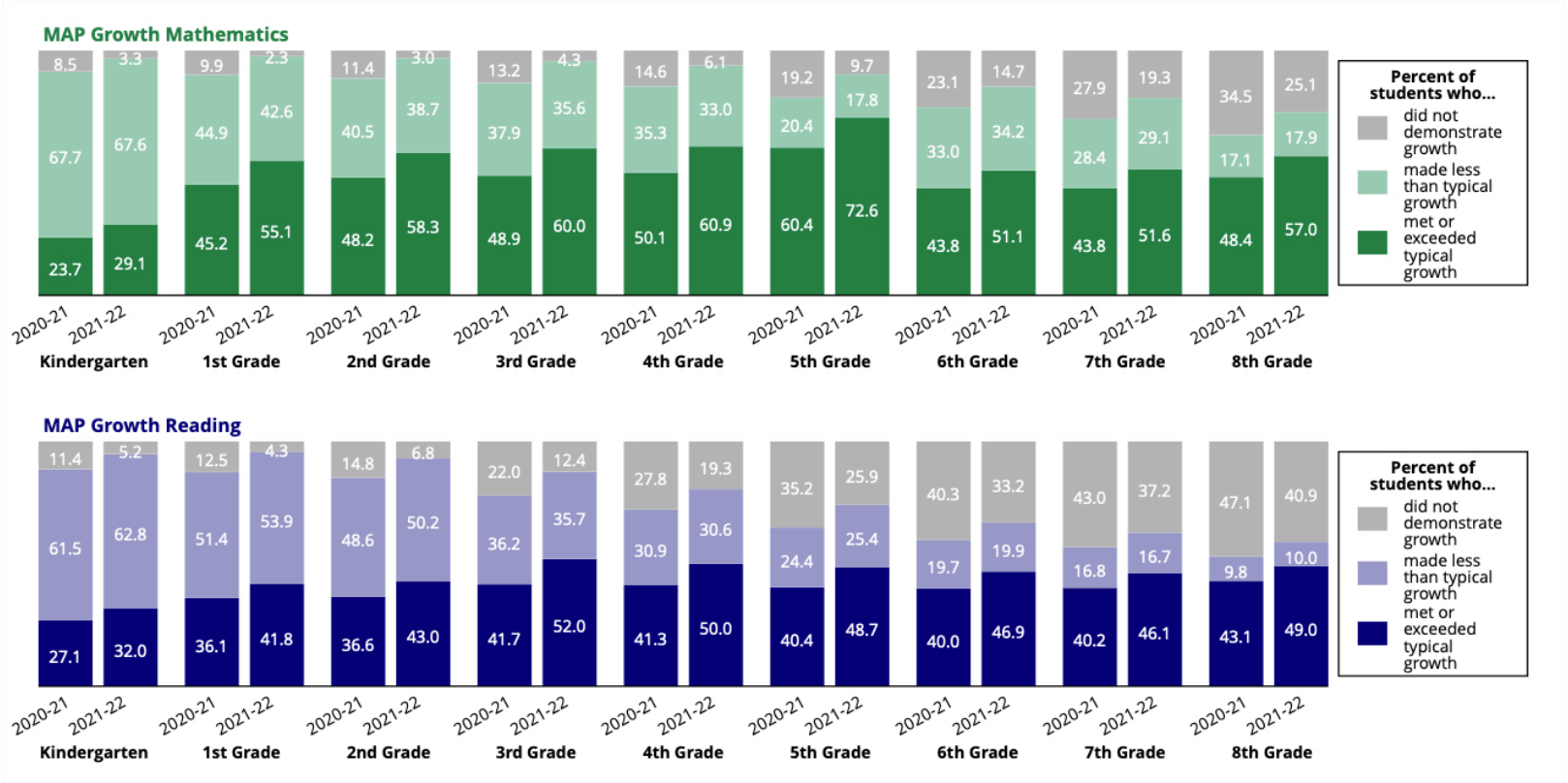
Notably, lower elementary students were the least likely to meet or exceed typical growth in either year. This was true across all subjects and assessment providers. Although larger shares of these students met or exceeded typical growth in 2021-22 than in the previous year, the percentages were still generally lower than 50%. This means that even though growth outcomes for these students improved, they did not improve enough to begin "catching up" to the pre-pandemic norm.

As discussed in our previous report, some K-2 students who completed their benchmark assessments remotely experienced an "at-home testing advantage." Fall 2020 benchmark assessment scores for these students were likely not a true reflection of their achievement at that time, possibly because they received assistance from caregivers as they were taking assessments administered virtually. As a result, many were unable to demonstrate growth throughout the 2020-21 school year based on their initial (unrealistically high) achievement levels. This is reflected in Figures 3.3.10 through 3.3.13; early elementary students in the top performance level for fall 2020 were far more likely than any other students to demonstrate no growth in 2020-21.

This "at-home testing advantage" should no longer have existed during the 2021-22 school year once in-person instruction became the predominant mode of instruction and most students took their fall assessments in-person without caregiver assistance. We therefore should expect fewer early elementary students with the highest initial fall scores to demonstrate no growth in the 2021-22 school year relative to 2020-21 and the differences in the percentages of students demonstrating no growth between initial fall achievement levels should diminish. As expected, far fewer early elementary students with the highest initial benchmark assessment scores in fall 2021 demonstrated no growth over the 2021-22 school year and they were only marginally more likely than their peers to do so. A significant portion of i-Ready students in the top fall 2021 math performance level demonstrated no growth during the 2021-22 school year (between 10% and 20% across grade levels), however, these shares are much smaller than 2020-21 and the disparities between the top two performance levels are also much smaller than the prior year.

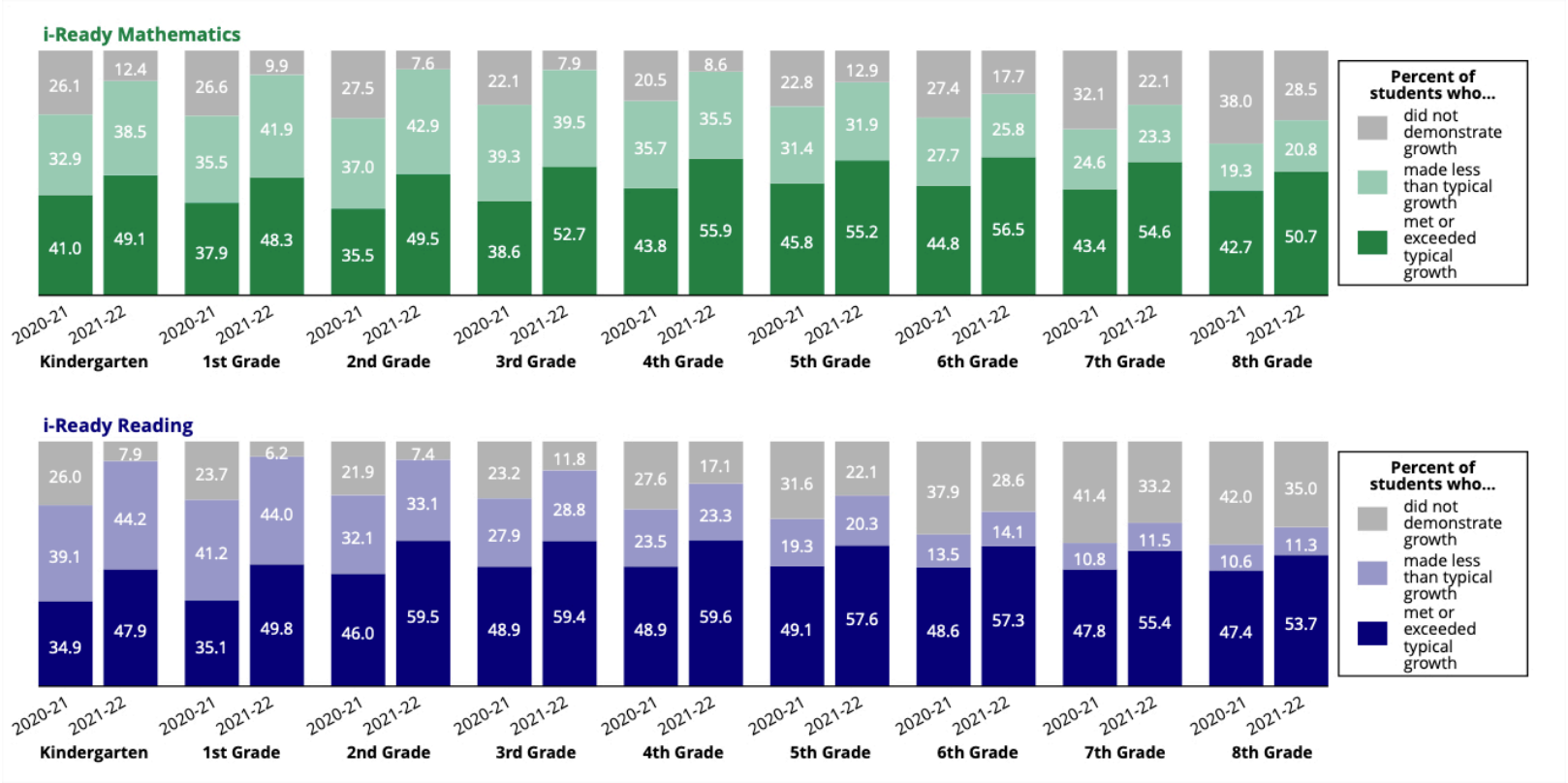
However, given that the "at-home testing advantage" no longer explains why Michigan's youngest students are so much less likely to meet or exceed typical growth than students in any other grade, it is concerning that this pattern persisted into the 2021-22 school year, in particular in districts administering the NWEA MAP Growth assessments. Put plainly, these patterns suggest that early elementary students in districts offering the NWEA assessment are for the most part not demonstrating enough progress to "make up" for interrupted learning in the 2020-21 school year. In the case of kindergarteners, who appear the most affected in 2021-22, these disparities may be related to other interruptions that resulted from the COVID-19 pandemic, including limited access to preschool and decisions about whether to delay kindergarten enrollment.

Figure 3.3.8. Yearly Growth in 2020-21 and 2021-22 by Grade, NWEA MAP Growth



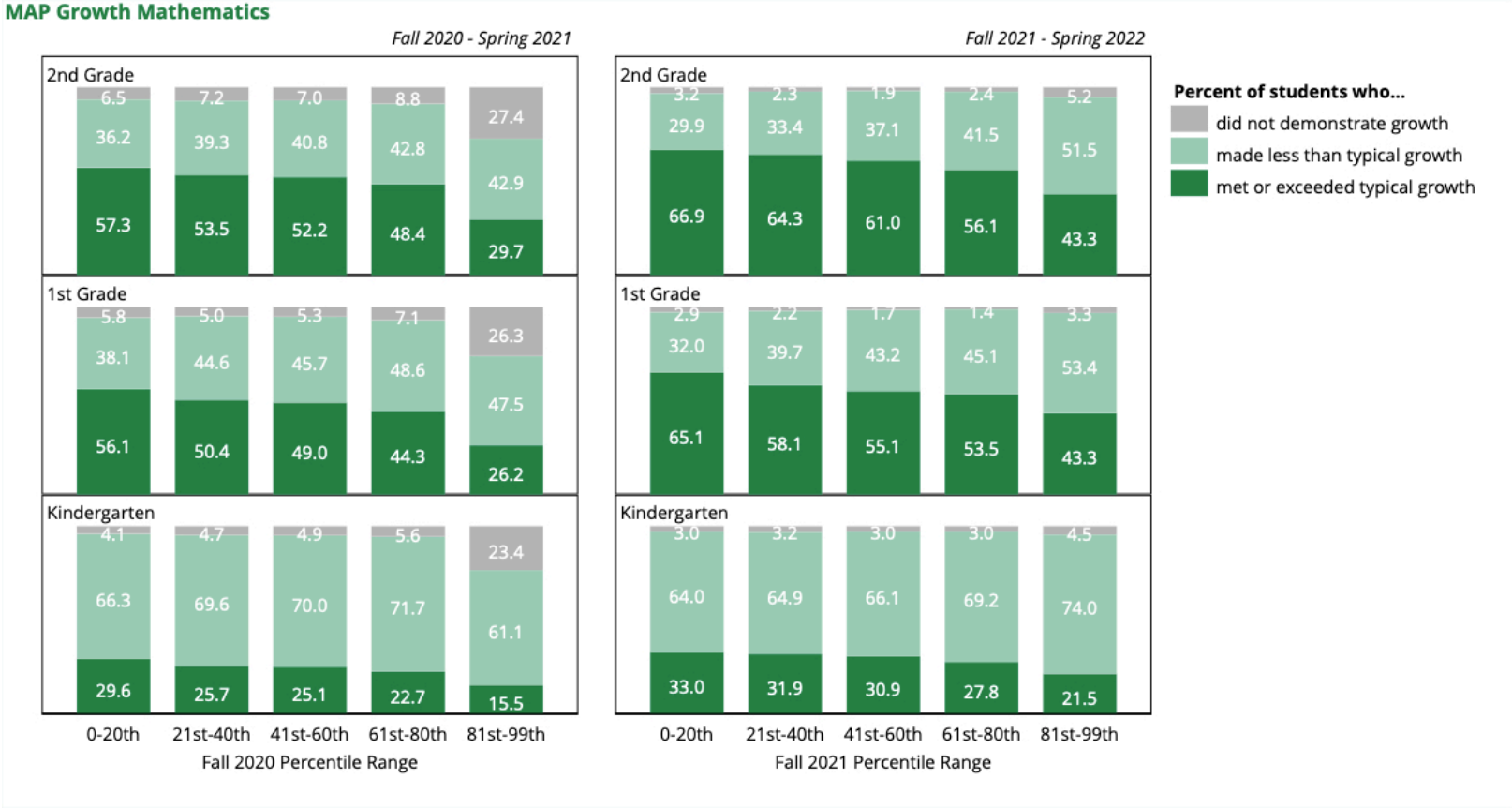
Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. We determine growth expectations using the 50th percentile of NWEA's fall-to-spring conditional growth distribution for each grade level, subject, and students' initial achievement level (Thum & Kuhfeld, 2020).

Figure 3.3.9. Yearly Growth in 2020-21 and 2021-22 by Grade, Curriculum Associates' i-Ready



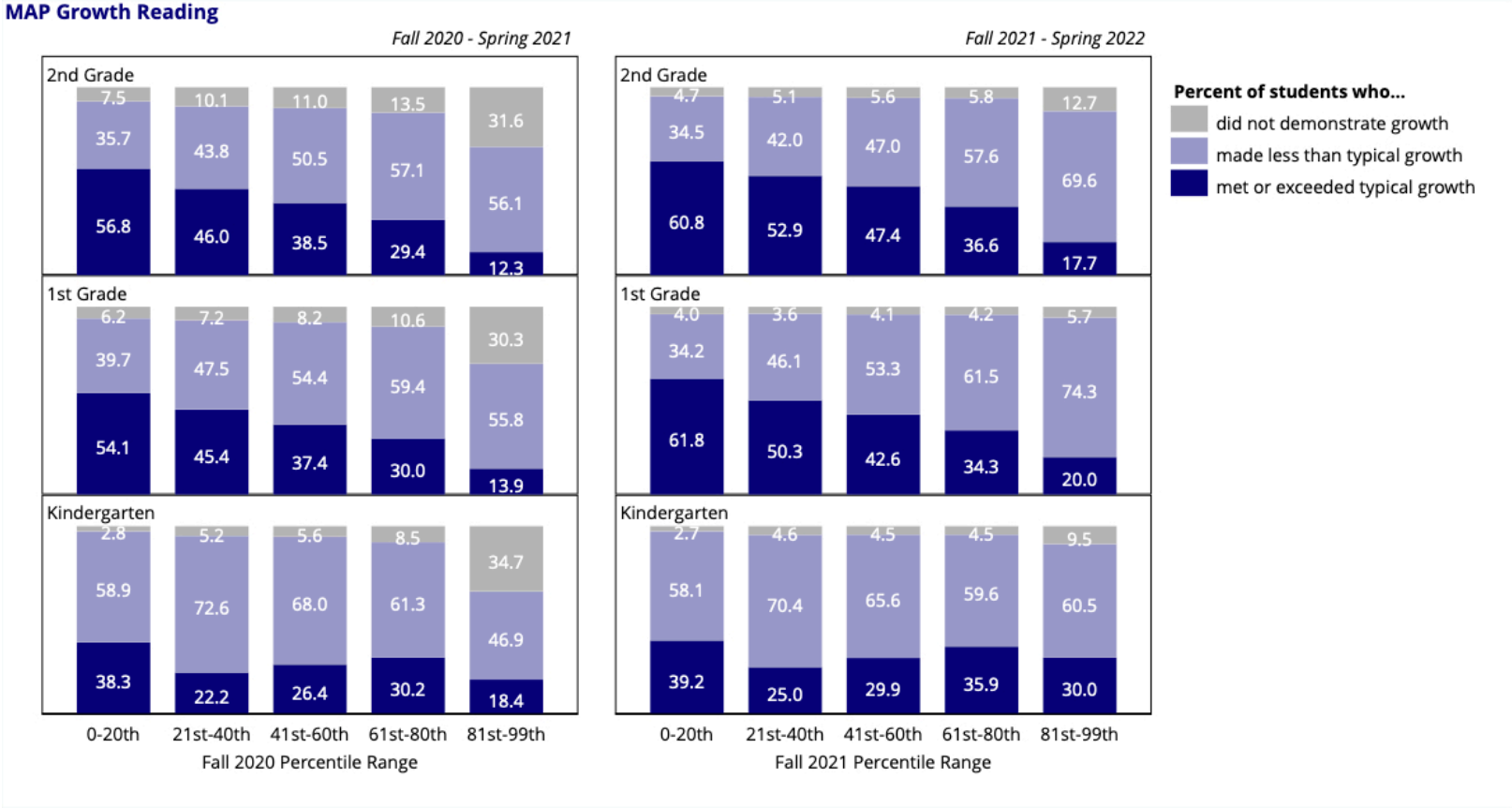
Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. Typical growth targets come from the i-Ready assessment growth models which indicate median math and reading growth for students at each grade level and initial achievement level (Curriculum Associates, 2020b).

Figure 3.3.10. Yearly Growth in 2020-21 and 2021-22 by Initial Performance Level, K-2 MAP Growth Math



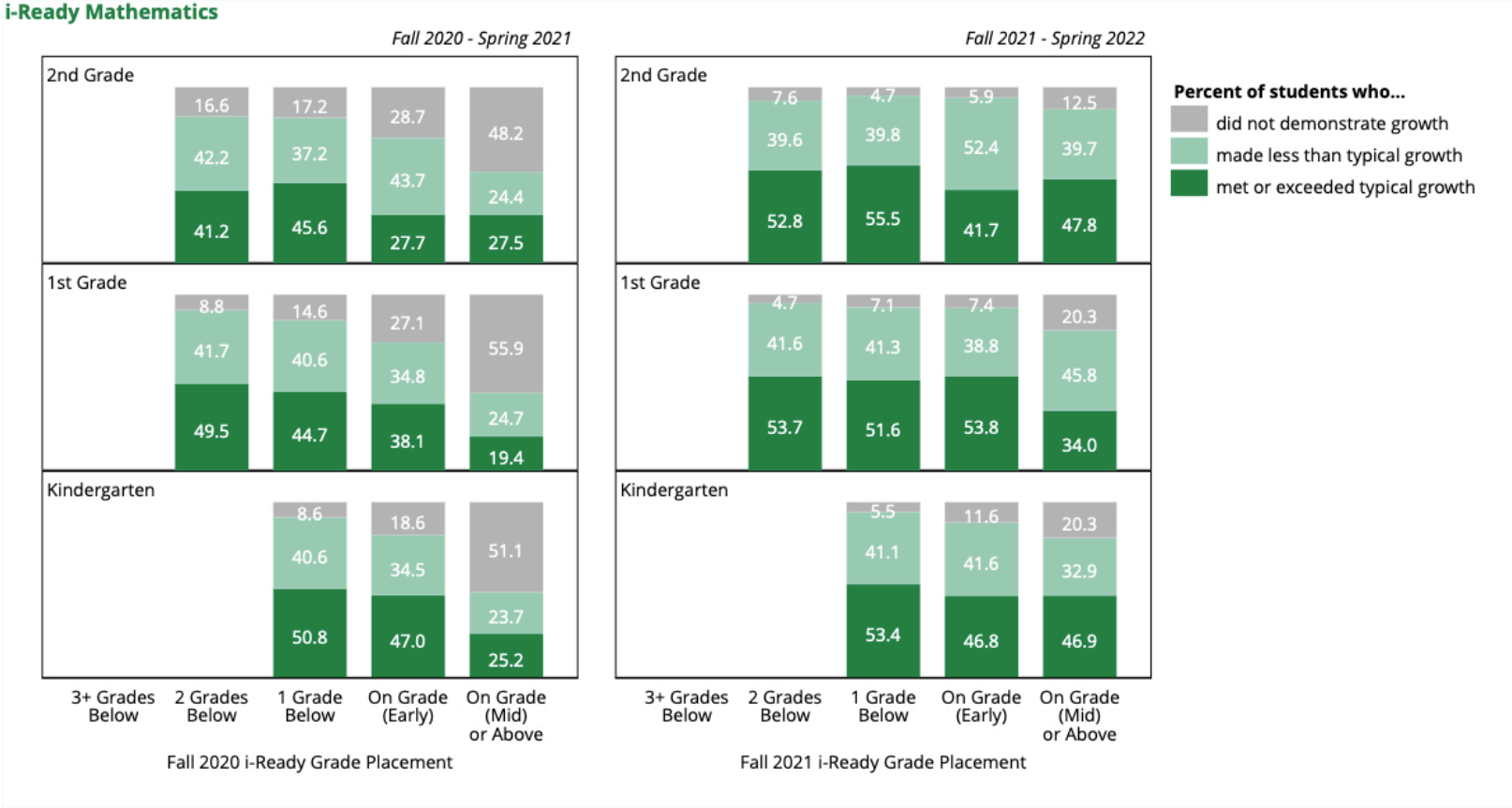
Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. We determine growth expectations using the 50th percentile of NWEA's fall-to-spring conditional math growth distribution for each grade level and students' initial achievement level (Thum & Kuhfeld, 2020).

Figure 3.3.11. Yearly Growth in 2020-21 and 2021-22 by Initial Performance Level, K-2 MAP Growth Reading



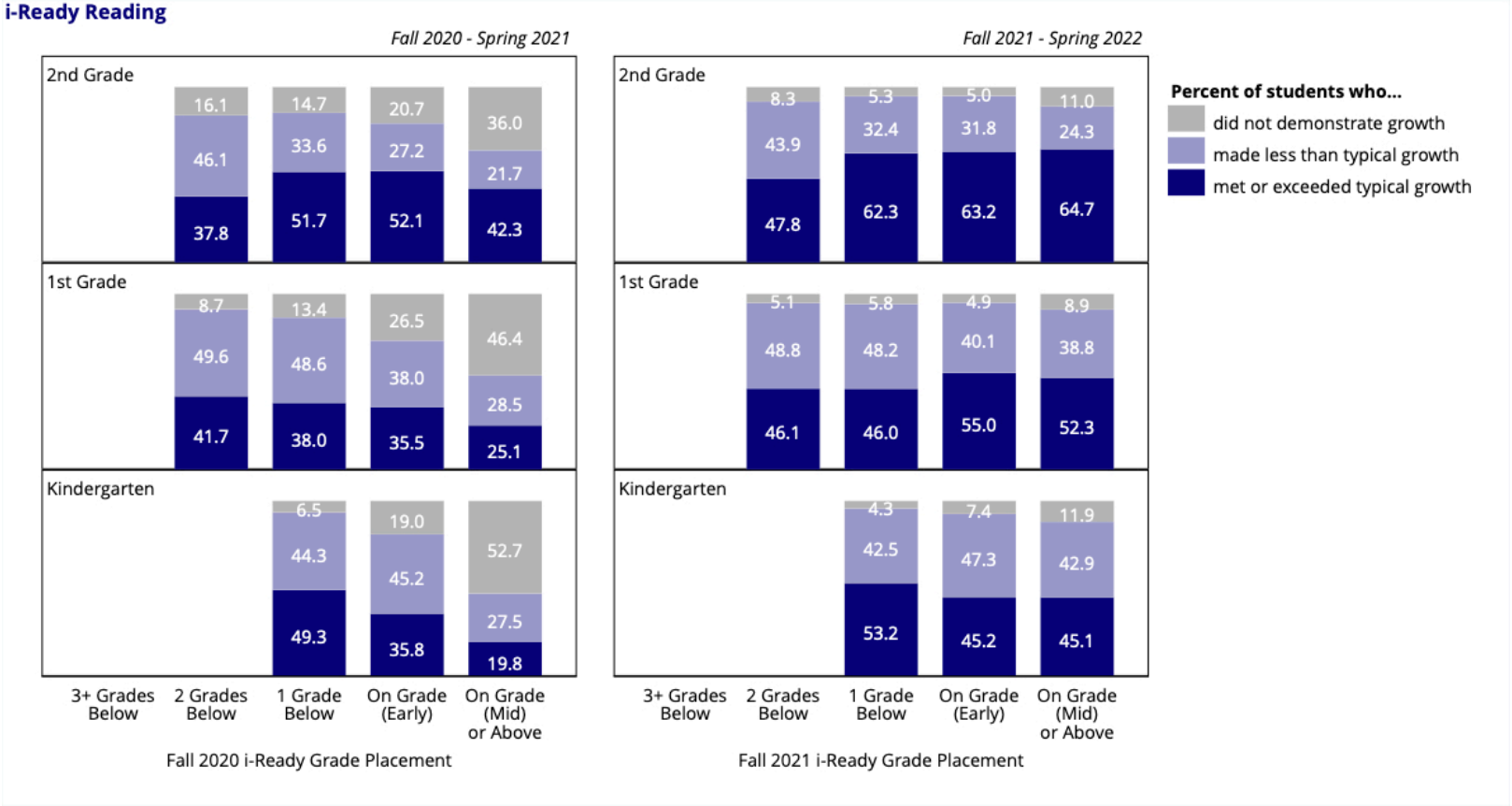
Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. We determine growth expectations using the 50th percentile of NWEA's fall-to-spring conditional reading growth distribution for each grade level and students' initial achievement level (Thum & Kuhfeld, 2020).

Figure 3.3.12. Yearly Growth in 2020-21 and 2021-22 by Initial Performance Level, K-2 i-Ready Math



Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. Typical growth targets come from the i-Ready assessment growth models which indicate median math growth for students at each grade level and initial achievement level (Curriculum Associates, 2020b).

Figure 3.3.13. Yearly Growth in 2020-21 and 2021-22 by Initial Performance Level, K-2 i-Ready Reading



Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. Typical growth targets come from the i-Ready assessment growth models which indicate median reading growth for students at each grade level and initial achievement level (Curriculum Associates, 2020b).

Yearly Growth by Demographic Characteristics

Figures 3.3.14 through 3.3.17 show the percentages of students in MAP Growth and i-Ready districts who achieved typical growth in 2020-21 and 2021-22 across different student demographic characteristics. The first two figures summarize differences by gender, economically disadvantaged status, and special education status, while the second set of figures document disparities by race/ethnicity.

Across all assessment providers, subject areas, and demographic subgroups, the percentages of students who met or exceeded typical growth all increased and the percentages of students who did not demonstrate any growth at all decreased between the 2020-21 and 2021-22 school years.

While Economically Disadvantaged and Special Education Students Remained Less Likely Than Their Peers to Meet or Exceed Typical Growth in 2021-22, Disparities Decreased Over Time

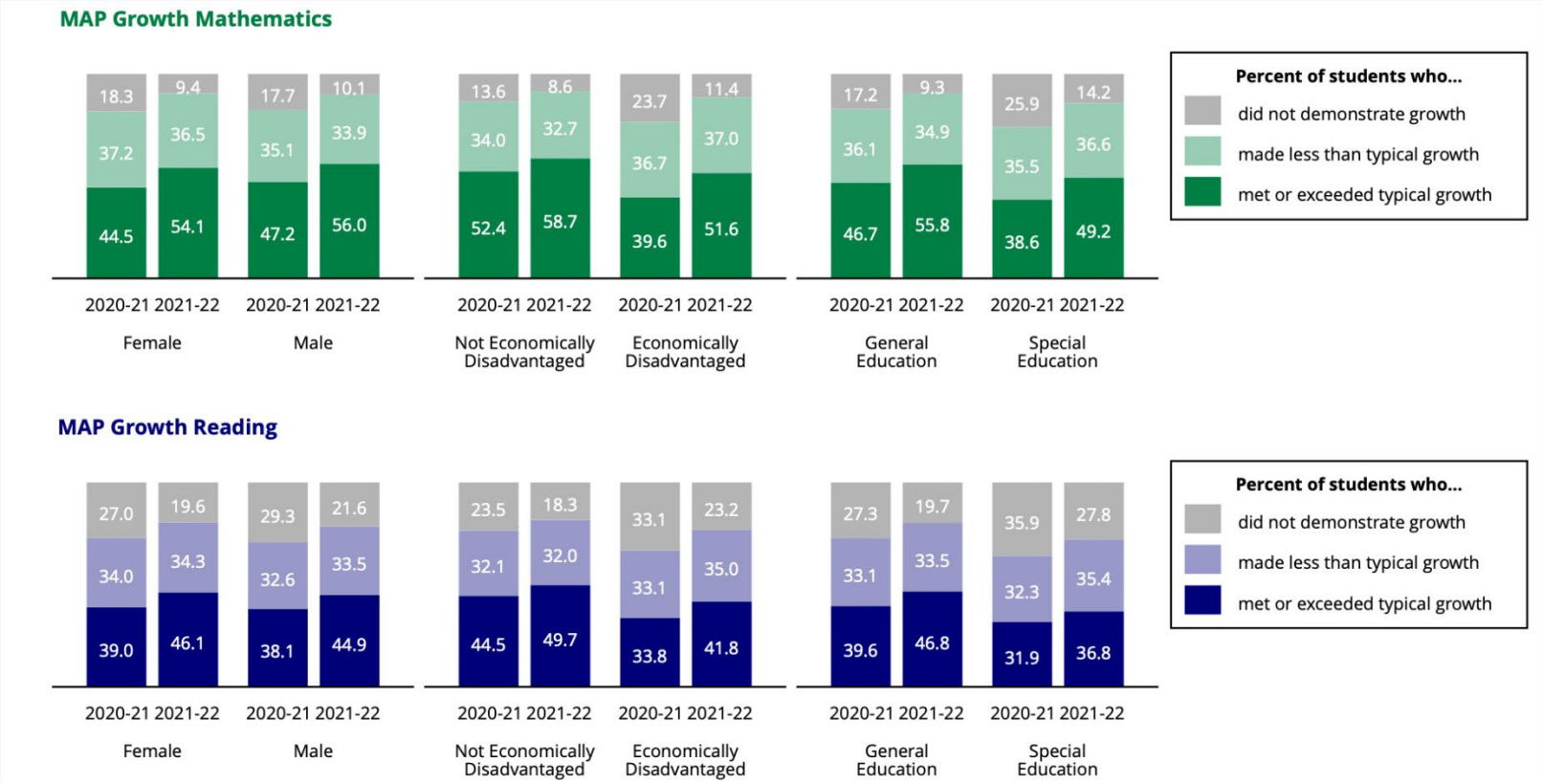
Figure 3.3.14 and Figure 3.3.15 show that, compared to their peers, economically disadvantaged students and students with disabilities were less likely to meet or exceed typical growth in either subject in both 2020-21 and 2021-22. Between these two school years, however, math outcomes for both economically disadvantaged students and students with disabilities improved at a faster rate compared to their counterparts, decreasing the disparities in their growth outcomes.

In both 2020-21 and 2021-22, male students were slightly more likely to meet or exceed typical growth in math compared to female students, while the opposite was true for reading. These gender disparities remained relatively consistent across the two school years.

While Black and Latino Students Remained Less Likely Than Their Peers to Meet or Exceed Typical Growth in 2021-22, Disparities Decreased Over Time

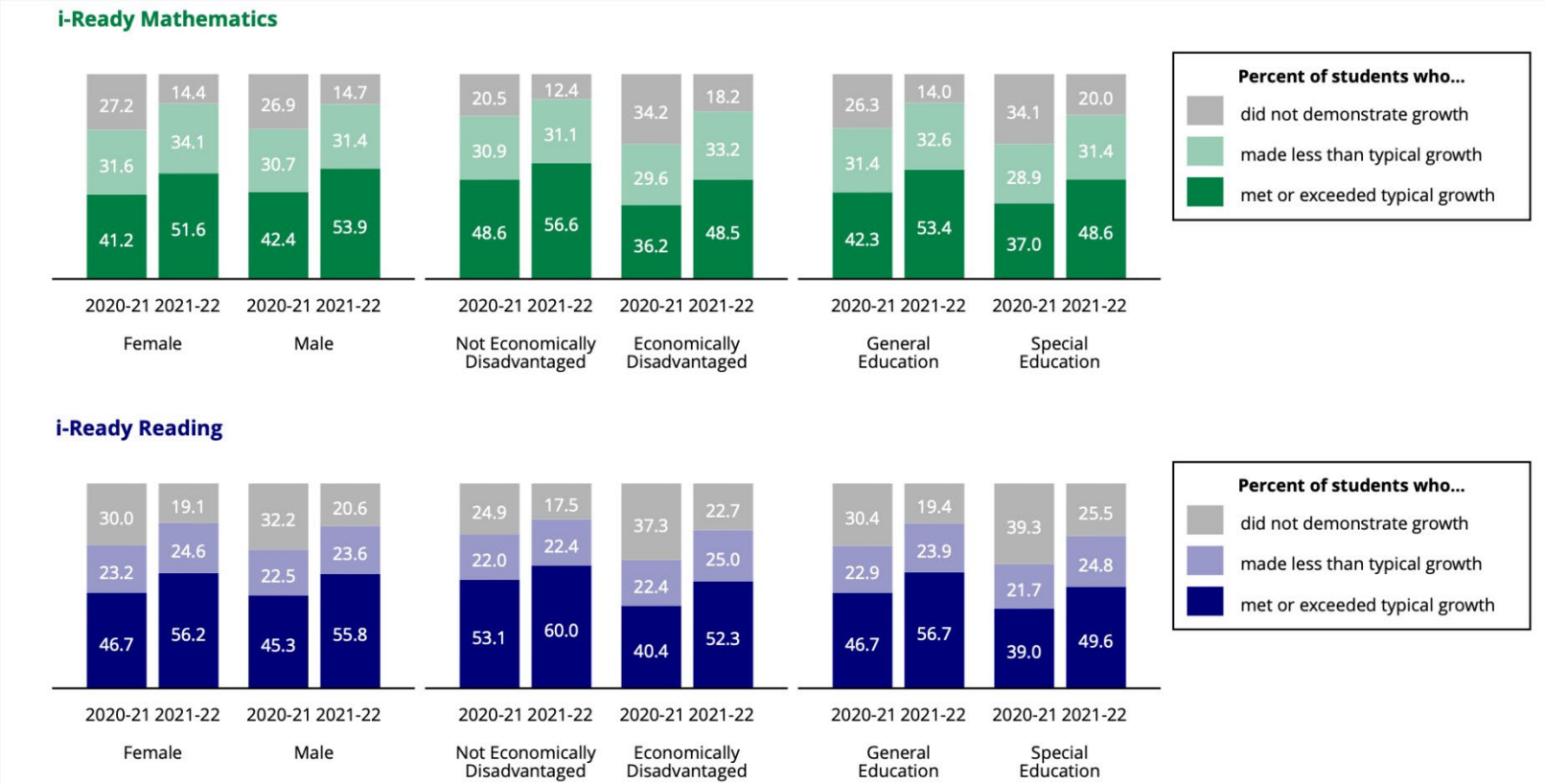
Figure 3.3.16 and Figure 3.3.17 show that Black and Latino students were the least likely of any racial or ethnic group to meet or exceed typical growth in either subject or school year, while White and Asian students were the most likely. However, Black and Latino students experienced the largest increases in the percentages of students who met or exceeded typical growth in both subjects between 2020-21 and 2021-22. As a result, the disparities in growth outcomes across race/ethnicity subgroups shrunk considerably by spring 2022.

Figure 3.3.14. Yearly Growth in 2020-21 and 2021-22 by Gender, Economically Disadvantaged Status, and Special Education Status, NWEA MAP Growth



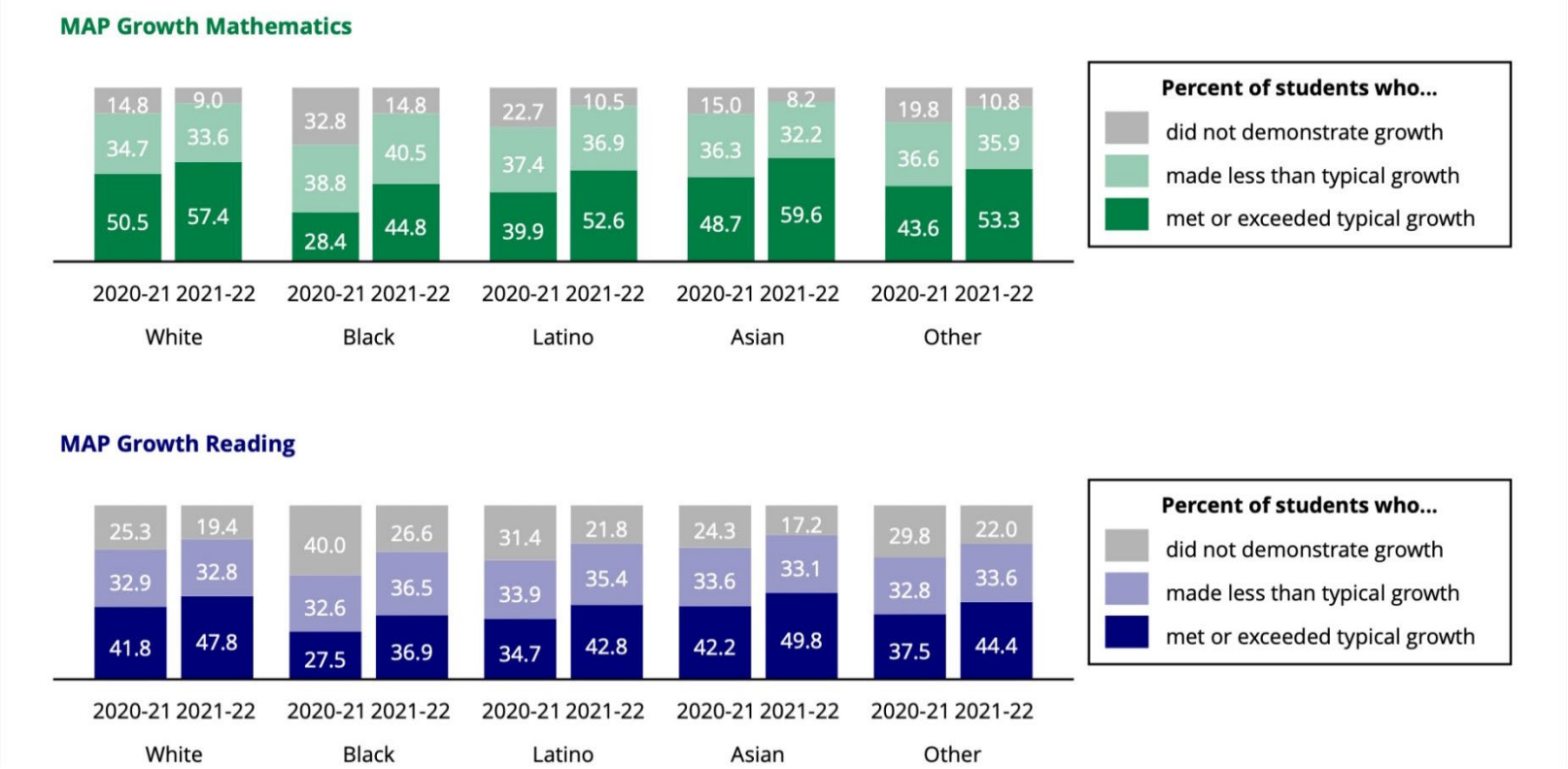
Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. We determine growth expectations using the 50th percentile of NWEA's fall-to-spring conditional growth distribution for each grade level, subject, and students' initial achievement level (Thum & Kuhfeld, 2020).

Figure 3.3.15. Yearly Growth in 2020-21 and 2021-22 by Gender, Economically Disadvantaged Status, and Special Education Status, Curriculum Associates' i-Ready



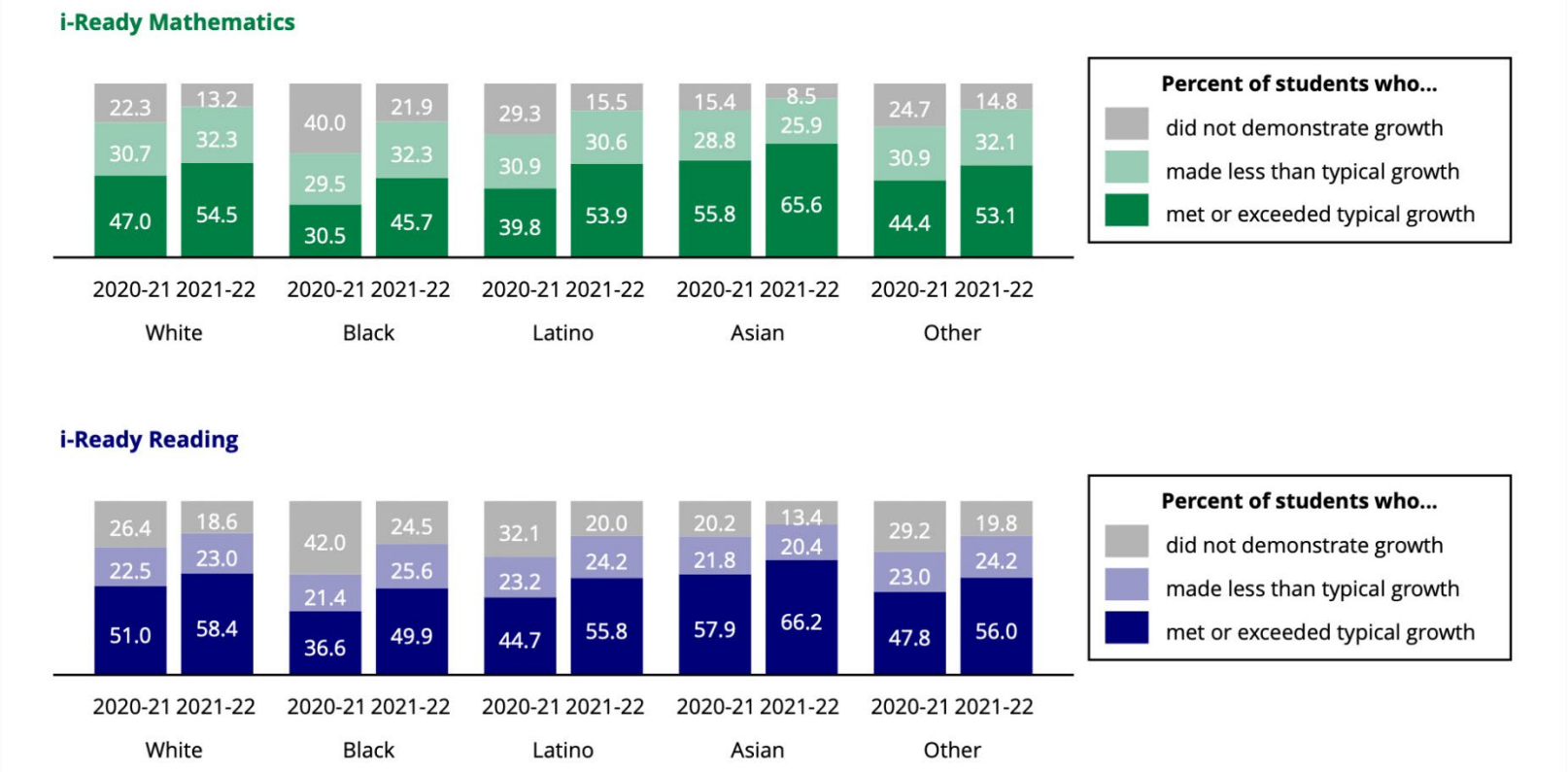
Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. Typical growth targets come from the i-Ready assessment growth models which indicate median math and reading growth for students at each grade level and initial achievement level (Curriculum Associates, 2020b).

Figure 3.3.16. Yearly Growth in 2020-21 and 2021-22 by Race/Ethnicity, NWEA MAP Growth



Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. We determine growth expectations using the 50th percentile of NWEA's fall-to-spring conditional growth distribution for each grade level, subject, and students' initial achievement level (Thum & Kuhfeld, 2020).

Figure 3.3.17. Yearly Growth in 2020-21 and 2021-22 by Race/Ethnicity, Curriculum Associates' i-Ready



Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. Typical growth targets come from the i-Ready assessment growth models which indicate median math and reading growth for students at each grade level and initial achievement level (Curriculum Associates, 2020b).

2020-21 and 2021-22 Instructional Modality

Figures 3.3.18 through 3.3.21 show similar breakdowns of growth outcomes by the instructional modalities *available to students* during the 2020-21 school year and by the instructional modalities *students received* during the 2021-22 school year.

Although students in districts that did not offer in-person instruction at all in 2020-21 were the least likely to meet or exceed typical growth in both years, performance gaps shrunk substantially once all districts returned to in-person learning in 2021-22

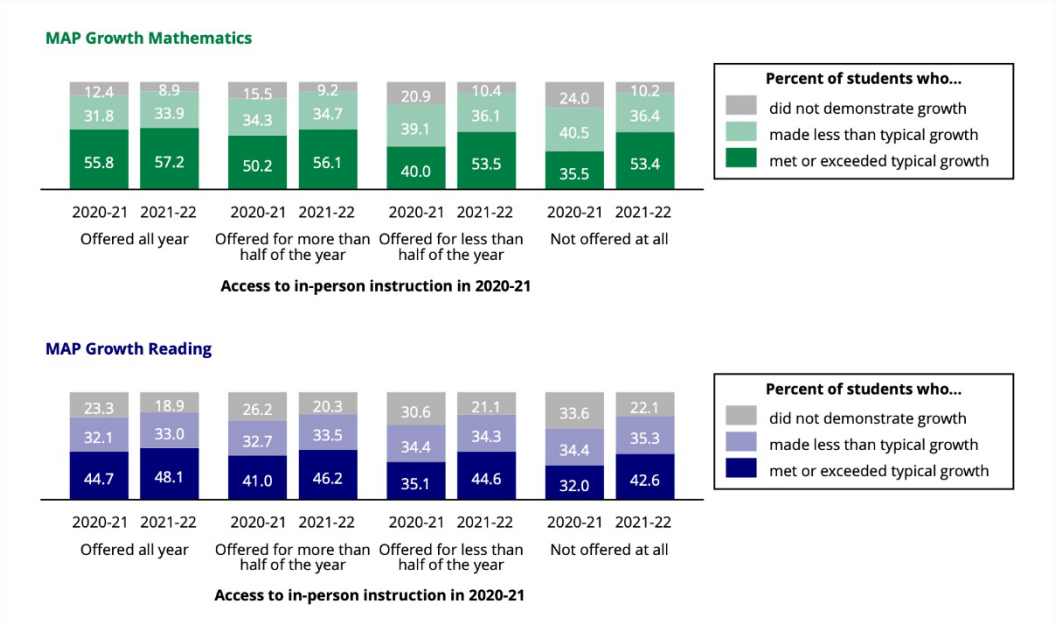
Figure 3.3.18 and Figure 3.3.19 show growth outcomes for students across subgroups based on their access to in-person instruction during the 2020-21 school year.[5] For students whose districts offered in-person instruction throughout the entire 2020-21 school year, growth outcomes changed very little from 2020-21 to 2021-22. It is perhaps unsurprising that students in these districts performed similarly across both full COVID-19 pandemic school years (2020-21 and 2021-22), as their learning environment—at least as evidenced by modality—changed the least. Students in these districts were more likely than those with less access to in-person instruction to meet or exceed typical growth in either subject, with approximately one-half to two-thirds achieving typical growth each year.

Students who had access to in-person instruction for only part of the 2020-21 school year made clear improvements in both math and reading once they had greater access to in-person instruction in 2021-22. The percentages of students who met or exceeded typical math and reading growth increased between 2020-21 and 2021-22, while the percentages of students who did not demonstrate growth decreased. Notably, students in NWEA MAP Growth districts that offered in-person instruction for more than half of the year were more likely to meet or exceed typical growth in both years, but students in districts that offered in-person instruction for less than half the year experienced greater improvements between years. As a result, gaps decreased between these two groups of districts over the course of the COVID-19 pandemic. This pattern was reversed in i-Ready districts, but gaps nonetheless diminished in 2021-22.

Students in districts that did not offer in-person instruction at any time during the 2020-21 school year were the least likely to meet or exceed typical growth and the most likely to not demonstrate any growth in both years—especially in the 2020-21 school year. These districts, on average, were in areas of the state that were hit hardest by the COVID-19 pandemic and were those with lowest pre-pandemic achievement levels (Hopkins et al., 2021). However, this group of districts generally saw the largest improvements in growth outcomes between the two years. Therefore, gaps between students in districts that never offered in-person instruction and those in districts that

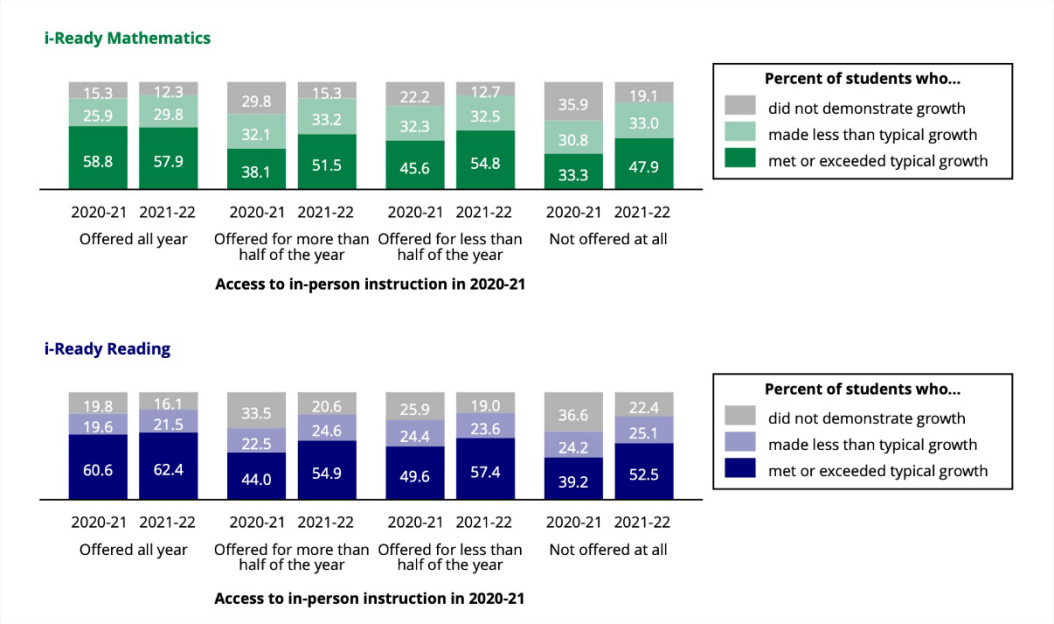
offered in-person instruction at least some of the year closed substantially in 2021-22, when all districts resumed in-person instruction.

Figure 3.3.18. Yearly Growth by Access to In-Person Instruction in 2020-21, NWEA MAP Growth



Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. We determine growth expectations using the 50th percentile of NWEA's fall-to-spring conditional growth distribution for each grade level, subject, and students' initial achievement level (Thum & Kuhfeld, 2020). The modality data from 2020-21 summarize what modes of instruction districts offered during each month of the school year.

Figure 3.3.19. Yearly Growth by Access to In-Person Instruction in 2020-21, Curriculum Associates' i-Ready

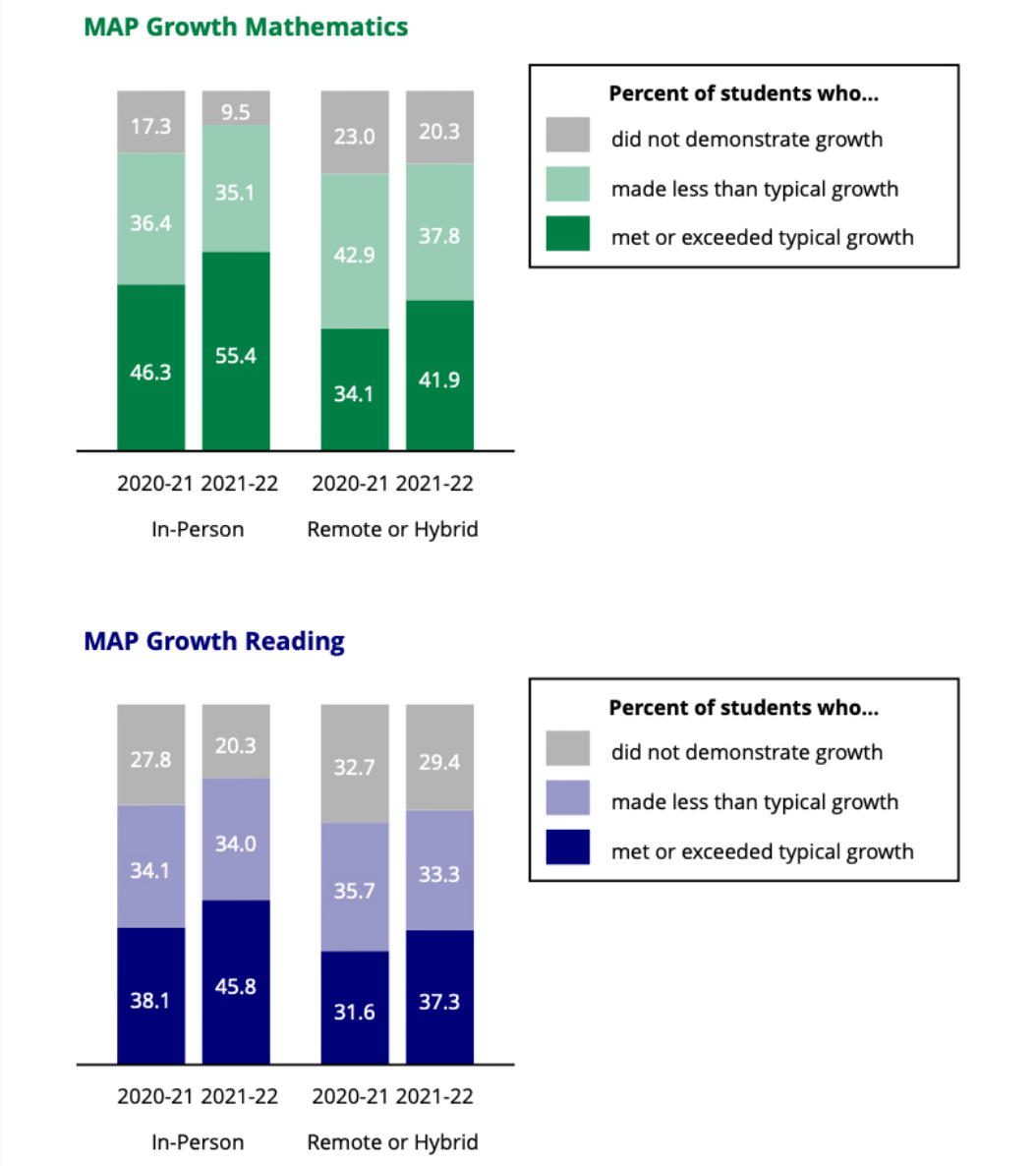


Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. Typical growth targets come from the i-Ready assessment growth models which indicate median math and reading growth for students at each grade level and initial achievement level (Curriculum Associates, 2020b). The modality data from 2020-21 summarize what modes of instruction districts offered during each month of the school year.

Students who learned remotely or in a hybrid context in 2021-22 were unlikely to have "caught up" to the same extent as peers learning in-person in 2021-22

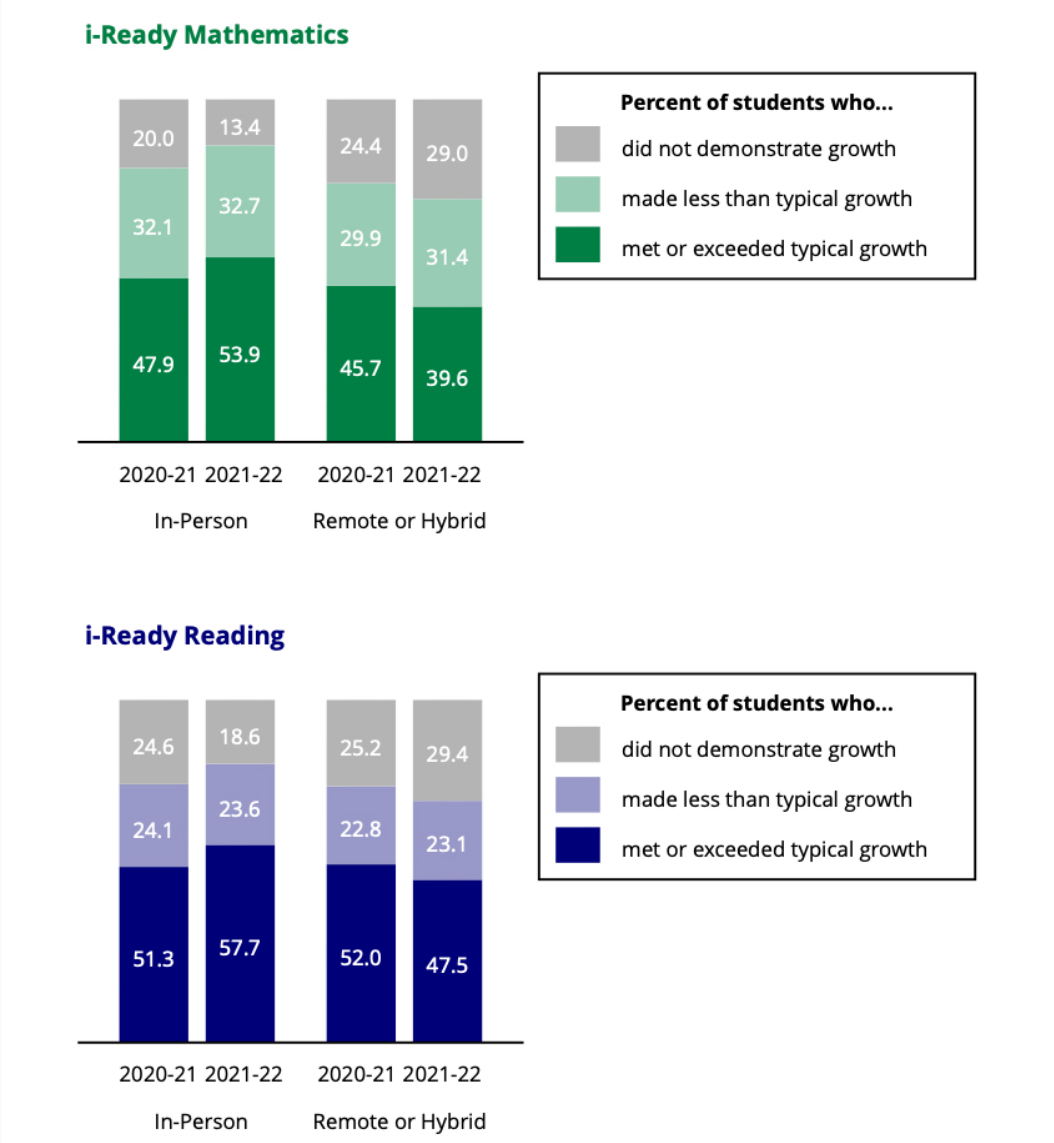
Figure 3.3.20 and Figure 3.3.21 show growth outcomes separately for students who primarily received in-person instruction in 2021-22 and those who received remote or hybrid instruction for a substantial portion of the 2021-22 school year. Notably, students who received hybrid or remote instruction in 2021-22 were already behind their peers at the beginning of the year, as these students were less likely to meet or exceed typical growth and more likely to not demonstrate any growth in the prior year, compared to those who were in-person in 2021-22.

Figure 3.3.20. Yearly Growth by 2021-22 Mode of Instruction, NWEA MAP Growth



Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. We determine growth expectations using the 50th percentile of NWEA's fall-to-spring conditional growth distribution for each grade level, subject, and students' initial achievement level (Thum & Kuhfeld, 2020). The modality data from 2021-22 summarize what mode of instruction a student received during school year.

Figure 3.3.21. Yearly Growth by 2021-22 Mode of Instruction, Curriculum Associates' i-Ready



Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. Typical growth targets come from the i-Ready assessment growth models which indicate median math and reading growth for students at each grade level and initial achievement level (Curriculum Associates, 2020b). The modality data from 2021-22 summarize what mode of instruction a student received during school year.

Improvements in growth outcomes on the MAP Growth assessments from 2020-21 to 2021-22 were consistently larger for students who received in-person instruction, though students who were remote or hybrid still experienced improvements, just to a

lesser extent. While we see similar changes from 2020-21 to 2021-22 for students in i-Ready districts who received in-person instruction in 2021-22, students in these districts who received remote or hybrid instruction were even *less likely* to meet or exceed typical growth and even *more likely* to demonstrate no growth in 2021-22 than they were in 2020-21. Overall, students who learned remotely or in a hybrid context in 2021-22 were substantially less likely to meet or exceed typical growth and more likely to demonstrate no growth. This suggests that these students, on average, were unlikely to have "caught up" to the same extent as peers learning in-person in 2021-22.

Each Additional Month That Students Had Access to In-Person Instruction in 2020-21 Increased Average Achievement by Up to 6% of a Typical Year's Growth by Spring of 2021

Based on the regressions described in Section Two that examine the relationship between months districts offered a given instructional modality and student performance on benchmark assessments, we can also provide an approximation of the change in average standardized achievement scores associated with each additional month that students had access to in-person instruction in 2020-21. The left panel of Table 3.3.1 shows results from this analysis. These numbers are scaled such that a value of 1 represents the standard deviation of scale scores for students from a nationally representative sample who took the same assessment in the same grade level and subject before the COVID-19 pandemic. Numbers between 0 and 1 can be interpreted as a percentage of this standard deviation. For example, an effect of 0.1 is equivalent to 10% of the size of the standard deviation for the pre-pandemic national norming sample.

To help convey the size of these effects in terms that are more meaningful to educators and stakeholders, we also convert them into percentages of a typical year's standardized achievement growth for students in each grade and subject attributable to an additional month of access to in-person learning in the 2020-21 school year.[6] We show the average across all grade levels. Results for each grade level individually are available in Appendix Table A4.

On the spring 2021 MAP Growth assessments, average scores increased by about 0.01 standard deviations in both math and reading for each additional month that a district offered in-person instruction. These effects were even larger for the i-Ready assessment; average scores increased by 0.03 and 0.02 standard deviations in math and reading, respectively, for each additional month that districts offered in-person instruction. This is likely because students in i-Ready districts, on average, had lower achievement scores before the COVID-19 pandemic, and students with lower baseline scores tend to make more growth on standardized assessments. When we translate these seemingly small changes into a proportion of a year's typical achievement gains, we see that these are indeed sizable effects. Each month of in-person instruction was

associated with an increase in spring 2021 student achievement equivalent to 3% to 4% of a year's learning on the MAP Growth assessments and 5% to 6% of a year's learning on the i-Ready assessments.

Table 3.3.1. Achievement Gains for Each Month a District Offered In-Person Instruction in 2020-21

Assessment	Increase in Average Standardized Scores		Increase in Percentage of a Typical Year's Growth	
	Math	Reading	Math	Reading
MAP Growth				
Spring 2021	0.011	0.012	2.58%	3.50%
Spring 2022	0.014	0.011	2.29%	1.89%
i-Ready				
Spring 2021	0.026	0.022	4.97%	5.59%
Spring 2022	0.034	0.034	5.34%	6.77%

Note: These regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to NWEA's and Curriculum Associates' pre-pandemic national norms, and estimates are converted to a percentile rank.

Students Continued to Benefit in 2021-22 from the In-Person Instruction Their Districts Provided in 2020-21

On average, students scored about 0.01 standard deviations higher on their spring 2022 MAP Growth math and reading assessments for each additional month of in-person instruction their districts offered the year before. After considering the differences in standardized achievement growth norms across grades and subjects, this is equivalent to about 2% of a year's learning. On the spring 2022 i-Ready assessments, students scored about 0.03 and 0.02 standard deviations higher in math and reading, respectively, for each month that their districts offered in-person instruction in 2020-21. This is about 5% of a typical year's math growth and 7% of a typical year's reading growth.

End-of-Year Achievement

Achieving less than typical growth, or no growth at all, has different implications for students depending on where they started. For example, students who began the year with top scores may still be close to or even above grade-level norms after a year of less than typical growth, while students who were already behind at the beginning of the year would be even further behind after a year of less than typical growth. To examine where Michigan students stood at the end of each year in terms of grade-

level performance standards, Figures 3.3.24 through 3.3.26 show end-of-year achievement and proficiency outcomes separately for students in each of the growth categories: Students who did not demonstrate growth; students who achieved less than typical growth; and students who met or exceeded typical growth.

HOW TO INTERPRET FIGURES 3.3.22 TO 3.3.26

Figure 3.3.22 shows the percentage of Michigan students in each growth category who scored above the pre-pandemic national average for their grade level on the spring 2021 and spring 2022 MAP Growth and i-Ready assessments.

- The **height of each bar** represents the percentage of students whose scores were **above** the pre-pandemic national average, across all grade levels for all students who had a particular growth outcome that year.
- The **light and dark shades of green** represent math results for the 2020-21 and 2021-22 school years, respectively. The **light and dark shades of blue** represent reading results for these same years.

Figures 3.3.23 through 3.3.26 show the distribution of estimated M-STEP proficiency levels across students in each growth category. As a comparison point, we also show the distribution of M-STEP proficiency levels for students in the same districts in 2018-19, the last year the test was administered before the COVID-19 pandemic.

- In these figures, each **shade of green** (for math) **or blue** (for reading) represents a different proficiency level. The **lightest shade** represents the lowest proficiency level (not proficient) and the **darkest shade** represents the highest proficiency level (advanced).
- For each school year depicted in the figure, there is a bar made up of four segments (one segment for each proficiency level), all of which add to 100%. The **height of each segment** represents the percentage of students who were classified into a particular proficiency category based on their M-STEP (2018-19) or benchmark assessment (2020-21 and 2021-22) scores.

Regardless of How Much Growth They Achieved, Students Were Less Likely to Score Above the National Average in 2021-22 Than in 2020-21

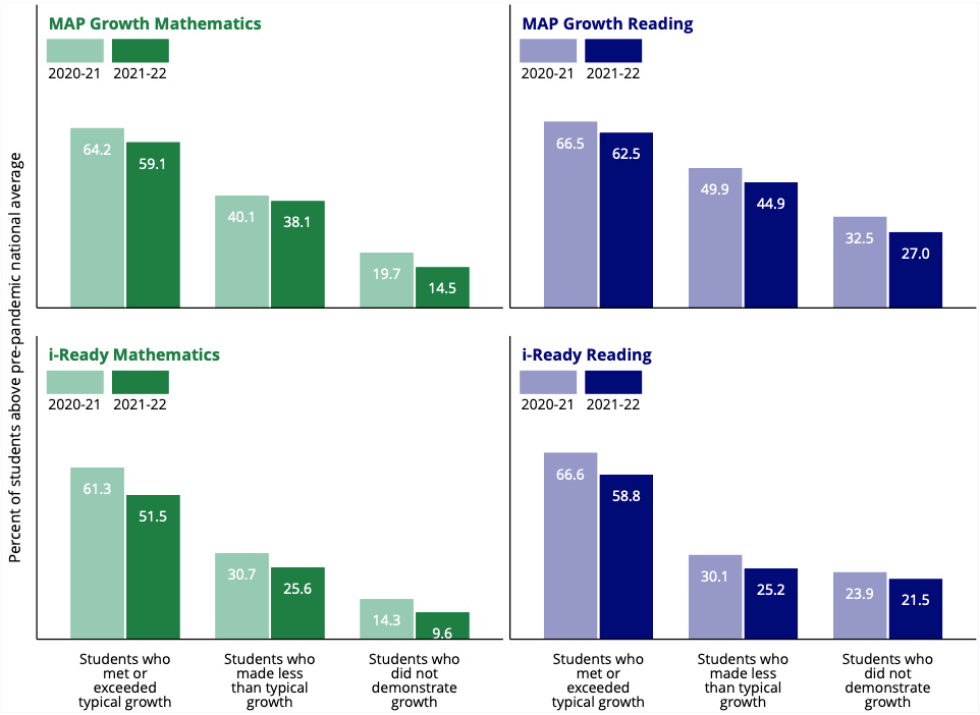
Across all growth categories, subjects, and assessment providers, the percentages of students scoring above pre-pandemic national averages decreased from 2020-21 to 2021-22. This is shown in Figure 3.3.22 and is likely because, on average, students started the 2021-22 school year behind where students in the same grade level were in 2020-21 (see for reference Figure 3.2.6 and Figure 3.2.7). This trend makes clear that the improvements in student achievement growth between 2020-21 and 2021-22

were not sufficient to bring enough students above the national average to reach the same end-of-year achievement level as the previous cohort.

High Achievers Who Were Already Ahead at the Beginning of the Year Account for Very Few of the Students Who Did Not Demonstrate Growth

Students with higher initial achievement typically demonstrate less growth over the course of a year than students with lower initial achievement; this pattern is evident in the pre-pandemic growth norms for the MAP Growth and i-Ready assessments as well as in our analysis of growth from fall 2020 to fall 2021 in our previous report.

Figure 3.3.22. Percent of Students Who Scored Above the Pre-Pandemic National Average for their Grade Level by Growth Category, NWEA MAP Growth and Curriculum Associates i-Ready



Note: These percentages include only students with benchmark assessment scores in both semesters of 2020-21 or 2021-22. Grade-specific national averages represent average achievement by students in each assessment provider's national norming sample.

However, Figure 3.3.22 shows that higher-achieving students, who theoretically may have less "room to grow," account for relatively few of the students who did not demonstrate growth in either the 2020-21 or 2021-22 school years. Depending on the year, subject, and assessment provider, between one-tenth and one-third of students who did not demonstrate growth scored above the pre-pandemic national average on

their spring benchmark assessments. Between one-quarter and one-half of the students who made less than typical growth ended the year above the pre-pandemic national average. In comparison, between 52% and 67% of students who met or exceeded typical growth ended the year above the pre-pandemic national average for their grade level.

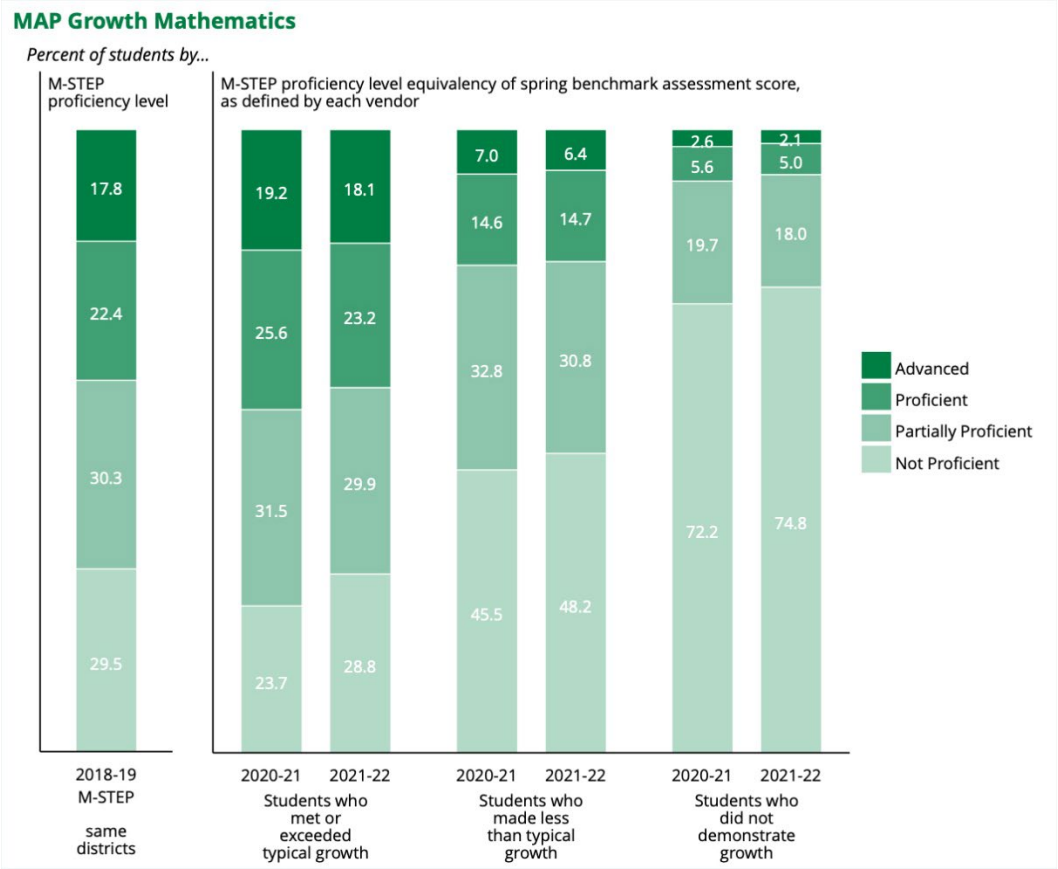
Regardless of Growth Category, Proficiency Rates Declined Between 2020-21 and 2021-22

Figures 3.3.23 through 3.3.26 make clear that estimated grade-level proficiency rates decreased slightly between 2020-21 and 2021-22. Importantly, this pattern is consistent in both MAP Growth and i-Ready districts and nearly all subject-growth category combinations. Estimated grade-level proficiency rates increased only for students who did not demonstrate any growth in math, and only very marginally by one-half and two percentage points in MAP Growth and i-Ready districts, respectively. However, this might be expected as this group of students already had very low proficiency rates and as a result their performance on the assessments could not substantially decrease between years.

Students Who Did Not Achieve a Typical Year's Growth Rarely Met State Standards for End-of-Year Grade-Level Proficiency

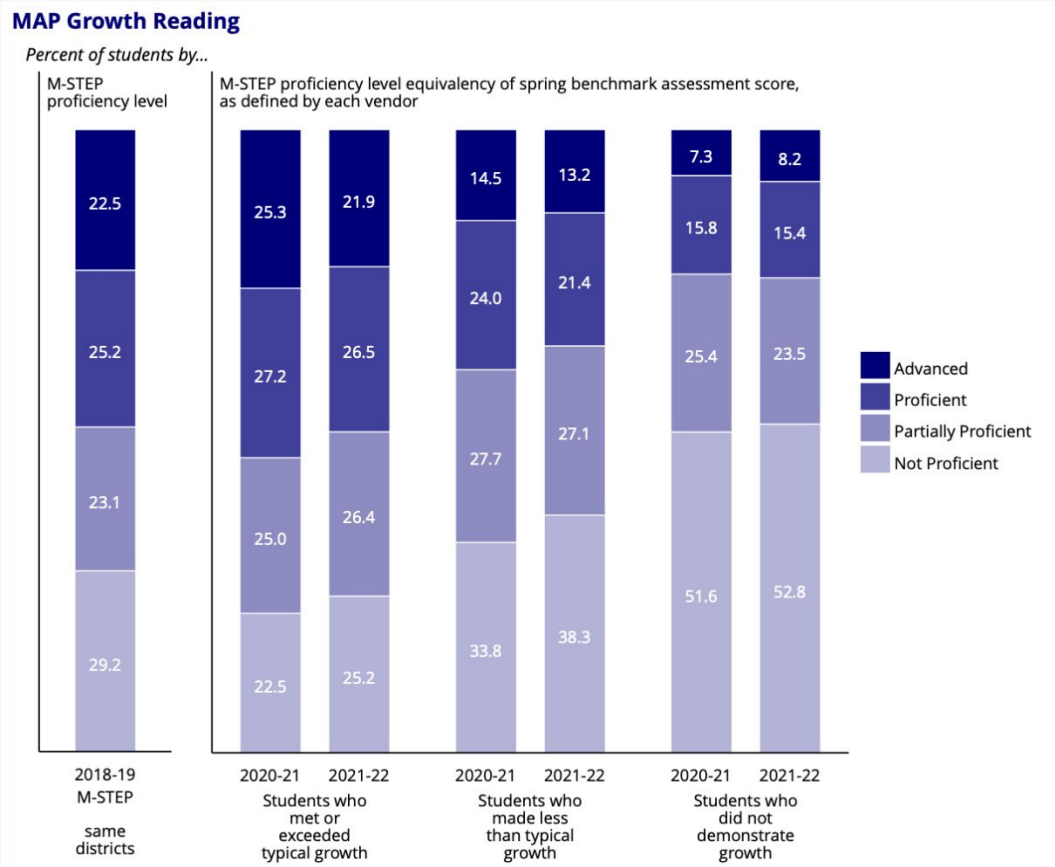
Proficiency levels for students who met or exceeded typical growth, estimated based on their spring MAP Growth or i-Ready benchmark assessment scores, were similar to the overall averages of all students in the same districts on the 2018-19 M-STEP. Students who made less than typical growth, on the other hand, were considerably more likely to score within the lowest proficiency level (not proficient) and less likely to score in the proficient or advanced categories. This is true to an even greater extent for students who did not demonstrate any growth at all, particularly on the math assessments. Notably, the proficiency rates for the 50% of students who achieved a typical year's growth in 2020-21 and 2021-22 mirrored the proficiency rates of all students in the same districts in 2018-19 while the other 50% performed substantially lower.

Figure 3.3.23. M-STEP Proficiency Levels and Vendor-Defined Equivalencies by Typical Growth Category, MAP Growth Math



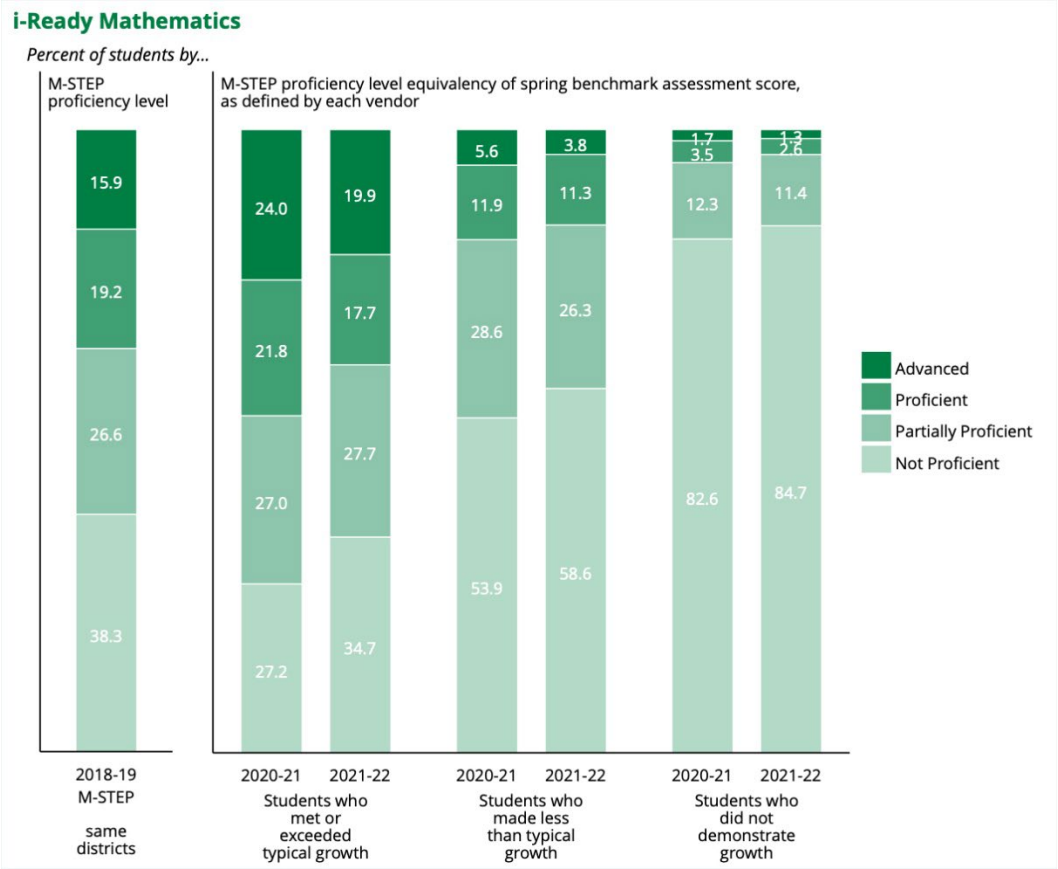
Note: These percentages include only 3rd- through 7th-grade students with benchmark assessment scores in spring 2020 and spring 2021. Benchmark assessment scores are converted to an equivalent M-STEP proficiency category using linking scores provided by each assessment provider. Proficiency rates from the 2018-19 M-STEP include all students in the same districts that use a particular benchmark assessment.

Figure 3.3.24. M-STEP Proficiency Levels and Vendor-Defined Equivalencies by Typical Growth Category, MAP Growth Reading



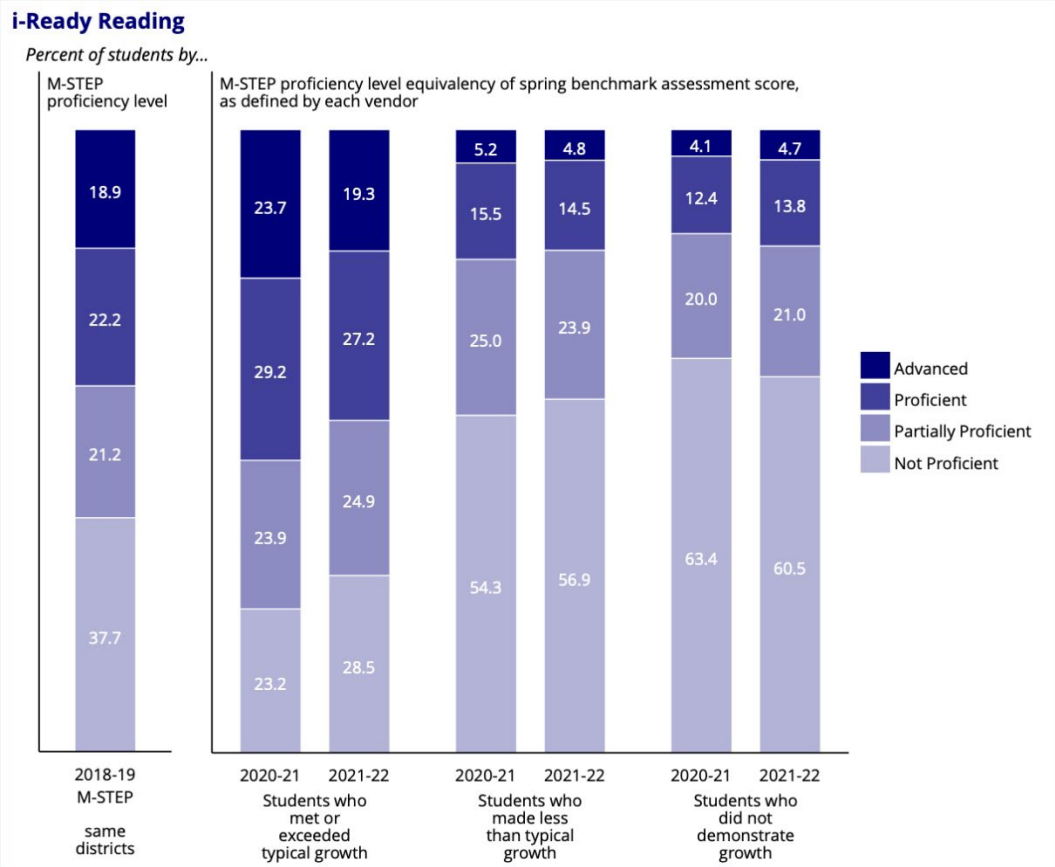
Note: These percentages include only 3rd- through 7th-grade students with benchmark assessment scores in spring 2020 and spring 2021. Benchmark assessment scores are converted to an equivalent M-STEP proficiency category using linking scores provided by each assessment provider. Proficiency rates from the 2018-19 M-STEP include all students in the same districts that use a particular benchmark assessment.

Figure 3.3.25. M-STEP Proficiency Levels and Vendor-Defined Equivalencies by Typical Growth Category, i-Ready Math



Note: These percentages include only 3rd- through 7th-grade students with benchmark assessment scores in spring 2020 and spring 2021. Benchmark assessment scores are converted to an equivalent M-STEP proficiency category using linking scores provided by each assessment provider. Proficiency rates from the 2018-19 M-STEP include all students in the same districts that use a particular benchmark assessment.

Figure 3.3.26. M-STEP Proficiency Levels and Vendor-Defined Equivalencies by Typical Growth Category, i-Ready Reading



Note: These percentages include only 3rd- through 7th-grade students with benchmark assessment scores in spring 2020 and spring 2021. Benchmark assessment scores are converted to an equivalent M-STEP proficiency category using linking scores provided by each assessment provider. Proficiency rates from the 2018-19 M-STEP include all students in the same districts that use a particular benchmark assessment.

Section Four: Takeaways and Implications

This report furthers our analyses of Michigan student learning during the COVID-19 pandemic by examining math and reading benchmark outcomes throughout the past two school years, from fall 2020 to spring 2022. In particular, we explore overall student achievement and achievement trajectories compared to pre-pandemic norms, progress toward appropriate growth targets, and differences in performance across subgroups of students with varying characteristics.

While this report helps to deepen our understanding of how Michigan public school students progressed and learned during the 2020-21 and 2021-22 school years, there are several limitations of the data that must be considered when interpreting results. Most importantly, the analyses presented in this report are based on imperfect and incomplete data. The students included in our analyses represent only a subset of the K-8 population across the state, and prior research has shown that the COVID-19 pandemic has had a greater negative effect on achievement and achievement growth for the specific student populations who are underrepresented in these analyses. Additionally, given that many districts administered benchmark testing virtually in the fall of 2020, it is difficult to assess fall 2020 performance and growth measures that incorporate fall 2020 achievement as a baseline to contextualize student progress. Lastly, the data on instructional modalities offered by school districts in the 2020-21 school year reflect only what the districts themselves offered not how individual students learned. Previous work in Michigan and elsewhere has shown that districts retained local autonomy over the choice of modality offerings, and parents and families were almost always able to make their own choices to learn remotely if they chose to do so.

Many key findings described below show that students have started to recover from the missed learning opportunities the COVID-19 pandemic caused. Policymakers, educators, and stakeholders should use these data to guide local and state education agencies as they continue to work to address the challenges wrought by the COVID-19 pandemic.

KEY FINDINGS

- **On average, students in 2021-22 started the school year behind but ended the year ahead of students who were in the same grade levels in 2020-21.** In most grade levels, students in fall 2021 had lower average benchmark assessment scores than students who were in the same grade in fall 2020. However, the 2021-22 cohort of students experienced greater achievement growth over the course of the school year and ultimately surpassed spring 2021 achievement levels in nearly all grades by spring 2022.
- **Students were more likely to demonstrate growth, more likely to make substantial progress toward yearly growth targets, and more likely to achieve a full year's growth in 2021-22 than in 2020-21.** Across subjects, grade levels, and assessment providers, the percentages of students who met or exceeded a typical year's growth increased, the percentages who did not demonstrate any growth at all decreased, and the students who made less than a typical year's growth made more progress, on average, toward their growth targets.
- **Improvements in achievement growth in 2021-22 were often not enough to counteract the effects of unfinished learning in 2020-21.** In 2020-21, average benchmark assessment scores across nearly all grade levels, subjects, and assessment providers increased at a slower rate than would have been expected before the COVID-19 pandemic, and students who started the year below pre-pandemic norms fell further behind. In 2021-22, math and reading scores increased at faster rates than the year before, but often still lagged behind pre-pandemic growth rates. While in some cases accelerated growth in 2021-22 resulted in average scores nearing—or sometimes even exceeding—pre-pandemic norms by spring 2022, in other cases, average scores fell further below pre-pandemic norms. Overall, fewer Michigan students scored above pre-pandemic national averages in 2021-22 than in 2020-21.
- **Fifteen percent of students did not demonstrate growth on their benchmark assessments in 2021-22. Most of these students scored far below grade-level proficiency standards at the end of the year.** While student growth, on average, improved relative to 2020-21, there were still far more students who did not demonstrate any growth at all in 2021-22 than in the pre-pandemic national norming samples. On average across grade levels, subjects, and assessment providers, about 15% of students did not demonstrate growth on their 2021-22 benchmark assessments. This is far less than the 24% of students who did not demonstrate growth in 2020-21 but still

more than the 11% of students who did not demonstrate growth in a typical pre-pandemic year. This is particularly concerning because the far majority of students who made no progress on their benchmark assessments were amongst the lowest performing in the state; while these students occasionally scored within a range that the assessment providers consider equivalent to the partially proficient level on the M-STEP, the vast majority fell within the lowest possible M-STEP proficiency level (not proficient), especially in math.

- **Students with less access to in-person instruction in 2020-21 experienced less achievement growth than those with more access. These gaps began to diminish for students who returned to in-person learning in 2021-22.** Students whose districts offered in-person instruction for only part of the 2020-21 school year made clear improvements in both math and reading once in-person instruction was more readily available in 2021-22. Students in districts that did not offer in-person instruction at any time during the 2020-21 school year were the least likely to achieve a typical year's growth and the most likely to demonstrate no growth in either year, though performance gaps shrunk substantially once most districts returned to in-person learning in 2021-22. Improvements in growth outcomes between 2020-21 to 2021-22 were consistently larger for students who received in-person instruction in 2021-22. On the MAP Growth assessments, students who continued to learn in a hybrid or remote format in 2021-22 experienced slight improvements, but to a lesser extent than those who returned to in-person instruction. However, on the i-Ready assessments, students who received remote or hybrid instruction in 2021-22 fared even worse than they did in 2020-21.
- **Students whose districts offered in-person instruction in 2020-21 continued to benefit academically, even a year later.** Each additional month that a district offered in-person instruction in 2020-21 was associated with an increase in average spring 2021 achievement of about 3 to 6% of a typical year's learning. In spring 2022, average achievement still increased by between 2% and 7% of a year's learning for every month a district offered in-person instruction the year before.
- **After most students returned to in-person learning in 2021-22, demographic gaps in achievement growth shrunk substantially.** Economically disadvantaged students, students with disabilities, Black students, and Latino students were less likely than their peers to achieve a typical year's math or reading achievement growth in both 2020-21 and 2021-22. Between these two school years, however, growth outcomes for each of these student subgroups improved at faster rates compared to their counterparts, decreasing the disparities.

IMPLICATIONS

Our findings make clear that interruptions to student learning and instruction during the COVID-19 pandemic negatively affected average student achievement and growth outcomes for Michigan students, and disruptions to learning were not spread out equitably amongst the student population. However, our results also show that many students improved academically over the past school year. This research highlights a few important considerations for policymakers moving forward:

- **The tremendous effect that the COVID-19 pandemic has had, and continues to have, on student learning will not be addressed quickly or without a substantial and sustained influx of resources to support education in Michigan.** As noted earlier, average achievement and achievement growth for many students was lower during the COVID-19 pandemic than would have been the case in a typical school year, and substantial proportions of students demonstrated no achievement growth during the 2020-21 *and* 2021-22 school years. This is not unique to Michigan, but rather mirrors recent findings from across the U.S. (e.g., Curriculum Associates, 2022; Goldhaber et al., 2022; Kuhfeld & Lewis, 2022). Results from the 2021-22 school year make clear that the road to academic recovery will not be quick and a return to "business as normal" will be insufficient to improve student achievement to pre-pandemic levels. It will be critical for local, state, and the federal governments to prioritize both short- and longer-term investments into public education, in Michigan and elsewhere, as educators and students work to recover from the trauma of the COVID-19 pandemic.
- **Educators and policymakers must continue to monitor learning outcomes for all students, and especially for groups that were disproportionately affected by the COVID-19 pandemic.** The mandated use and reporting of benchmark assessments make it possible for state and local policymakers to understand where progress is (and is not) being made towards academic recovery. It will be critical to continue collecting data that allow policymakers, educators, and stakeholders to assess progress in the coming years. In particular, research exploring trends in academic achievement over the past two years makes clear that the COVID-19 pandemic has had a greater and more negative effect on economically disadvantaged, Black, and Latino students as well as students enrolled in special education. While we do find that outcomes for these students increased at a faster rate compared to their respective peers between 2020-21 and 2021-22, disparities between each group persist. Any decisions to reduce monitoring of student learning progress may exacerbate longstanding achievement gaps.

- **Stakeholders should continue to interpret and use fall 2020 benchmark assessment results with caution.** Without giving due consideration to the "at-home testing advantage" for early elementary students, stakeholders could come to misleading conclusions about student performance, growth, and achievement gaps based on the fall 2020 data. We find that most abnormal patterns in the benchmark assessment data are concentrated among kindergarten, 1st-, and 2nd-grade students with fall 2020 scores above the 80th percentile. Focusing on students with fall 2020 scores at or below the 80th percentile will provide a more reliable (albeit less representative) picture of student performance outcomes. This is true in our data and analyses, as well as in national reporting from assessment vendors (Huff, 2020; Kuhfeld et al., 2020; Renaissance Learning, 2021a).

References

Amplify Education. (2021). COVID-19 means more students not learning to read. Amplify Education Research Brief. https://readytogether.sde.ok.gov/sites/default/files/2022-04/Amplify-mCLASS_MOY-COVID-Learning-Loss-Research-Brief_022421.pdf

Appleton, A. (2022, July 13). *Modest gains for this year's Indiana test scores after pandemic decline.* Chalkbeat Indiana. https://in.chalkbeat.org/2022/7/13/23205866/ilearn-indiana-state-testing-scores-2022-pandemic-recovery

Baird, M. D., & Pane, J. F. (2019). Translating standardized effects of education programs into more interpretable metrics. Educational Researcher, 48(4), 217-228. https://journals.sagepub.com/doi/abs/10.3102/0013189X19848729

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness*, 1(4), 289–328. https://www.tandfonline.com/doi/abs/10.1080/19345740802400072

Curriculum Associates (2020a, May). Scores on i-Ready Diagnostic that are equivalent to performance levels on the Michigan Student Test of Educational Progress (M-STEP). Curriculum Associates Research Report No. RR 2020-29.

Curriculum Associates. (2020b, August). i-Ready Diagnostic National Norms Tables for Reading and Mathematics for Grades K-8 2020-21.

Curriculum Associates. (2022, September). The state of student learning in 2022. *Curriculum Associates Annual Report.* https://www.curriculumassociates.com/-/media/mainsite/files/corporate/state-of-student-learning-2022.pdf?_ga=2.249214663.1791340512.1663161051-741382319.1663161051.

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020, December 8). COVID-19 and learning loss—disparities grow and students need help. *McKinsey & Company*. https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-learning-loss-disparities-grow-and-students-need-help.

Greater Fort Lauderdale Alliance. (2022, June 29). State releases 2022 Florida standards assessment and end of course exam results. *The Alliance*. https://www.gflalliance.org/news/2022/06/29/education-news/state-releases-2022-florida-standards-assessment-and-end-of-course-exam-results/#:%7E:text=State%20Releases%202022%20Florida%20Standards%20Assessment%20and%20End%20of%20Course%20Exam%20Results,-Scores%20Show%20BCPS&text=On%20Tuesday%2C%20June%2028%2C%202022,Mathematics%20grades%203%20through%208.

Goldhaber, D., Kane, T. J., McEachin, A., & Morton, E. (2022). A comprehensive picture of achievement across the COVID-19 pandemic years: Examining variation in test levels and growth across districts, schools, grades, and students. Working Paper No. 265-0522. *National Center for Analysis of Longitudinal Data in Education Research (CALDER)*. https://caldercenter.org/publications/comprehensive-picture-achievement-across-covid-19-pandemic-years-examining-variation

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

Idaho State Department of Education. (2022, July 6). *Early reading test shows statewide gains from previous spring and fall.* [Press release]. https://www.sde.idaho.gov/communications/files/news-releases/07-06-2022-Idaho-early-reading-test-shows-strong-statewide-gains-from-previous-spring-and-fall.pdf

Kilbride, T., Hopkins, B., & Strunk, K.O. (2021) Michigan's 2020-21 Benchmark Assessments. Education Policy Innovation Collaborative. https://epicedpolicy.org/wp-content/uploads/2022/02/Benchmark-Rptvl_Aug2021.pdf

Kilbride, T., Hopkins, B., Strunk, K.O., & Imberman, S. (2021) K-8 student achievement and achievement gaps on Michigan's 2020-21 benchmark and summative assessments. Education Policy Innovation Collaborative. https://epicedpolicy.org/k-8-student-achievement-and-achievement-gaps-on-michigans-2020-21-benchmark-and-summative-assessments/

Kilbride, T., Hopkins, B., Strunk, K.O., & Yu, D. (2022) Michigan's Fall 2021 Benchmark Assessments. Education Policy Innovation Collaborative. https://epicedpolicy.org/wp-content/uploads/2022/04/Benchmark_Report_April2022.pdf

Kogan, V. (2022, June). Academic achievement and pandemic recovery in Ohio: An update from fall third grade English language arts assessments. https://glenn.osu.edu/academic-achievement-pandemic-recovery.

Kogan, V., & Lavertu, S. (2021, January 27). *The COVID-19 pandemic and student achievement on Ohio's third-grade English language arts assessment*. John Glenn College of Public Affairs, Ohio State University. http://glenn.osu.edu/educational-governance/reports/reports-attributes/ODE_ThirdGradeELA_KL_1-27-2021.pdf.

Kuhfeld, M., & Lewis, K. (2022). Student achievement in 2021-22: Cause for hope and continued urgency. NWEA. https://www.nwea.org/research/publication/student-achievement-in-2021-22-cause-for-hope-and-continued-urgency

Kuhfeld, M., Lewis, K, Meyer, P., & Tarasawa, B. (2020, November). Comparability analysis of remote and in-person MAP Growth testing in fall 2020. NWEA. https://www.nwea.org/research/publication/comparability-analysis-ofremote-and-in-person-map-growth-testing-in-fall-2020/.

McEachin, A. & Atteberry, A. (2017). The impact of summer learning loss on measures of school performance. Education Finance and Policy, 12(4), 468-491. https://direct.mit.edu/edfp/article-abstract/12/4/468/10289/The-Impact-of-Summer-Learning-Loss-on-Measures-of

Michigan Department of Education (2019). *Technical Report, Spring 2019 Michigan Student Test of Educational Progress (M-STEP).* https://www.michigan.gov/mde/-/media/Project/Websites/mde/Year/2020/05/14/Spring_2019_M-STEP_Technical_Report_Main_report.pdf?rev=0c3bbf9629ea411395157233c92439e0&hash=7889BF8A50495DC27B1FD1AD62820C04

Michigan Public Act 147 of 2020, Mich., MCL § 388.1621f (2020) http://www.legislature.mi.gov/(S(rhqyboiv2ivksh4xj0wbiqpl))/mileg.aspx?page=GetObject&objectname=mcl-388-1621f.

Michigan Public Act 148 of 2020, Mich. MCL § 388.1701 (2020) http://www.legislature.mi.gov/(S(hotihihl5cde5dkkjzjvx4g3))/mileg.aspx?page=getObject&objectName=mcl-388-1701.

Michigan Public Act 149 of 2020, Mich., MCL § 388.1606 (2020) http://legislature.mi.gov/doc.aspx?mcl-388-1606.

Michigan Public Act 48 of 2021, Mich., MCL § 388.1704a (2021) http://legislature.mi.gov/doc.aspx?mcl-388-1704a.

National Center for Education Statistics (2022). *NAEP long-term trend assessment results: Reading and mathematics*. The Nation's Report Card. https://www.nationsreportcard.gov/highlights/ltt/2022/

NWEA (2020, December). *Linking study report: Predicting performance on the Michigan state assessment system in grades 3-8 ELA and mathematics based on NWEA MAP Growth scores.* NWEA Psychometric Solutions. https://www.nwea.org/content/uploads/2016/12/MI-MAP-Growth-Linking-Study-Report_NWEA_2020-12-22.pdf

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

Quinn, D. M., Cooc, N., McIntyre, J., & Gomez, C. J. (2016). Seasonal dynamics of academic achievement inequality by socioeconomic status and race/ethnicity: Updating and extending past research with new national data. Educational Researcher, 45(8), 443-453. https://scholar.harvard.edu/dmq/publications/seasonal-dynamics-academic-achievement-inequality-socioeconomic-status-and

Renaissance Learning (2019, July). Pathway to proficiency: Linking Star Reading and Star Math to the Michigan Student Test of Educational Progress (M-STEP). *Renaissance Technical Papers.* *https://doc.renlearn.com/KMNet/R0053649AF33BE0B.pdf*

Renaissance Learning. (2021a). How kids are performing: Tracking the school-year impact of COVID-19 on reading and mathematics achievement. *Special Report Series.* https://renaissance.widen.net/s/t8trbrt2tc/r63444

Renaissance Learning. (2021b). Renaissance: Star assessments for early literacy technical manual.

Renaissance Learning. (2021c). Renaissance: Star assessments for math technical manual.

Renaissance Learning. (2021d). Renaissance: Star assessments for reading technical manual.

Sass, T. & Goldring, T. (2021, May). Student achievement growth during the COVID-19 pandemic: Insights from Metro-Atlanta School Districts. Georgia State University. https://gpl.gsu.edu/download/student-achievement-growth-during-the-covid-19%1epandemic%1ereport/?ind=1620902643226&filename=Achievement%20Growth%20During%20COVID%1e19_Report_20210512_FINAL.pdf&wpdmdl=2105&refresh=609e93df82b691621005279.

Smarter Balanced Assessment Consortium (2020). *2018-19 summative technical report.* https://technicalreports.smarterbalanced.org/2018-19_summative-report/_book/

Smarter Balanced Assessment Consortium (2021). *2020-21 interim technical report.*

Smarter Balanced Validity Research. (2020, October 14). Percentiles. Retrieved from https://validity.smarterbalanced.org/reports-and-specifications/percentiles/

Texas Education Agency. (2022, July 1). *TEA Releases 2022 Grades 3–8 STAAR Results | Texas Education Agency*. https://tea.texas.gov/about-tea/news-and-multimedia/news-releases/news-2022/tea-releases-2022-grades-3-8-staar-results

Tennessee Department of Education. (2022, June 14). *Tennessee Releases 2021–22 TCAP State-Level Results Highlighting Significant Learning Acceleration*. https://www.tn.gov/education/news/2022/6/14/tennessee-releases-2021-22-tcap-state-level-results-highlighting-significant-learning-acceleration-.html

Thum, Y. M. & Kuhfeld, M. (2020, July). *NWEA 2020 MAP growth: Achievement status and growth norms for students and schools*. NWEA. https://teach.mapnwea.org/impl/normsResearchStudy.pdf

von Hippel, P. T., & Hamrock, C. (2019). Do test score gaps grow before, during, or between the school years? Measurement artifacts and what we can know in spite of them. Sociological Science, 6, 43-80.

Appendix

Table A1. Average Scale Score Trends on NWEA's MAP Growth, Curriculum Associates' i-Ready, and Renaissance Learning's Star Mathematics Assessments; Two-Year Growth Samples and National or State Norms

Grade 2020	Semester	MAP Growth		i-Ready		Star 360	
		Average	Norm	Average	MI Norm	Average	Norm
K ('21)	F21	143.2	139.6	336.5	338		
	S22	160.1	157.1	368.0	378		
K	F20	149.7	139.6	359.6	338		
	S21	162.8	157.1	380.6	378		
	F21	161.2	160.1	370.4	373		
	S22	178.1	176.4	399.8	407		
1st	F20	163.7	160.1	381.5	373	290.3	266
	S21	177.7	176.4	402.5	407	416.9	395
	F21	173.3	175.0	392.8	397	293.1	266
	S22	189.3	189.4	421.0	429	413.1	395
2nd	F20	175.6	175.0	400.3	397	405.1	417
	S21	188.6	189.4	419.4	429	518.0	506
	F21	185.7	188.5	415.2	420	405.4	417
	S22	199.5	201.1	442.0	451	518.8	506
3rd	F20	186.9	188.5	420.7	420	503.1	503
	S21	198.5	201.1	440.1	451	601.5	597
	F21	197.2	199.6	434.5	442	503.7	503
	S22	209.0	210.5	459.2	471	589.0	597
4th	F20	197.9	199.6	438.8	442	578.0	590
	S21	207.9	210.5	457.3	471	665.0	660
	F21	206.4	209.1	452.0	461	580.2	590
	S22	216.4	218.7	471.4	484	657.5	660
5th	F20	207.2	209.1	455.8	461	646.6	656
	S21	215.4	218.7	470.2	484	715.6	715
	F21	211.1	214.8	466.0	468	644.7	656
	S22	218.8	222.9	481.6	487	712.5	715
6th	F20	212.5	214.8	468.4	468	696.7	724
	S21	218.9	222.9	479.2	487	738.2	770
	F21	217.9	220.2	475.7	479	696.9	724
	S22	224.3	226.7	489.1	494	732.6	770
7th	F20	218.9	220.2	480.5	479	721.8	766
	S21	223.9	226.7	489.8	494	761.6	797
	F21	222.9	224.9	485.9	488	730.4	766
	S22	228.1	230.3	497.1	499	763.0	797
8th	F20	224.2	224.9	487.5	488	754.7	798
	S21	227.5	230.3	494.6	499	783.9	824

Note: These averages only include students with benchmark assessment scores for every possible testing period. The comparison norms represent the 50th percentile of NWEA's conditional growth distribution, median scores for Michigan i-Ready students in 2018-19, and median scores for Renaissance Learning's norming sample.

Table A2. Trends in Average Scale Scores on NWEA's MAP Growth, Curriculum Associates' i-Ready, and Renaissance Learning's Star Reading Assessments; Two-Year Growth Samples and National or State Norms

Grade 2020	Semester	MAP Growth		i-Ready		Star 360	
		Average	Norm	Average	MI Norm	Average	Norm
K ('21)	F21	139.0	136.7	340.6	337	508.0*	522*
	S22	154.9	153.1	389.5	399	688.3*	687*
K	F20	146.0	136.7	376.7	337	536.9*	522*
	S21	158.2	153.1	407.1	399	687.2*	687*
	F21	156.8	155.9	397.2	398	618.1*	641*
	S22	172.0	171.4	442.3	447	748.2*	781*
1st	F20	159.9	155.9	413.9	398	624.0*	641*
	S21	172.3	171.4	445.2	447	750.8*	781*
	F21	170.5	172.4	440.9	444		
	S22	185.0	185.6	483.3	498		
2nd	F20	173.7	172.4	453.7	444	216.7	219
	S21	185.3	185.6	484.0	498	349.8	317
	F21	185.1	186.6	479.7	488	330.4	362
	S22	196.7	197.1	513.3	525	458.6	435
3rd	F20	188.7	186.6	490.4	488	333.4	362
	S21	196.9	197.1	514.4	525	453.3	435
	F21	196.5	196.7	511.1	520	445.8	465
	S22	204.5	204.8	538.2	548	554.6	522
4th	F20	198.6	196.7	518.2	520	451.5	465
	S21	204.3	204.8	537.2	548	553.4	522
	F21	204.0	204.5	536.1	543	544.2	570
	S22	209.8	211.0	558.1	568	635.6	640
5th	F20	205.4	204.5	540.6	543	544.3	570
	S21	209.2	211.0	555.0	568	631.3	640
	F21	209.4	210.2	554.2	554	627.2	684
	S22	213.7	215.4	571.3	575	686.9	795
6th	F20	211.4	210.2	559.7	554	628.7	684
	S21	213.8	215.4	568.7	575	684.9	795
	F21	213.7	214.2	569.3	570	684.8	811
	S22	217.1	218.4	582.9	591	735.3	895
7th	F20	215.6	214.2	575.6	570	712.9	811
	S21	217.4	218.4	583.4	591	748.0	895
	F21	217.5	218.0	583.5	584	761.4	921
	S22	220.0	221.7	596.5	607	791.3	994
8th	F20	218.5	218.0	583.1	584	780.2	921
	S21	219.2	221.7	590.3	607	791.6	994

*Note: These averages only include students with benchmark assessment scores for every possible testing period. The comparison norms represent the 50th percentile of NWEA's conditional growth distribution, median scores for Michigan i-Ready students in 2018-19, and median scores for Renaissance Learning's norming sample. The * indicates Early Literacy scores or norms.*

Table A3. Regression Adjusted Standardized Scale Scores by Grade; NWEA MAP Growth, Curriculum Associates i-Ready, Renaissance Learning Star 360 Districts

	Mathematics		Reading	
	MAP Growth (1)	i-Ready (2)	MAP Growth (3)	i-Ready (4)
Grade K*Fall 2020	0.949***	1.220***	0.677***	1.292***
	(0.031)	(0.122)	(0.037)	(0.177)
Grade K*Spring 2021	0.678***	0.733***	0.488***	0.501***
	(0.027)	(0.185)	(0.023)	(0.137)
Grade K*Fall 2021	0.310***	0.375***	0.140***	0.212***
	(0.017)	(0.052)	(0.014)	(0.033)
Grade K*Spring 2022	0.309***	0.232***	0.155***	0.056*
	(0.018)	(0.055)	(0.014)	(0.024)
Grade 1*Fall 2020	0.423***	0.671***	0.216***	0.480***
	(0.023)	(0.079)	(0.024)	(0.064)
Grade 1*Spring 2021	0.289***	0.385***	0.119***	0.094*
	(0.019)	(0.077)	(0.016)	(0.041)
Grade 1*Fall 2021	0.084***	0.152***	-0.068***	-0.045+
	(0.021)	(0.025)	(0.015)	(0.025)
Grade 1*Spring 2022	0.166***	0.119***	0.057***	-0.025
	(0.021)	(0.031)	(0.017)	(0.055)
Grade 2*Fall 2020	0.187***	0.316***	0.008	0.084**
	(0.018)	(0.036)	(0.017)	(0.027)
Grade 2*Spring 2021	0.134***	0.138***	0.046***	-0.020
	(0.019)	(0.032)	(0.013)	(0.027)
Grade 2*Fall 2021	0.003	0.136***	-0.025*	-0.006
	(0.013)	(0.022)	(0.011)	(0.029)
Grade 2*Spring 2022	0.055***	0.057+	0.081***	0.025
	(0.015)	(0.032)	(0.014)	(0.042)
Grade 3*Fall 2020	0.006	0.199***	0.017+	0.055***
	(0.009)	(0.023)	(0.009)	(0.012)
Grade 3*Spring 2021	-0.006	0.060*	0.029***	0.020
	(0.010)	(0.023)	(0.008)	(0.018)
Grade 3*Fall 2021	0.030***	0.004	0.032***	0.010
	(0.009)	(0.024)	(0.008)	(0.029)
Grade 3*Spring 2022	0.056***	0.002	0.068***	0.046
	(0.012)	(0.031)	(0.010)	(0.034)
Fall 2020	-0.350***	-0.450***	-0.087***	-0.176***
	(0.017)	(0.046)	(0.014)	(0.040)
Spring 2021	-0.059**	-0.165*	-0.153***	-0.098**
	(0.019)	(0.069)	(0.010)	(0.030)
Fall 2021	0.003	-0.071	-0.089***	-0.099**
	(0.015)	(0.070)	(0.009)	(0.036)
Spring 2022	0.045**	-0.022	-0.130***	-0.076*
	(0.017)	(0.054)	(0.011)	(0.029)
Grade 5*Fall 2020	-0.025**	-0.052*	-0.061***	-0.065***
	(0.009)	(0.021)	(0.008)	(0.018)
Grade 5*Spring 2021	-0.041***	-0.002	-0.078***	-0.090***
	(0.012)	(0.033)	(0.009)	(0.025)
Grade 5*Fall 2021	-0.050***	0.096**	-0.015	0.043
	(0.010)	(0.034)	(0.009)	(0.037)

Grade 5*Spring 2022	-0.095***	0.086+	-0.030**	0.045
	(0.014)	(0.046)	(0.011)	(0.042)
Grade 6*Fall 2020	-0.068***	0.077***	-0.078***	0.026
	(0.011)	(0.022)	(0.009)	(0.024)
Grade 6*Spring 2021	-0.100***	0.142*	-0.097***	0.003
	(0.013)	(0.071)	(0.011)	(0.051)
Grade 6*Fall 2021	0.025*	0.113+	-0.018+	0.047
	(0.012)	(0.059)	(0.009)	(0.058)
Grade 6*Spring 2022	-0.013	0.145+	-0.025*	0.001
	(0.015)	(0.079)	(0.011)	(0.064)
Grade 7*Fall 2020	-0.006	0.110***	-0.075***	0.014
	(0.015)	(0.030)	(0.010)	(0.038)
Grade 7*Spring 2021	-0.026	0.228+	-0.066***	-0.015
	(0.019)	(0.114)	(0.013)	(0.082)
Grade 7*Fall 2021	0.043**	0.163*	-0.036**	0.036
	(0.016)	(0.081)	(0.011)	(0.069)
Grade 7*Spring 2022	-0.005	0.165	-0.061***	0.015
	(0.019)	(0.106)	(0.013)	(0.075)
District-grade-Level Student Characteristics	Y	Y	Y	Y
Observations	15858	1542	15754	1414
R^2	0.722	0.870	0.679	0.863

Note: These regression estimates only include students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to NWEA's and Curriculum Associates' pre-pandemic national norms, and estimates are converted to a percentile ranks.

Table A4. Achievement Gains for Each Month a District Offered In-Person Instruction in 2020-21 by Grade					
Grade	**Semester**	**Change in Average Standardized Scores**		**Change in Percentage of a Typical Year's Growth**	
		Math	**Reading**	**Math**	**Reading**
MAP Growth					
K	S21	-0.015	-0.009	-1.31	-0.56
	S22	0.010	0.014	0.92	0.95
1st	S21	0.010	0.014	0.95	1.40
	S22	0.016	0.023	1.56	2.34
2nd	S21	0.016	0.019	1.77	3.23
	S22	0.023	0.019	2.55	3.21
3rd	S21	0.029	0.021	5.51	5.88
	S22	0.027	0.013	5.14	3.61
4th	S21	0.011	0.012	2.58	3.50
	S22	0.014	0.011	2.29	1.89
5th	S21	0.027	0.020	6.47	6.12
	S22	0.011	0.007	2.74	2.16
6th	S21	0.012	0.014	3.92	5.98
	S22	0.009	0.007	3.14	2.88
7th	S21	0.002	0.006	0.74	2.47
	S22	0.000	-0.005	-0.04	-1.94
i-Ready					
K	S21	-0.003	0.002	-0.30	0.15
	S22	0.036	0.042	3.19	2.78
1st	S21	0.021	0.023	2.05	2.33
	S22	0.037	0.053	3.60	5.45
2nd	S21	0.039	0.030	4.44	5.08
	S22	0.055	0.041	6.18	6.88
3rd	S21	0.060	0.050	11.54	13.77
	S22	0.071	0.056	13.70	15.42
4th	S21	0.026	0.022	4.97	5.59
	S22	0.034	0.034	5.34	6.77
5th	S21	0.048	0.022	11.63	7.01
	S22	0.030	0.023	7.23	7.34
6th	S21	0.017	0.020	5.66	8.80
	S22	0.007	0.007	2.26	3.00
7th	S21	-0.001	0.005	-0.24	2.02
	S22	0.004	0.017	1.22	6.49

Note: These regression estimates only include students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to NWEA's and Curriculum Associates' pre-pandemic national norms, and estimates are converted to a percentile rank.

Endnotes

[1] NWEA, Curriculum Associates, and Renaissance Learning each developed their own crosswalks between benchmark assessment scores on the MAP Growth, i-Ready, and Star 360 assessments, respectively, using an equipercentile linking method (NWEA, 2020; Curriculum Associates, 2020a; Renaissance Learning, 2019). This means that they defined their crosswalks based on a sample of students who took both the M-STEP and a particular benchmark assessment and determined the benchmark assessment scale score ranges that would result in the same percentages of students scoring in each proficiency level across the two assessments (e.g., if 20% of students in this sample scored in the "advanced" level on the M-STEP, 20% would also score in the "advanced" level based on their benchmark assessment scores). This process isn't necessary for the Smarter Balanced ICA, as MDE developed the M-STEP in partnership with DRC and Smarter Balanced, and scores for both assessments are derived from the same underlying scale (Michigan Department of Education, 2019; Smarter Balanced Assessment Consortium, 2020; Smarter Balanced Assessment Consortium, 2021), allowing us to easily convert Smarter Balanced scores and M-STEP scores to each other's scales.

[2] We also conducted this analysis after limiting the analytic sample to students with valid benchmark assessment scores in every possible testing period. We found similar results. The lack of a difference in results across analyses likely confirms that the unusual results we find for students who completed a Smarter Balanced assessment are driven by differences in the groups of students who took a Smarter Balanced ICA assessment each semester.

[3] At the time of this writing, EPIC does not yet have access to 2021-22 M-STEP scores, so we are unable to compare end-of-year performance on benchmark assessments from spring 2022 to M-STEPs taken at approximately the same time.

[4] We only include the pre-pandemic national norm in these figures for the MAP Growth assessments, as these are the only assessments for which a conditional growth *distribution* (rather than just the *median* conditional growth) is available.

[5] For reference, regardless of the modality that districts offered for a majority of the year, districts consistently offered remote instruction as at least an option (on average, districts offered remote instruction for at least 94% of the 2020-21 school year). In districts that offered in-person instruction for a majority of the school year, hybrid instruction was also offered to students for approximately one quarter of the school

year (23%). Districts that offered hybrid instruction for a majority of the year also offered students the opportunity to learn in-person for approximately half of the school year (49%). Finally, districts that offered remote instruction for more than half of the school year also offered in-person instruction for approximately two-thirds of the school year (63%) and offered hybrid instruction for a third of the year.

[6] To calculate these percentages, we use benchmarks for "yearly growth" on standardized achievement tests that researchers developed based on data from seven different nationally normed assessments (Bloom et al., 2008). These are not specific to the MAP Growth or i-Ready assessments, but rather, they are general guidelines for translating standardized effect sizes to more practical terms to improve their interpretability.